BRAIN RESEARCH ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Brain Research

Review

Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses

*Leonard Varghese**, *Hari M. Bharadwaj*, *Barbara G. Shinn-Cunningham*

Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Accepted 24 June 2015

Keywords:
Frequency-following response
FFR
Envelope-following response
Attention
Auditory processing

ABSTRACT

Auditory brainstem responses (ABRs) and their steady-state counterpart (subcortical steady-state responses, SSSRs) are generally thought to be insensitive to cognitive demands. However, a handful of studies report that SSSRs are modulated depending on the subject's focus of attention, either towards or away from an auditory stimulus. Here, we explored whether attentional focus affects the envelope-following response (EFR), which is a particular kind of SSSR, and if so, whether the effects are specific to which sound elements in a sound mixture a subject is attending (selective auditory attentional modulation), specific to attended sensory input (inter-modal attentional modulation), or insensitive to attentional focus. We compared the strength of EFR-stimulus phase locking in human listeners under various tasks: listening to a monaural stimulus, selectively attending to a particular ear during dichotic stimulus presentation, and attending to visual stimuli while ignoring dichotic auditory inputs. We observed no systematic changes in the EFR across experimental manipulations, even though cortical EEG revealed attention-related modulations of alpha activity during the task. We conclude that attentional effects, if any, on human subcortical representation of sounds cannot be observed robustly using EFRs.

This article is part of a Special Issue entitled SI: Prediction and Attention.

© 2015 Published by Elsevier B.V.

Contents

*Correspondence to: Boston University, Center for Computational Neuroscience and Neural Technology, 677 Beacon Street, Boston, MA 02215, USA.
E-mail address: lennyv@bu.edu (L. Varghese).

http://dx.doi.org/10.1016/j.brainres.2015.06.038
0006-8993/© 2015 Published by Elsevier B.V.

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

1. Introduction

The ability to selectively attend to a particular talker in "cocktail party" situations depends on the fidelity of sensory encoding throughout the auditory pathway (Shinn-Cunningham and Best, 2008). In human listeners, numerous electroencephalography (EEG), magnetoencephalography (MEG), and electrocorticography (ECoG) studies over the past four decades have shown that selective attention (i.e., "selecting" and focusing on a particular sound source from among multiple sound sources; Shinn-Cunningham, 2008) modulates how sound is encoded in primary auditory cortex. The P1-N1-P2 complex in the auditory-evoked potential (AEP), which is a stereotyped response occurring approximately 100 ms after sound onset that localizes to auditory cortex (Scherg et al., 1989), is enhanced when a listener actively pays attention to the evoking sound (e.g., Hillyard et al., 1973). Conversely, this response is suppressed when sounds are ignored in dichotic listening tasks (Sussman et al., 2005; Bidet-Caulet et al., 2007; Choi et al., 2013). The auditory steady-state response (ASSR), also originating in auditory cortex, has been reported to behave similarly during selective listening (Bharadwaj et al., 2014). "Inter-modal" selective attention (e.g., paying attention to visual stimuli while ignoring simultaneously presented auditory stimuli) has also been shown to modulate the strength of cortical auditory responses. Specifically, the magnitude of the AEP (Hackley et al., 1990; Choi et al., 2013) and ASSR (ASSR; Wittekindt et al., 2014) both have been observed to increase when subjects are actively listening for an auditory stimulus compared to when they perform a visual task and are ignoring the same auditory inputs.

In contrast to the well-described effects of attention on auditory-related neuroelectrical responses originating in the cortex, it is less clear whether selective listening or inter-modal attentional shifts modulates responses originating in subcortical auditory structures. At least one previous study has suggested that attentional modulation of phase-locked neural activity may not occur at processing stages below auditory cortex (Gutschalk et al., 2008). However, there exist corticofugal projections from auditory cortex to subcortical structures that have the potential to modulate the function of lower nuclei (see Winer, 2006 for a concise review). What remains unclear is whether the actions and specific anatomical targets of these projections are specific enough to support selective attention, or even whether these projections to lower nuclei *actively* sculpt neural processing based on task demands at all.

In animals, efferent projections from auditory cortex play a role in the long-term plasticity of the neural firing properties of a number of different subcortical structures, including outer hair cells (Xiao and Suga, 2002), neurons in inferior colliculus (Yan and Suga, 1996, 1999; Bajo et al., 2010), and possibly at later subcortical processing stages as well. Similar long-term changes have also been reported in humans, seen in modifications of subcortical steady-state responses (SSSRs) obtained from the brainstem (e.g., Skoe and Kraus, 2010b). Here, we focus not on long-term plasticity, but rather on the question of whether cortical feedback shapes sub-cortical processing to aid performance "online" flexibly and in an immediate, task-dependent manner. Some awake behaving animal studies suggest that this occurs. Attention to visual stimuli was found to reduce the amplitude of transient evoked responses elicited by broadband auditory stimuli in the cochlear nucleus (Hernandez-Peon et al., 1956; Oatman, 1971, 1976) and the auditory nerve (Oatman, 1971, 1976), as well as of transient (Oatman and Anderson, 1977) and steady-state (Oatman and Anderson, 1980) responses to pure tones at a variety of frequencies in the cochlear nucleus in awake felines. Cochlear sensitivity, as measured by compound action potentials obtained from a round-window electrode

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

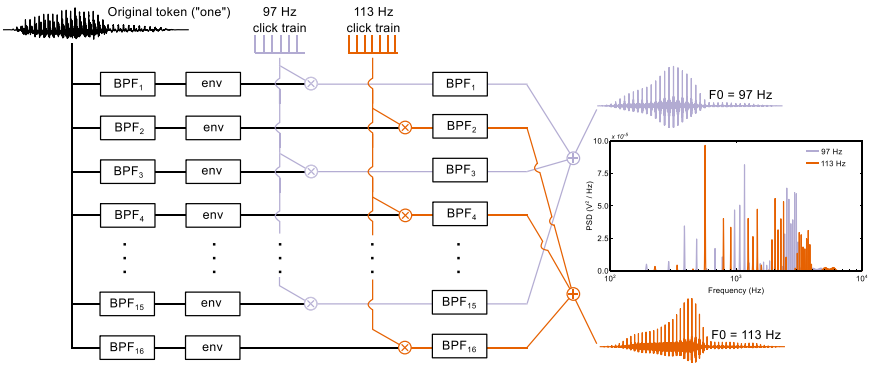

Fig. 1 – Example of vocoding procedure for the speech token "one". The original natural speech token was filtered into 16 frequency band using 1/3-octave-wide filters. The filter output from odd-numbered bands was half-wave rectified and low-pass filtered, multiplied with a click train at 97 Hz, and filtered once more using the same bandpass filter. The output from this final stage was summed to obtain the final stimulus. Construction of the 113 Hz stimulus was similar, but used even numbered bands and a click train at 113 Hz. The vocoding procedure resulted in speech streams with a fundamental frequency of 97 Hz or 113 Hz and with minimally overlapping power spectral density functions (inset, right). On each dichotic experimental trial, a digit stream with one of the two carrier frequencies was sent to one ear, and another digit stream with the other carrier frequency was presented to the other ear.

in the chinchilla, was reduced when animals deployed visual attention compared to when animals performed an auditory-only task (Delano et al., 2007). The spectrotemporal tuning of neurons in IC has also been observed to change during detection tasks relative to when animals are passively listening (Slee and David, 2015). Yet despite these results from animal studies, there are not consistent reports of attention-related modulation of subcortical neuroelectrical activity in human listeners.

Correlates of subcortical auditory function can be observed non-invasively in humans by measuring compound neuroelectrical activity at scalp locations in response to auditory stimulation. The best known among these measures is the auditory brainstem response (ABR; Sohmer and Feinmesser, 1967; Jewett and Williston, 1971; Picton et al., 1981), which is a stereotypical response occurring in the first 10 ms or so after presentation of a brief stimulus. Individual ABR peaks correspond to the neural responses from different ascending stages of the auditory pathway, from the auditory nerve (wave I) through the lemnisci and the inferior colliculus (IC, wave V; Melcher and Kiang, 1996; Scherg and von Cramon 1985). Neurons in the structures from which the ABR arises can also respond in a phase-locked manner to periodic stimulation, resulting in a measured potential known as the subcortical steady-state response (SSSR; also called the "frequency-following response"; Worden and Marsh, 1968; Skoe and Kraus, 2010a). By combining SSSRs in response to periodic stimuli presented in opposite polarities, one can separate responses that are phase-locked to the temporal fine structure (TFS) of the stimulus from those that are phase-locked to its envelope (Aiken and Picton, 2008). The response that is phase-locked to the relatively slow-varying auditory stimulus envelope specifically has been called the "amplitude-modulated following response" (AMFR; Kuwada et al., 2002) or the "envelope-following response" (EFR; Dolphin and Mountain, 1992); the latter terminology will be used throughout the rest of this paper. The strength of the EFR, measured at the fundamental frequency of the envelope and its harmonics via spectral estimation techniques, is a sum of activity from a number of different peripheral frequency auditory channels and possibly a number of different subcortical nuclei (Chandrasekaran and Kraus, 2010; Shinn-Cunningham et al., 2013; Zhu et al., 2013).

The prevailing practice within the audiology and auditory research community is to obtain ABR/SSSR recordings while subjects watch a silent movie (Skoe and Kraus, 2010a) or even while they are sedated or sleeping, with the latter method routinely used when conducting clinical ABR measurements (Burkard and McNeary, 2009; Sininger and Hyde, 2009). Such practices suggest that an individual's cognitive state does not pose a concern to auditory researchers and clinicians interested in quantifying peripheral neural activity. The literature on the modulation of ABRs by attentional focus supports such practices, since the majority of studies examining the effect of attention on the ABR in human listeners have failed to find any effects of either inter-modal attention (e.g. Connolly et al., 1989; Gregory et al., 1989; Hackley et al., 1990) or selective auditory attention (e.g., Picton and Hillyard, 1974; Woldorff et al., 1987; Gregory et al., 1989). Similarly, EFRs specifically have been found to be insensitive to sleep and even general anesthesia (Cohen et al., 1991; Lins et al., 1996; Lins and Picton, 1995).

However, examination of the literature reveals several papers that appear to show the opposite, that ABRs and SSSRs *are* sensitive to a listener's attentive state. Decreased amplitude and increased latency of ABR wave I and wave V have been reported when subjects attended to a visual stimulus compared to when there was no visual stimulus (Lukas, 1980; Brix, 1984). Several studies by Galbraith and colleagues reported that the amplitudes of SSSRs are modulated by both inter-modal attention (Galbraith et al., 2003) and selective auditory attention (Galbraith and Arroyo, 1993; Galbraith and Doan, 1995; Galbraith et al., 1998). More recently, Hairston and colleagues (2003) found that the SSSR amplitude to task-irrelevant tones decreased during an

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

auditory task, but did not change during a visual task, potentially indicating a subcortical suppression of irrelevant stimuli in challenging listening situations. Finally, Lehmann and Schönwiesner (2014) reported changes in the EFR as a result of selective attention.

Given the presence of anatomical projections capable of supporting top-down modulation of peripheral structures, as well as the reports of top-down modulation observed in animal studies, the inconsistencies in the literature describing subcortical attentional effects in human scalp-recorded potentials are puzzling and deserve further consideration. This is especially true given the observation of subcortical attentional effects using other types of functional measurements of subcortical auditory processing, including otoacoustic emissions (OAEs; Giard et al., 1994; Maison et al., 2001; Wittekindt et al., 2014) and fMRI BOLD responses in IC (Rinne et al., 2008).

Here, we conducted two experiments designed to test the hypothesis that attention modulates subcortical neuroelectrical activity. Since cortical feedback could be targeting any number of subcortical auditory processing sites, including different tonotopic channels, the likelihood of observing such feedback effects seems higher when presenting broadband periodic stimuli and quantifying neural activity via EFRs compared to other stimulus-response combinations that could be used to assay subcortical neuroelectrical activity. As such, auditory stimuli in both experiments consisted of spoken digits, and the subcortical neural responses to these stimuli were quantified using the complex-domain phase-locking value (PLV; Lachaux, et al., 1999; Bharadwaj and Shinn-Cunningham, 2014) derived from the EFR. Auditory stimuli were presented either monaurally or dichotically (one to each ear); during dichotic presentation, digits presented to each ear had asynchronous onsets and had different fundamental frequencies to allow the observed neural responses to attended and ignored digits to be resolved in time and frequency. Furthermore, in order to give selective auditory attentional feedback the "best shot" at independently modulating responses to attended and ignored digits, auditory stimuli were specifically processed such that they had non-overlapping frequency content spanning a 6 kHz frequency range (via click-train "vocoding"; see Fig. 1 and Section 4). Our overall hypothesis was that selective attention to a particular auditory stimulus stream would enhance the phase locking of the EFR to the when the stream was attended and/or diminish the phase locking of the EFR when the stream is ignored.

2. Results

2.1. *Experiment 1*

In the first experiment, subjects were asked to listen to streams of spoken digits and respond whenever two consecutive, increasing digits were heard in the attended ear

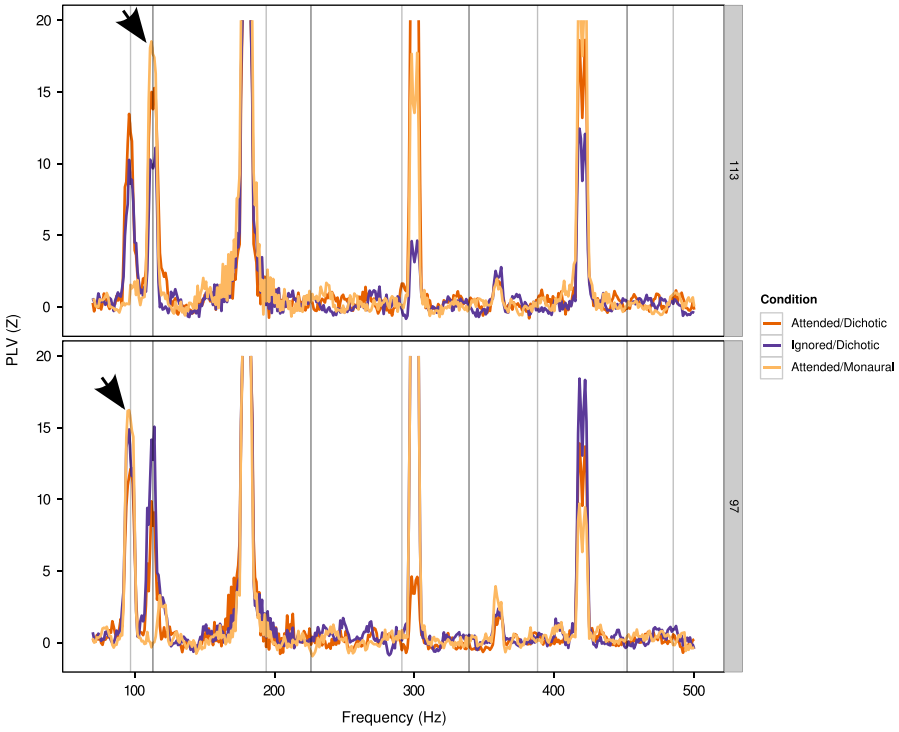

Fig. 2 – Average PLV z-score as a function of frequency in Experiment 1. Due to the high z-scores at odd multiples of 60 Hz (20 > z > 40; see text for explanation), the y-axes are truncated to the range of interest around the stimulus fundamental frequency values (see text). Hz Top row: PLV (z-score) as a function of frequency for 113 Hz stimuli. Bottom: PLV (z-score) as a function of frequency for 97 Hz stimuli. Integer multiples of 97 Hz and 113 Hz are indicated by vertical lines. Arrows represent peaks for which individual subject data is shown in Fig. 3. The opposite frequency peak (i.e., the peak at 97 Hz in the top plot and the peak at 113 Hz in the bottom plot) arise from the lack of jitter in digit onset timings; such peaks were not present in Experiment 2 (see Fig. 4 and text for details).

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

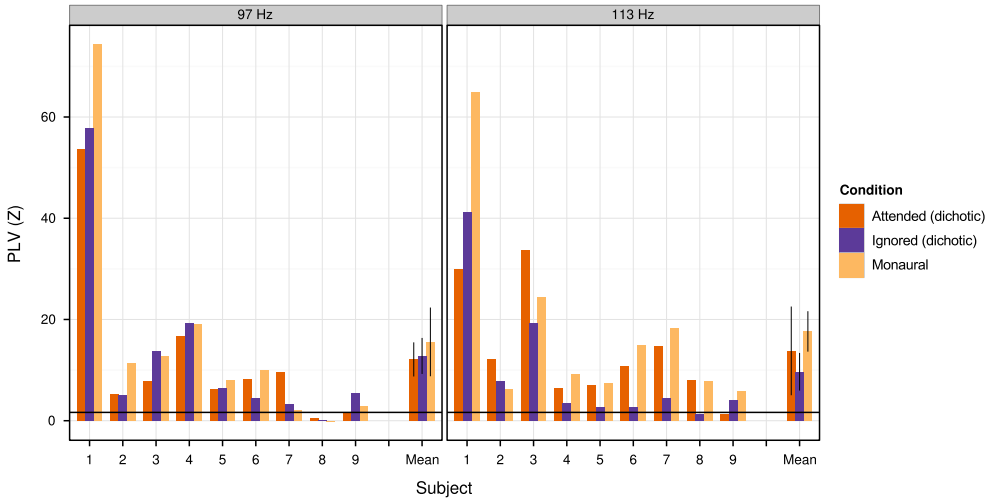

Fig. 3 – PLV z-scores at the fundamental frequency of each stimulus shown for individual subjects in Experiment 1. Left column: Data for epochs corresponding to 97 Hz stimulus onsets (PLV z-score at 97 Hz). Right column: data for epochs corresponding to 113 Hz stimulus onsets (PLV z-score at 113 Hz). The horizontal line indicates the 95th percentile of the standardized noise distribution. Error bars represent 95% confidence intervals about the mean with a within-subject correction applied (Morey, 2008; Chang, 2015).

(e.g., "3" followed by "4"); this task was performed both in quiet (i.e., during monaural presentation) or while ignoring digits presented to the other ear (i.e., during selective listening in a dichotic setting). EEG was recorded during the task for the purpose of quantifying EFR-stimulus phase locking.

2.1.1. *Behavioral results*

Analysis of behavioral performance data indicated that subjects performed the task as instructed, and they performed similarly regardless of whether they were presented with a monaural or dichotic stimulus. The mean *d'* when subjects were attending to one sound stream while ignoring another (in the attended/dichotic condition) was 3.18 (range: 2.41-4.16). The mean *d'* when subjects were presented with a single sound stream (in the attended/monaural condition) was 3.16 (1.88-4.63). This difference between conditions was not statistically significant [two-tailed, paired *t*-test; $t(8)=0.09$, $p=0.94$]. Examination of the natural logarithm of the bias parameter, $\log(\beta)$, suggested that subjects were more likely overall to categorize a digit in the attended ear as a non-target in each condition. Mean $\log(\beta)$ values were 4.47 (2.82-5.96) for the dichotic condition and 3.93 (2.17-5.29) in the monaural condition; this difference was not found to be statistically significant [two-tailed, paired t-test; $t(8)=1.82$, $p=0.11$].

2.1.2. *Phase-locking results*

Grand-averaged plots of PLV as a function of frequency are shown in Fig. 2, with PLV z-scores for individual subjects at the fundamental frequency of each stimulus shown in Fig. 3. Regardless of which stimulus was being attended (97 Hz or 113 Hz click frequency), there was significant phase locking (i.e., $z>1.64$) to the fundamental frequency of each of the two audio streams. Responses to higher harmonics were near the noise floor for almost all subjects. We also observed significant phase-locking to odd harmonics of 60 Hz due to line noise, because we did not jitter between the digit onsets within a stimulus and the regular 2 Hz repetition rate of the digits within the stream, resulting in line noise that was at the same phase at the onset of each digit token within a single trial. The regularized across-stream timing also resulted in a peak at the opposite click train frequency in epochs from dichotic trials (i.e., a peak at 97 Hz in the epochs binned at 113 Hz digit onsets, and a peak at 113 Hz in the epochs binned at 97 Hz digit onsets).

Visual comparison of z-scores across conditions in Fig. 3 reveals no consistent effects of attention on phase locking. Although phase locking to the 113 Hz stimuli tended to be higher at 113 Hz for attended conditions compared to the ignored condition (Fig. 3, right; red and yellow bars tend to be higher than blue bars), the effect was small; moreover, no similar trend was observed for the 97 Hz stimuli (Fig. 3, left; no clear ordering to red, blue, and yellow bars across subjects).

ANOVA on PLV z-scores were conducted, with experimental condition (attend during dichotic presentation, ignore during dichotic presentation, "attend" during monaural presentation), stimulus carrier frequency (i.e., click rate; 97 Hz or 113 Hz), and the interaction of the two factors treated as fixed effects. Subject effects were included in the model as random effects. None of the fixed effects were found to be significant [experimental condition: $F(2, 16)=2.76$, $p=0.09$, generalized effect size $\eta^2_G=1.90\text{E}-2$ (see Section 4); stimulus frequency: $F(1, 8)=4.38\times10^{-3}$, $p=0.95$, $\eta^2_G=4.22\times10^{-5}$; interaction: $F(2, 16)=2.62$, $p=.10$, $\eta^2_G=5.37\text{E}-3$].

Given that our experimental manipulations had no statistically significant effect on PLVs, we sought to quantify the likelihood that the null hypothesis is true, given our results (as opposed to simply failing to reject the null hypothesis). PLV z-scores were subject to a Bayes Factor analysis (Kass and Raftery, 1995; Rouder et al., 2012). Given our data, a linear model that excludes experimental condition is 2.98 times as

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

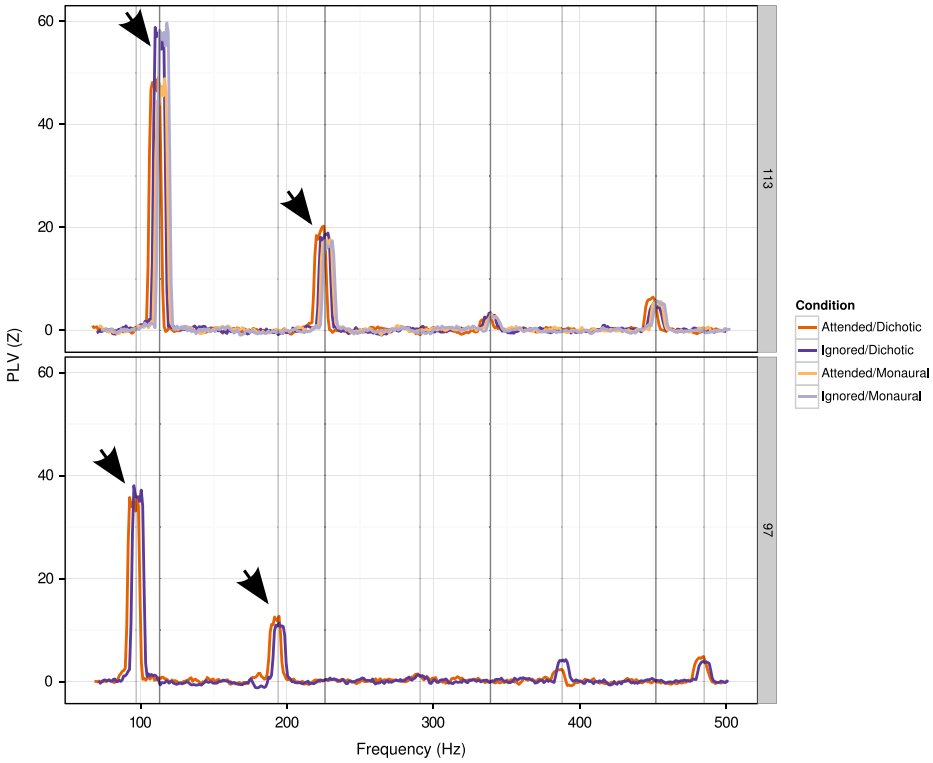

Fig. 4 – Average PLV z-score as a function of frequency in Experiment 2. Traces have been shifted slightly in the horizontal direction to aid visualization. Top row: PLV (z-score) as a function of frequency for 113 Hz stimuli. Bottom: PLV (z-score) as a function of frequency for 97 Hz stimuli. Integer multiples of 97 Hz and 113 Hz are indicated by vertical lines. Arrows represent peaks for which individual subject data is shown in Fig. 5.

likely as one taking into account both experimental condition and the interaction between experimental condition and stimulus frequency. Based on published guidelines (Kass and Raftery, 1995), our data provides at best slight evidence in favor of the null hypothesis; i.e., if a priori, an effect of task on the EFR strength is equally likely as no effect of task, the no-effect hypothesis is about 3 times as likely to be true as the hypothesis that PLVs are affected by experimental manipulation.

2.2. *Experiment 2*

Results of Experiment 1 suggest that task conditions did not affect phase-locking in the EFR. Experiment 2 was conducted to test the same set of hypotheses regarding auditory attention, but with digit onset timings randomized within a single stream. This resulted in the randomization of relative onset timings of digits across ears when stimuli were presented dichotically. Furthermore, given that previous reports of modality-specific attentional modulation of subcortical responses are more consistent than reports of modulation due to selective, within-modality attention shifts (e.g., Lukas, 1980; Oatman and Anderson, 1980; Lukas, 1981; Bauer and Bayles, 1990; Galbraith et al., 2003; Delano et al., 2007; Wittekindt et al., 2014), we also introduced an additional condition in which participants attended to visual stimuli while ignoring monaurally presented auditory stimuli. The detection task employed in the second experiment was similar to the task utilized in Experiment 1 (detect two consecutive, increasing digits in the attended ear or, for the visual task, on a computer monitor), but with a small financial bonus contingent on performance levels to encourage participants' maximum engagement with the task. We hypothesized that switching attention to a different sensory modality altogether would decrease the strength of EFR phase locking to auditory inputs, even if within-modality shifts of attention did not result in similar changes (as suggested by the results of Experiment 1).

2.2.1. *Behavioral task performance*

The mean *d'* values were 3.10 (range: 2.62-3.86), 3.12 (2.04–4.62), and 3.27 (1.97–3.84) for the selective listening, monaural, and attend-visual tasks, respectively, indicating that all subjects were able to perform the task reasonably well. The difference in performance between conditions was not statistically significant [$F(2, 22)=0.67$, $p=0.52$, $\eta_G^2=0.02$]. Mean log(β) values were 4.52 (3.43–6.51), 4.42 (2.86–5.87), and 3.85 (1.77–5.80) for the selective listening, monaural, and attend-visual tasks, indicating that subjects had an overall bias towards classifying digits in the attended ear as non-targets. The difference between conditions was not statistically significant after applying the Greenhouse-Geisser correction for violation of sphericity [$F(2, 22)=3.72$, Greenhuose-Geisser corrected $p=0.07$, $\eta_G^2=0.09$].

2.2.2. *Phase-locking results*

Grand-averaged plots of PLV as a function of frequency are shown in Fig. 4, with PLV z-scores for individual subjects at

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

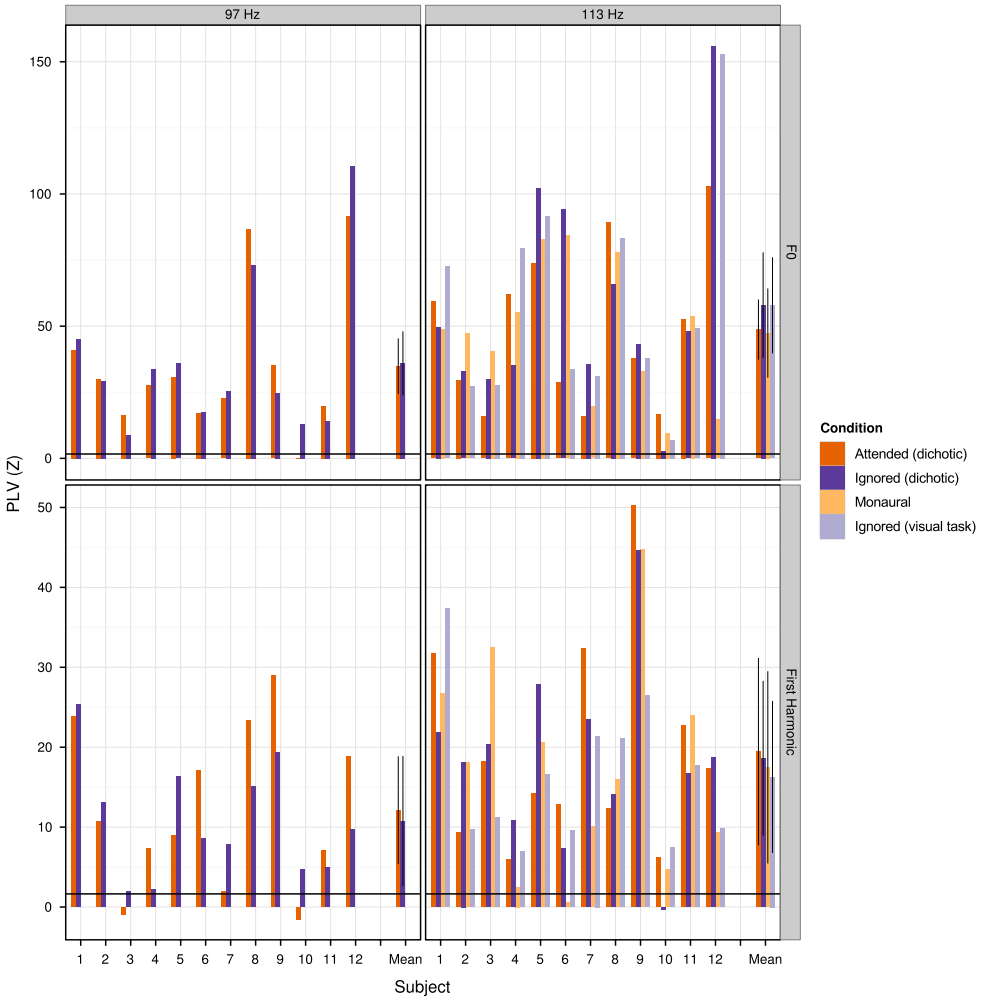

Fig. 5 – PLV z-scores at the fundamental frequency of each stimulus shown for individual subjects in Experiment 2. Top row: Z-scores at stimulus fundamental frequency. Bottom row: Z-scores at first harmonic. Left column: Data for epochs corresponding to 97 Hz stimulus onsets (PLV z-score at 97 Hz or 194 Hz). Right column: data for epochs corresponding to 113 Hz stimulus onsets (PLV z-score at 113 Hz or 226 Hz). Horizontal lines in each panel indicate the 95th percentile of the standardized noise distribution. Error bars represent 95% confidence intervals about the mean with a within-subject correction applied (Morey, 2008; Chang, 2015).

the fundamental frequency and first harmonic frequency of each stimulus shown in Fig. 5. In contrast to Experiment 1, subjects exhibited significant phase-locked activity at the fundamental frequency of both streams (113 Hz or 97 Hz), as well as at integer multiples of these fundamentals (Fig. 4, peaks for each PLV curve align at the frequencies indicated by darker vertical lines). Additionally, there was no cross-frequency peak for the selective attention conditions; i.e. there was no significant phase-locking to the 113 Hz stimuli in the 97 Hz epochs, and vice versa (compare Fig. 2 to Fig. 4). The absence of the peak at the opposing stimulus frequency as well as the more robust responses at harmonics of the fundamental is most likely due to the introduction of the onset timing jitter within each digit stream, which has the effect of reducing phase locking to the fundamental frequency of the competing stream and to the line noise, since the phases of these stimulus components are not aligned across the analysis epochs.

Table 1 – Pairwise *t*-test results for 97 Hz vs. 113 Hz comparison for PLV data from Experiment 2. All *p*-values shown after Bonferroni-Holm correction have been applied. Bold indicates statistical significance using α=0.95. D=Dichotic, M=Monaural, A=Attend, I=Ignore.

	97 Hz D/I	113 Hz D/A	113 Hz D/I	113 Hz M/A	113 Hz V/I
97 Hz D/A	1.000	**0.018**	0.066	1.000	**0.034**
97 Hz D/I		**0.011**	**0.032**	1.000	**0.007**
113 Hz D/A			1.000	1.000	1.000
113 Hz D/I				1.000	1.000
113 Hz M/A					1.000

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

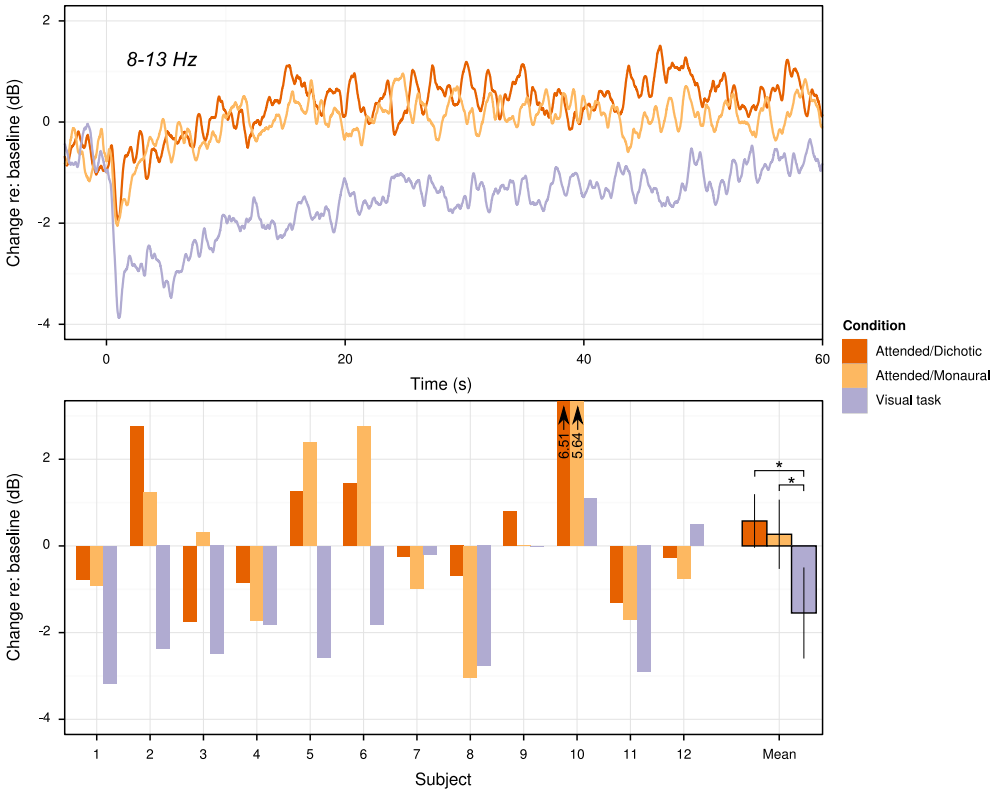

Fig. 6 – Summary of parieto-occipital alpha activity obtained from Experiment 2. Top: mean alpha amplitude across subjects over time, expressed in dB relative to the amplitude in the pretrial period (−3.5 to 0 s), shown for the pretrial period and the first 60 s of the trials. Bottom: change in alpha amplitude during stimulus presentation, expressed as dB relative to baseline amplitude, shown for individual subjects and for the group mean. Error bars represent 95% confidence intervals about the mean with a within-subject correction applied (Morey, 2008; Chang, 2015). Stars indicate significant pairwise differences (α=0.95) obtained via t-test after Bonferroni-Holm corrections were applied to p values.

When comparing across individual subjects, there did not appear to be a consistent effect of attentional condition on PLV strength (in Fig. 5, there is no consistent ordering of red, blue, yellow, or gray bars across subjects for the 97 Hz stimuli or in the ordering of red and blue bars across subjects for the 113 Hz stimuli). The mean PLV z-scores for the 113 Hz stimuli at F0 tend to be higher in the "ignore" cases than the "attend" cases (Fig. 5, upper right; blue and gray bars are higher than red and yellow bars). However, this trend is driven almost entirely by the results of subject 12, who showed very large responses in the "ignored" trials; this tendency was not otherwise present across subjects.

Due to the imbalance of stimulus frequency and attentional condition (four attentional conditions for 113 Hz stimuli, and two attentional conditions for 97 Hz stimuli), we chose to treat stimulus frequency and attentional condition as a single fixed effect ("experimental condition") for the purpose of conducting ANOVA. Stimulus harmonic (F0 versus second harmonic) was treated as an additional fixed effect. Subject was treated as a random effect. ANOVA indicated that the interaction between experimental effect and harmonic was not significant [$F(5,55)=1.53$, $p=0.20$, $\eta^2_G=0.02$]. The effect of stimulus harmonic was found to be significant [$F(1, 11)=16.11$, $p<0.01$, $\eta^2_G=0.31$], confirming that phase-locking at F0 is greater than phase-locking at the first harmonic. The effect of experimental condition was also found to be significant [$F(5, 55)=4.09$, Greenhouse-Geisser corrected $p=0.02$, $\eta^2_G=0.06$]. However, post-hoc, pair-wise t-tests with Bonferroni-Holm correction indicated that the only significant differences occurred when comparing data from 97 Hz stimulus presentations against data from 113 Hz stimulus presentations (Table 1). This outcome can be interpreted as PLVs to 113 Hz stimuli being greater than PLVs to 97 Hz stimuli, which is evident visually by comparing left and right panels in Fig. 5. We note that a given natural speech token vocoded with a 113 Hz click train has more F0 cycles than the same token vocoded with a 97 Hz click train. This inherent difference in the physical stimuli may contributeto the difference in absolute PLV strength at these two frequencies.

We computed Bayes factors on this dataset to quantify evidence in favor of the null hypothesis, using the same fixed and random effect structure used in the ANOVA. Results indicated that a model that does not include experimental condition is 6.05 times more likely than a model that includes both experimental condition and the interaction between harmonic and experimental condition, indicating "positive" evidence against the latter (Kass and Raftery, 1995). When this procedure was repeated on the data from just the four attentional conditions for 113 Hz stimulus presentation, the model without attentional condition was found to be 70.2 times more likely than the model including attentional condition and its interaction with harmonic. This constitutes

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

"strong" evidence against our experimental manipulations having an effect on the EFR.

2.2.3. *Post-hoc cortical data analysis*

In the absence of any subcortical attentional effects in Experiments 2, we turned to cortical responses (which dominate the low-frequency portions of the recorded signals) to see if there was evidence that subjects were directing their attentional focus appropriately. In conducting this post-hoc analysis, we noted some issues that complicate the interpretation of auditory-evoked cortical activity in our paradigm. First, over the course of each long trial, adaptation is likely to diminish the amplitude of canonical AEP components (Woods and Elmasian, 1986). Second, each individual token is sufficiently different in its spectrotemporal content that the latency and morphology of auditory-evoked peaks may differ substantially (Hansen et al., 1983; Sharma et al., 2000; Frye et al., 2007). Indeed, cursory inspection of cortical-evoked AEP components time-locked to both the epoch start as well as the approximate onset of voicing in the speech stimuli exhibited features consistent with traditional AEP components in some, but not all subjects; even when typical ERP components were present, their latencies and amplitudes varied from token to token and from subject to subject. Some recent studies have shown that the EEG signal correlates with the speech energy as a function of time; indeed, in mixtures of two speech waveforms, it is possible to use the strength of the correlation between the cortical response and the envelope of each of the competing speech waveforms to decode which stream a listener is attending (e.g., Lalor et al., 2009; Kerlin et al., 2010). Unfortunately, broadband stimulus envelopes and onset times of our two competing streams were not independent of each other (indeed, they were roughly alternating) and relatively regular compared to natural, continuous speech, so such analysis was not suitable for our data set.

Given that our paradigm was ill suited for evoked potential analysis, we investigated ongoing alpha band (8–13 Hz) oscillatory activity in Experiment 2 as a marker of attention. A number of recent studies associate alpha band activity with sensory suppression in a variety of tasks, including selective attention tasks (Foxe and Snyder, 2011). One recent study reported that during an inter-modal selective attention task using auditory and visual stimuli, alpha activity was reduced over frontal and parieto-occipital EEG sensor regions when subjects attended to visual stimuli compared to when they attended to audio stimuli (Wittekindt et al., 2014). Increased alpha activity over central and parietal sensors has also been associated with task difficulty on auditory-only tasks (Obleser et al., 2012; Obleser and Weisz, 2012).

Based on previous literature on alpha band activity and on our own experience with EEG, we chose to focus on alpha activity within the parieto-occipital sensor region. We hypothesized that the magnitude of alpha activity in these sensors would be reduced when participants performed the visual task compared to when they performed the auditory task. We additionally hypothesized that alpha band activity would be greater in the dichotic, auditory selective attention condition, where selective attention requires suppression of the competing stream, than in the monaural, auditory-only condition.

The time course of parieto-occipital sensor alpha activity averaged over subjects is shown in the upper panel of Fig. 6, with the mean difference between baseline alpha activity and alpha activity during the stimulus presentation period shown in the bottom panel for each subject. Inspection of the time course indicates a sharp drop in alpha activity in the first second after stimulus onset, with the activity during the visual task dropping most sharply. Activity in the auditory conditions returns to near- or above-baseline levels within about 5 s; in contrast, activity during the visual task remains below baseline over the entire course of the time period analyzed.

The mean change between baseline alpha amplitude and alpha amplitude during performance of the task was +0.58 dB for the attended/dichotic condition, +0.27 dB for the attended/monaural condition, and −1.55 dB for the attend-visual condition. One-way, repeated measures ANOVA on the alpha amplitude change revealed a significant difference in the magnitude of change relative to baseline across conditions [$F(2, 22)=9.01$, $p=0.01$, $\eta_G^2=0.18$]. Pairwise, paired two-tailed t-tests with Bonferroni-Holm correction indicated that the alpha activity in the attend-visual condition was significantly lower than when subjects selectively attended to a stream in the dichotic condition ($p=0.01$), as well as when they attended to the monaurally presented auditory stimuli ($p=0.04$). The difference in the change in alpha amplitude between the dichotic and auditory-only monaural conditions was not significant ($p=0.41$).

3. Discussion

Selective attention involves selecting a particular source of information while suppressing competing sources (e.g., Shinn-Cunningham, 2008). It is well established that selective attention modulates cortical responses to sound, so that responses to the attended source are strong compared to responses to an ignored sound (e.g., Fritz et al., 2007). Yet it is not clear whether attention modulates responses via corticofugal pathways in a way that can be consistently observed from non-invasive measures of subcortical responses, specifically in EFRs. Even if *selective* auditory attention does not lead to a relative enhancement of EFRs elicited by attended over unattended sources in a sound mixture, it may be that EFRs are modified by whether a subject is attending to sound vs. attending to some other sensory modality.

Here, in two separate experiments, subjects either selectively listened to one sound stream or selectively attended to visual stimuli while ignoring auditory inputs. We observed no systematic variation in EFR PLV when listeners selectively attended vs. ignored a sound stream in a mixture of competing streams, even though we designed our auditory stimuli such that the two streams were separated in frequency content, pitch, and ear of presentation. Furthermore, the lack of systematic variation in EFR PLV with task demands, or even presentation scheme (i.e., monaural vs. dichotic), cannot be due to poor overall SNR, as overall levels of phase

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

locking were well above the noise floor at the fundamental frequency of each digit stream in both experiments.

3.1. *Experiment 1*

For almost all subjects and conditions, there was significant phase locking to the fundamental frequency of the stimuli. However, we observed no consistent changes in the strength of phase locking that could be attributed to attentional focus. A Bayes factor analysis indicated that our data favored the null hypothesis by a factor of about 3.

While clear phase-locking peaks were found at the fundamental frequencies of 97 Hz and 113 Hz, we did not observe consistent, strong phase locking at the harmonics of these fundamental frequencies. When PLVs are used to analyze phase locking to the stimulus envelope, clear peaks can often be seen at the first few harmonics of the fundamental frequency of the stimulus, regardless of whether PLVs were obtained from a single electrode (e.g., Ruggles et al., 2012; Zhu et al., 2013), or derived from multiple electrodes (Bharadwaj and Shinn-Cunningham, 2014); however, in general, the strength of these peaks is idiosyncratic across subjects. The weak phase locking at stimulus frequency harmonics observed here may have been due to interactions between the responses evoked by the two streams with different fundamental frequencies. The magnitude of SSSR harmonics increases with presentation level (Krishnan, 2002). Moreover, higher harmonics of SSSRs at higher stimulus harmonics are more sensitive to SNR than the SSSR at the fundamental frequency (Zhu et al., 2013). Given that our listeners set the presentation level to a comfortable listening level, presentation level may also have influenced the amount of phase locking observed at the higher harmonics compared to what has been found in previous studies.

At the fundamental frequencies of the two competing streams, the PLV was significantly above the noise floor in almost all subjects and conditions, yet we found no consistent effects on the PLV that could be attributed to selective listening. Behavioral data ($d' > 3$) suggests that our inability to find an effect is not due to the possibility that subjects were passively listening, rather than engaging in the task, during the experiment. One potential explanation for the null result could be that listeners were rapidly switching their attention back and forth between streams in dichotic trials rather than selectively listening to one stream, thereby cancelling out any expected neural effects of selective listening. The time to switch the focus of attention from one sound source to another is thought to be in the range of 200–300 ms (Shinn-Cunningham, 2008). Given that the digits in alternating streams were 250 ms apart, it is possible that subjects adopted such a strategy. However, not only would such a strategy interfere with rather than help performing the task, sustaining this kind of listening over a one-minute long experimental trial would be quite taxing, making it unlikely that subjects adopted this strategy. Our pilot testing prior to the start of the experiment indicated that the least effortful way to perform the task was to direct attention to the signal presented in the to-be-attended ear and sustain focus on the target stream throughout each trial.

3.2. *Experiment 2*

There was significant phase locking to the fundamental frequency of the stimuli for all subjects in all conditions, as well as at the first harmonic frequency of the stimulus for most subjects and conditions. In spite of these gains, which could be attributed to better overall SNR relative to Experiment 1, we once again observed no consistent changes in the strength of phase locking that could be attributed to attentional focus, with a Bayes factor analysis indicating that our data provided "positive" to "strong" evidence (Kass and Raftery, 1995) in favor of the null hypothesis.

Behavioral performance on the dichotic and attend-monaural conditions was similar to what was observed in Experiment 1. However, it is difficult to directly compare task performance across the two experimental sessions. The financial incentive reward for correct responses may have pushed listeners to try harder on the task, but any resulting performance gains might have been offset by the fact that the digit onsets were temporally unpredictable due to the timing jitter introduced, which may have made the task harder.

We observed strong phase locking of the EFR to the stimulus fundamental frequency and at multiple harmonics across subjects, and overall, PLV z-scores were higher in Experiment 2 than in Experiment 1. We believe that the additional cautions taken in the design and presentation of the stimulus and the analyses helped increase the strength of the phase locking that we observed. Specifically, we fixed the RMS presentation level in Experiment 2, rather than allowing subjects to control the stimulus level (thus raising the possibility that they set it too low to observe reliable phase-locking at the harmonics). Presentation levels were also normalized on a per-token basis, compared to a per-stream basis as in Experiment 1. There may also have been less interference from the responses to the competing stream due to the introduction of temporal jitter in the digit onsets. Finally, equating the number of epochs corresponding to each digit and polarity in PLV computation, as opposed to simply equating number of epochs per polarity in Experiment 1, may also have led to more robust phase locking estimates across conditions, averaging out differences in phase-locking caused by differences in the temporal envelope of different digits.

Examination of the alpha activity indicated that alpha activity remained lower than baseline alpha activity in the attend-visual trials, while returning to near- or above- baseline levels when subjects attended to auditory stimuli. These observations lend support to the argument that our failure to see changes in EFR strength across conditions is not due to the possibility that participants simply failed to direct attention or follow instructions. Neural oscillations in the alpha band have been associated with suppression of sensory inputs (Obleser et al., 2012; Wittekindt et al., 2014). Reduced parieto-occipital alpha activity during attention to a visual stimulus has been reported in numerous studies on visual information processing (e.g., Foxe et al., 1998) as well as during attention to visual stimuli during auditory stimulation (Wittekindt et al., 2014). Consistent with these reports, we observed greater cortical parieto-occipital alpha activity during selective listening compared to when listeners performed a visual task while ignoring auditory stimuli.

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

We had hypothesized that alpha activity would be higher during the attended/dichotic condition due to the additional cognitive load associated with selective listening. However, we did not observe a significant difference in alpha activity between attend/monaural and attend/dichotic conditions. Previous reports associate the strength of alpha activity with task difficulty (Obleser and Weisz, 2012), and may index the direction of spatial attention during selective listening tasks (Kerlin et al., 2010). As yet, it is unclear as to whether such lateralized effects during auditory-only processing are due to anticipatory effects (e.g., from a pre-stimulus visual cue) or whether changes in alpha activity are directly related to where attention is directed in a sound scene (see Strauss et al., 2014). Our paradigm was not designed to examine this particular question; therefore, we did not try to localize activity to neural sources in order to determine whether lateralized alpha changes were present during auditory-only conditions.

In summary, even though EFR phase locking was stronger in Experiment 2, the qualitative outcome was identical to Experiment 1: EFR-derived PLVs were insensitive to task demands. Specifically, PLVs were unaffected by whether attention was directed to one stream versus another, whether attention was focused on a visual task rather than an auditory task, or even whether stimuli were presented dichotically versus monaurally. Based on the behavioral results, post-hoc analysis of cortical alpha activity, and informal reports from our subjects both during the breaks and after the experimental sessions were complete, we conclude that our failure to find attentional effects on the EFR is not due to simple, uninteresting possibility that participants failed to follow instructions and direct their focus of attention appropriately.

3.3. *Reconciling our negative findings with previous studies*

Several previous studies have reported that attention does not affect the scalp-recorded ABR (Picton and Hillyard, 1974; Woldorff et al., 1987; Gregory et al., 1989; Connolly et al., 1989; Hackley et al., 1990). As the ABR and EFR are both generated by brainstem auditory structures, it is perhaps unsurprising that we found no significant effects of attention on the EFR. However, ABR and SSSR responses (including the EFR) tend to emphasize different features of neural responses. The ABR measures the transient response to the onset of a sound (typically a broadband click; Sohmer and Feinmesser, 1967; Jewett and Williston, 1971; Picton et al., 1981), with each peak representing neural activity from a particular structure along the auditory pathway (Melcher and Kiang, 1996). In contrast, SSSRs measure responses to ongoing neural activity that is phase-locked to the periodic structure of an input stimulus, reflecting steady-state activity that likely originates from multiple brainstem structures. SSSRs will not reflect modulation of neural populations that do not respond in a phase-locked manner to the input stimulus (e.g., Palombi et al., 2001, as cited in Carcagno et al., 2014).

While some earlier studies performed in humans have reported that there are no robust effects of attention on SSSRs (Hillyard and Picton, 1979; Galbraith and Kane, 1993), other studies have reported the opposite (Galbraith and Arroyo, 1993; Galbraith and Doan, 1995; Galbraith et al., 1998, 2003; Hairston et al., 2013; Lehmann and Schönwiesner, 2014). Yet on close examination, several of these later reports suffer from issues that make it difficult to conclude that the reported effects are due to task-dependent shifts in the focus of attention or cognitive state.

Lehmann and Schönwiesner (2014) reported selective attention effects that are idiosyncratic across subjects and are non-specific in their direction: for some subjects, SSSR strength was greater when attention was directed towards the corresponding stimulus, while for others, it decreased. Despite discussing SSSRs in general terms, the authors seem to focus specifically on the EFR in their study, since their analysis combines responses to alternate-polarity stimuli when computing spectral measurements (Lehmann and Schönwiesner, 2014, pg. 3). Further, the bootstrap statistics used to assess attention effects at the individual level compared between-block effects of attention to an estimated variance of the EFR that depended primarily on *within* block variability (Lehmann and Schönwiesner, 2014, Fig. 4). Taken as a whole, these results seem to support the conclusion that the EFR is *not* sensitive to attention (despite the title of the report), in agreement with the present study and other recent studies that failed to find attentional effects on the EFR (Dai and Shinn-Cunningham, 2014; Ruggles et al., 2014).

Among the studies conducted by Galbraith and colleagues on the modulation of SSSRs by attention, one reported no effect (Galbraith and Kane, 1993), while there was no main effect of attention in any of the studies claiming attentional modulation effects (Galbraith and Arroyo, 1993; Galbraith and Doan, 1995; Galbraith et al., 1998, 2003). In those studies claiming a positive effect of attention, conclusions are based on complex patterns of interactions between stimulus, analysis window, and other factors such as the recording montage. We suspect that the inconsistent findings in these studies are due in part to insufficient SNR to allow for an accurate computation of spectral quantities. For example, in Galbraith and Arroyo (1993), attention was reported to affect the SSSR differently, depending on both which stimulus frequency was considered and the period over which the SSSR was analyzed. Specifically, attention enhanced the first half of the 25 ms-long response to the 400 Hz SSSR and not its second half, yet also reported that attention enhanced the second half of the 200 Hz SSSR and not its first half. We suggest that 25 ms or 12.5 ms analysis windows may be too short to reliably assess the presence or absence of attentional effects without a very large number of stimulus repetitions, especially for pure tone stimuli that excite a limited number of auditory channels and yield concomitantly weak responses at the scalp (Ananthanarayan and Durrant, 1992). Similarly, the SSSR results presented in Galbraith and Doan (1995) are based on 500 trials per condition, with approximately 43 ms of neural data from a single electrode analyzed per trial. The pattern of results, in which SSSR amplitudes in one of the conditions (intensity discrimination of a pure tone) showed a large and opposite trend compared to SSSR amplitudes in the other conditions, may reflect a genuine physiological effect. However, such a pattern of results could be entirely due to insufficient SNR (and false positives by chance).

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

Most studies that have reported an effect of attention on SSSRs (all except those of Galbraith et al., 2003), did not consider the spectral noise floor when computing spectral power, and whether it differed across subjects, conditions, or stimulus frequency (see Zhu et al., 2013). While a response metric incorporating the spectral noise floor was considered by Galbraith et al. (2003), the overall magnitude of change due to the attention effect was not reported. Because the signals driving the SSSR are quite small compared to measurement noise, all SSSR measures, especially magnitude measures, depend on background noise levels to some degree. Moreover, since cortical activity is arguably the largest source of noise at most frequencies, changes in listener state could cause changes in cortical function, which could translate into changes (or increased variability) in observed SSSR strength. Changes in the noise due to overall subject state could be particularly problematic in listening tasks requiring a sustained focus of attention over relatively long periods of time (e.g., a lapse of listener arousal could affect a large set of trials from one particular condition, biasing results). Indeed, the deviant detection tasks used in most previous SSSR studies, as well as the detection task used in the present study, may be particularly prone to such issues due to the relatively long trial durations typically used in these experiments.

This possibility, that scalp-recorded measures have too low an SNR to reliably reveal subcortical attentional effects, is supported by reports of subcortical effects of attention in animal models, as well as reports of attention-related subcortical activity modulations in humans using techniques with better spatial selectivity than SSSRs/EFRs. In cochlear nucleus responses, attention to visual stimuli decreases both transient activity (Hernandez-Peon, et al., 1956; Oatman, 1971, 1976; Oatman and Anderson, 1977) and the amplitude of the steady-state response (Oatman and Anderson, 1980). In chinchillas, the magnitude of the compound action potential decreases the cochlear microphonic amplitude increases during periods of visual attention (Delano et al., 2007). In the IC of awake behaving ferrets, there is task-related modulation of neural ressponses (Slee and David, 2015). Modulation of BOLD responses localized to IC during selective attention has also been reported in human listeners (Rinne et al., 2008). Subcortical attentional effects have also been noted at the level of the cochlea in humans. Stimulus-frequency OAE recordings collected during selective listening (Giard et al., 1994; Maison et al., 2001) and distortion-product OAE recordings collected in an inter-modal attention task (Wittekindt et al., 2014) also suggest that attention modulates responses in outer hair cells.

Based on the results of our study and others that have failed to find modulatory effects on the EFR, the observed effects on cochlear responses in humans may not be inherited by the subsequent neural structures that contribute to the EFR. Alternatively, such propagative effects may be small and subtle enough that they are masked by noise. Indeed, deep neural cellular currents must flow with high synchrony in an appropriate direction in order to produce voltages large-enough to be measureable on the scalp (Okada et al., 1997). Given past reports, it may be that attentional mechanisms target specific subcortical structures or particular neuronal populations rather than operate as a blanket enhancement or suppression affecting all levels and channels of subcortical auditory processing. Such modulation may not robustly affect the EFR, even if brainstem and midbrain responses are modulated by descending efferent signals.

3.4. *Caveats and conclusions*

In computing PLVs, we combined responses to different vocoded digits with the same carrier signal, but different envelopes. Our assumption was that the primary drivers of EFR phase-locking are the harmonic peaks in each frequency band that arise from the click train vocoding procedure. However, the amount of energy in each vocoder band, and thus the energy driving each peripheral auditory channel, will differ across digits of the same carrier frequency due to the different temporal envelopes of the digits. We attempted to minimize the effects of this source of noise by equating the overall RMS energy per digit and the number of epochs per digit considered in the PLV calculation in Experiment 2, yet we still observed no consistent effect of attention on our results. In addition to supporting the hypothesis that subcortical attentional effects are small and perhaps localized, this also implies that it would be difficult to observe consistent attentional effects in EFR responses to more ecologically valid stimuli, such as natural sounds or speech.

SSSRs (including EFRs) are increasingly used within the auditory research community as tools to measure subcortical auditory coding fidelity (e.g., Picton et al. 2003; Purcell et al., 2004; Bharadwaj et al., 2015). Somewhat ironically, they are useful as a tool for assessing sensory coding precision *because* they are relatively insensitive to a listener's state of arousal. In addition to the present data, other recent studies (Dai and Shinn-Cunningham, 2014; Lehmann and Schönwiesner, 2014; Ruggles et al., 2014) and previous reports (Cohen et al., 1991; Lins et al., 1996; Lins and Picton, 1995) support this idea. Even if the structures generating the ABR or EFR are truly altered by attention or cognitive state, our results indicate these changes are likely of sufficiently small magnitude that for all practical purposes, their effect on scalp recordings of EFRs can be ignored. Our results reinforce the notion that more sensitive and specific assays of subcortical function are necessary to better understand the contributions of online corticofugal modulation to everyday hearing abilities in human listeners.

4. Experimental procedures

4.1. *Experiment 1*

Ten subjects (4 males, 6 females, ages 18–28) recruited from the Boston University community participated in the experiment. Each subject was screened via audiometer to confirm normal hearing thresholds at standard audiometric frequencies between 250 Hz and 8 kHz (defined as $<$20 dB hearing level in either ear, or a 20 dB asymmetry across ears at any tested frequency). Subjects signed informed consent forms approved by the Boston University Charles River Campus IRB and were compensated $20/h for their participation. Each subject participated in three or four two-hour-long experimental sessions,

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

completing at most one session per day. An intermittent headphone connection problem led us to exclude data from one subject, as we could not reliably determine which of that subject's sessions were affected by the problem with audio presentation. All other subjects had at least 2 sessions where we were confident in the reliability of the headphone connection.

4.1.1. *Stimuli*

The spoken digits 1–5 were recorded by a male member of the laboratory using an Audio Technica (Tokyo, Japan) AT4033 microphone and digitized with 16 bit precision at a sampling rate of 44,100 Hz using an Apogee (Santa Monica, CA, USA) Duet audio interface with a 40 Hz high pass filter applied. The speech tokens were then cropped using Digital Performer 7 software (MOTU; Cambridge, MA, USA).

All stimulus processing took place in Matlab (Mathworks; Natick, MA, USA). A 50 ms cosine squared onset-offset ramp was applied to each of the original speech tokens. All tokens were downsampled to 24,414 Hz, then filtered with a first-order shelf filter (+3 dB at approximately 600 Hz, +13.5 dB at the Nyquist frequency) to boost the high-frequency content of the signal. The original speech was channel vocoded as illustrated in Fig. 1 using a regular click train carrier with fundamental frequency of either 97 or 113 Hz, one used for one stream, and the other for the competing stream. Narrowband envelope signals derived from the original speech tokens (16 in total) were imposed on click train, replacing the temporal fine structure of the speech with a steady, monotone fundamental frequency while preserving the intelligibility of the speech tokens (Shannon et al., 1995). By design, the two fundamental frequencies used for the two streams were separated enough that the EFR of each stream could be isolated in the frequency domain. To further maximize the separation of the peripheral auditory representations of the two competing streams, each token was synthesized using only half the vocoder bands (odd-numbered bands for 97 Hz stimuli, even-numbered bands for 113 Hz stimuli; see Fig. 1). With this design, when there were two competing streams (on dichotic trials), the streams were presented to different ears and also had complementary frequency bands, thus minimizing their spectral overlap and optimizing the potential to observe frequency-specific attentional modulation effects.

4.1.2. *Digit stream construction*

To create each dichotic stimulus (the "attended/dichotic" condition), one carrier click frequency was selected as the "attend" frequency and assigned to either the left or the right channel, and the other was designated the "distractor" frequency and assigned to the opposite channel; this was done randomly on a trial-by-trial basis. For the attended channel, the 120 digits composing the stream were selected pseudo-randomly from the set of 5 digits vocoded with the carrier frequency to be attended. In the sequence of 120 tokens in the attended stream, any pair of temporally abutting consecutive digits (e.g., 1–2, or 2–3, etc.) was designated a target. Constraints were imposed such that (1) there were 4 to 6 targets in each stream, (2) target start positions were separated by at least 10 other digits, and (3) the target start position never occurred in either the first three or the last three positions in the stream. The distractor digit stream was constructed similarly; 120 tokens were drawn from the set of 5 digits at the to-be-ignored carrier frequency, with the only constraint being that no consecutive, increasing digits could occur in the sequence. All digits within a single stream had a 500 ms inter-onset interval, and the to-be-ignored stream always started 250 ms after the to-be-attend stream, which resulted in the listeners hearing digits alternating in each ear every 250 ms. (Unfortunately, this construction of constant inter-token intervals meant that the 60 Hz power line noise was in the same phase for each of the 120 tokens in each of the two streams, leading to artifacts in the measurements at multiples of 60 Hz for this experiment – an issue that was corrected in Experiment 2.) An additional set of monaural stimuli was constructed in a similar fashion, except that the distractor stream was not included ("attended/monaural"). On half of all trials, the output stimulus was inverted in polarity to allow us to compute the subcortical response to the envelope rather than the response arising from the combination of phase-locked activity to the envelope and fine structure of the stimulus (Aiken and Picton, 2008; Skoe and Kraus, 2010a).

These manipulations (2 frequencies × 2 attended locations × 2 polarities for both dichotic and monaural stimuli) were counterbalanced across all experimental conditions for each session. Within an experimental session, each trial type was repeated twice, for a total of 32 trials per session.

4.1.3. *Stimulus presentation and task details*

Experimental flow was controlled in Matlab with the PsychToolbox 3 extension (Brainard, 1997; Kleinerand Brainard, 2007) installed to present on-screen instructions. Matlab was interfaced with Tucker-Davis Technologies (Alachua, FL, USA) System 3 hardware for D/A conversion and playback over Etymotic (Elk Grove Village, IL, USA) ER-1 insert earphones. Listeners were seated in a sound-attenuating booth (Eckel; Cambridge, MA, USA) for the duration of the experiment. The audio presentation was set to a comfortable listening level for each subject (in the range of 66–73 dB SPL, computed using a 1 kHz pure tone with the same RMS level) within each session.

On each trial, a small cross was presented in the middle of the screen, and subjects were instructed to fixate their gaze on it without blinking excessively. Prior to the start of sound playback, a small arrow was drawn on screen to indicate the side (left or right) to which the subject should direct his or her attention. After 1 s, the arrow disappeared and audio playback commenced. Subjects pressed a button on a response box whenever they detected a target sequence (two sequential numbers in a row in the target stream). Subjects were scored as having missed a target if they did not respond within 2 s of the occurrence of the second digit in a target sequence; responses outside the two-second window were labeled as false alarms, with a corresponding false alarm rate computed as the ratio of false responses to the total number of non-target stimuli presented. Once playback was complete, the cross in the center of the screen turned red, and listeners were able to rest before beginning the next trial at their

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

discretion. No feedback was given to subjects during the task. Each trial lasted a little longer than a minute, and subjects were given a short break after every 8 trials.

Subject performance was quantified using signal detection theory. We considered d' and β as a measure of sensitivity and response bias, respectively (Macmillan and Creelman, 2005; Abdi, 2007). In these computations, hit rate was defined as the ratio of correct detections to the total number of target digit presentations within a condition, and false alarm rate was defined as the ratio of the number of responses outside a target window to the total number of non-target digit presentations presented at the attended ear within a condition. Extreme proportions of 0 or 1 were corrected by recomputing the ratios as $0.5/N$ or $1–0.5/N$, respectively, where N represents the number of trials used to compute the rates (Macmillan and Kaplan, 1985; Hautus, 1995).

4.1.4. *EEG recording and EFR analysis procedures*

EEG data were collected from subjects during the task using a Biosemi (Amsterdam, Netherlands) ActiveTwo system sampling at 16,384 Hz. Responses were collected from 32 scalp electrodes. Data from a pair of reference electrodes, one affixed to each earlobe, was collected and used later in offline analysis. One additional channel recorded event markers indicating the time of each digit onset and subject response button presses. Recordings were monitored online to ensure that subjects were not closing their eyes during audio playback.

Initial assessment of signal quality

The signal quality of the channels was assessed visually by applying a 70–1,800 Hz first-order Butterworth filter using the EDFBrowser software package (http://www.teuniz.net/edfbrowser/). Any channel that was observed to have an excessive number of noisy data points (e.g., motion artifacts, muscle fiber activity, or artifacts resulting from poor electrode contact, all of which are visually distinguishable from the electrical signals of interest) was recorded and excluded from all subsequent analyses.

Phase-locking computations

After exclusion of channels identified as artifact-contaminated, the raw signals from the scalp and reference electrodes were filtered with a 70–1800 Hz FIR filter and downsampled to 4096 Hz, then time shifted to compensate for the group delay imposed by the bandpass filtering. The scalp data were then re-referenced to the average of the signal at the two earlobe electrodes. Data were split into 500-ms long epochs aligned to the starting point of the digit. These epochs were grouped according to whether (1) they corresponded to a digit from the attended stream, the ignored stream, or a monaural stream, (2) the stream had a carrier frequency (i.e., click rate) of 97 Hz or a carrier frequency of 113 Hz, and (3) the stimulus was presented in positive or negative polarity. We note that while our method of splitting the data resulted in responses from both the attended and the ignored digits being present in each epoch collected during dichotic stimulation, the responses to each digit should have distinct and dissociable frequency components, as the two digits had distinct, resolvable fundamental frequencies. Finally, trials showing deflections greater than 65 μV in any channel were marked as artifact-contaminated and discarded.

EFR strength was quantified using the phase-locking value (PLV; Lachaux et al., 1999; Zhu et al., 2013). As the vocoding procedure resulted in tokens with identical envelope phase relationships relative to one another, no token-specific adjustments of the response phase was necessary in the PLV computation. The SNR of PLV estimates were improved by utilizing the multi-taper and complex-valued eigenvalue decomposition method described in Bharadwaj and Shinn-Cunningham (2014), as implemented in the "ANLffr" version 0.1.0 software package (http://github.com/haribharadwaj/ANLffr). Briefly, the method involves multiplying the data from each epoch with a discrete prolate spheroidal sequence window (DPSS; also known as a Slepian sequence; Slepian, 1978), deriving a cross-spectral density matrix across channels from the data, and then performing an eigenvalue decomposition on this matrix to obtain the squared PLV (p^2) with spectral information derived from multiple channels. This procedure is then repeated on the data with different (orthogonal) tapers applied to the time domain signal. The final result is obtained by averaging over the estimates obtained from each taper. The first three DPSS tapers corresponding to a time half-bandwidth parameter of 2 were used for this analysis. A 4096-point FFT was utilized in computing the cross-spectral densities, resulting in a frequency axis sampled at 1 Hz.

The variance and bias of PLV estimates derived using the complex PCA method depends on the number of electrodes utilized in the computations (Bharadwaj and Shinn-Cunningham, 2014). Because all subjects had clean signals from a minimum of 20 electrodes, we fixed the number of electrodes utilized in the phase-locking calculations by choosing 14 electrode sites that all subjects had in common, and then adding additional electrodes at random from the remaining good set of electrodes for that subject until the total number reached 20.

The bias and variance of PLV estimates in general depends on the number of trials utilized for computation (Bokil et al., 2007; Zhu et al., 2013). As we were interested in within-subject effects of attention (i.e., the relative strength of the PLV across different attentional states for a given subject), we elected to analyze the same number of electrodes and trials across the conditions of interest within a subject. For each subject, this number was chosen as the minimum number of trials available across attention conditions that allowed for an equal number of positive and negative polarity trials to be included in each PLV calculation. As each subject performed the test over multiple sessions; there may have been differences in cap positioning, background noise levels, or overall subject state across sessions that make it difficult to combine data across days. To address this confound, we chose to perform PLV computations on data from a single recording session per subject. The particular session chosen for analysis was the session that maximized the number of artifact-free trials available for analysis. Overall, these selection procedures resulted in PLVs being computed on a minimum of 572 trials/condition (286/polarity) for the noisiest dataset, a maximum of 952 trials/condition (476/polarity) for the dataset

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

with the fewest number of artifacts. Bootstrapped estimates of PLVs in the frequency range 70–600 Hz were obtained by averaging the results of 240 PLV computations on resampled datasets obtained by sampling M trials with replacement from among the M trials available per polarity.

For each subject, condition, and frequency bin in this range, we empirically derived the noise floor distribution (i.e., the expected PLV using our analysis procedures in the absence of any actual phase-locked neural activity; the null model) by reversing the phase of half of the trials and computing the resulting PLV. We note that for PLVs computed on single-channel data, the noise floor for stationary noise is theoretically independent across frequency bins and depends only on the number of trials available for analysis; practically speaking, this theoretical floor is close to what we observe, suggesting that non-stationarities in the measured noise do not strongly impact results (Zhu et al., 2013). When using the frequency-domain eigenvalue decomposition to obtain PLVs from multiple channels, however, this assumption may not be correct at lower (cortical) frequencies if there is significant spatial correlation of activity between electrode sites. In the frequency range of interest in this study (above 70 Hz), these spatial correlations should be minimal. As we observed no significant differences in the noise floor across conditions, we estimated the noise floor in each frequency bin in the 70–600 Hz range over all conditions for each subject s, and constructed a subject-specific noise distribution by computing the mean $\mu_{s,\text{noise}}$ and variance $\sigma^2_{s,\text{noice}}$ of the middle 95% of these values.

The noise mean and variance were utilized to convert the bootstrapped estimates of phase locking from the original, phase-intact data to a z-score z for each condition c, which was then used for all further analyses:

$$z_{s,c}(f) = \frac{P^2_{s,c}(f) - \mu_{s,\text{noise}}}{\sigma_{s,\text{noise}}}$$

Using this scaling method, a peak may be considered "significantly above the noise floor" if a z-score is above 1.64, which corresponds to the 95th percentile of the standard normal distribution.

Statistical analyses of z-scores were conducted in R (R Core Team, 2015). Repeated-measures ANOVA was conducted using the "ezANOVA" package (Lawrence, 2015). Effect sizes for ANOVA factors were quantified using generalized eta-squared, η^2_G, a measure of effect size suitable for comparisons across different experimental designs (Olejnik and Algina, 2003; Bakeman, 2005). When the assumption of equal variance across conditions was violated, a Greenhouse-Geisser correction to the degrees of freedom of the reference F-distribution was applied to compute the p-values from the F statistics (Keselman et al., 2001).

PLV z-scores were subject to a Bayes Factor analysis (Kass and Raftery, 1995; Rouder et al., 2012) using the "BayesFactor" R package (Morey et al., 2015). For this analysis, a default, non-informative, Cauchy prior distribution was imposed on the standardized effect size (Rouder et al., 2012).

4.2. *Experiment 2*

Experiment 2 was similar to Experiment 1, except for some minor changes in stimulus construction and the addition of an "ignore-auditory" condition. These differences are described below.

Thirteen subjects (3 males, 10 females, age 20–29) recruited from the Boston University student population participated in Experiment 2; none had participated in Experiment 1. All subjects were native speakers of American English, and were screened to ensure that they had detection thresholds < 20 dB hearing loss in both ears for frequencies between 250 Hz and 8 kHz. Subjects signed informed consent documents approved by the Boston University Charles River Campus IRB, and were compensated $25/hour for their participation. Recordings were acquired in a single experimental session. The total recording time per subject was approximately 1.5 h, including breaks. Data from one subject was excluded due to unusually high PLVs (z-scores > 700), which suggests that some electromechanical artifact had compromised the measured signals.

4.2.1. *Stimulus presentation and task details*

Stimuli were generated as in Experiment 1 with the exception of the following key differences: (1) each stream contained 140, rather than 120 digits; (2) vocoded digits were equated in RMS energy on a per-token basis, rather than on a per-trial basis; (3) the digit onsets within each stream were jittered in time by up to 100 ms, but with the constraint that no onset in one stream could occur within 25 ms of another onset in the opposing stream; (4) there were always exactly 4 target digits in the attended stream, rather than a random number between 4 and 6; and (5) audio stimuli were presented at a fixed level across subjects (identical RMS to a 1 kHz pure tone presented at 75 dB SPL), rather than set to a "comfortable" level determined subjectively by the subject; (6) positive and negative polarity stimuli were included within the same trial, rather than all stimuli having the same polarity within a trial.

There were three conditions in Experiment 2: (1) "attend" to a monaural digit stream ("monaural"); (2) selectively attend to one stream in a dichotic mixture comprising one stream at 113 Hz and one stream at 97 Hz ("selective attention"); and (3) attend to a visual digit stream during presentation of a monaural digit stream ("visual task"). In the visual task, a digit stream was presented at the center of the computer screen; the visual stream was similar to the auditory streams in its timing and content. These visual trials were included solely to determine whether responses evoked by auditory stimuli are altered when subjects focus attention on a non-auditory stimulus. As such, visual stimulus parameters (e.g., visual angle, luminosity) were not controlled. In all cases, which ear received the stream to attend and which received the stream to ignore was counterbalanced across trials.

To ensure maximum engagement with the task and to incentivize selective listening on dichotic trials, subjects were paid a small financial bonus of about $0.05 for each correct target detection and penalized about $0.03 for each false alarm. These financial bonuses did not affect subjects' base compensation, and were capped at a maximum of $10.00 in the event of perfect performance.

Stimuli were presented like those in Experiment 1. At the beginning of each trial, a green arrow pointing either left or right (for attend-monaural and attend-dichotic trials), or the message "on-screen" (for attend-visual trials) appeared at the center of the

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

computer screen to indicate the task type. A simultaneously presented red arrow indicated that subjects should ignore any sounds originating from that direction. Subjects pressed a button to acknowledge the instructions, after which the arrows were replaced with a single gray fixation cross at the center of the screen. Stimulus presentation then commenced. Participants completed the same task as in Experiment 1, pressing a button whenever they heard two consecutive, increasing digits when attending to an audio stream, or whenever they saw two consecutive, increasing digits presented on screen during the attend-visual trials. Participants were instructed to keep their eyes open for the duration of the trial without blinking excessively, and to do their best to keep their gaze on the center of the screen for the duration of the trial. The fixation cross turning red signaled the end of a trial. Subjects' responses were scored as in Experiment 1, but they were given feedback after each trial in the form of number of correct responses, number of false presses, and their running "bonus" total (described above). Participants were given a short break after every 12 trials.

4.2.2. *EEG recording and EFR analysis procedures*

EEG was collected from subjects using a Biosemi ActiveTwo system sampling at 4096 Hz while they performed the task. Subcortical EFR data analyses procedures were similar to those in Experiment 1, with the following differences: (1) 23 electrodes were used in the computation of PLVs (15 electrode sites in common and 8 additional electrodes selected per subject), reflecting the number of "clean" electrodes available in the noisiest subject; (2) the number of trials utilized in the PLV calculations was fixed based on the minimum number of trials per data pool available when datasets were was broken down by attention condition, frequency, digit, side of presentation, and polarity for a given subject. Across subjects, this procedure resulted in a minimum of 720 trials (36/digit/side/polarity) and a maximum of 1060 trials (53/digit/side/polarity) utilized for PLV calculations across the different subjects. Resampled datasets of equal size were constructed by selecting an equal number of trials with replacement from each data pool (digit/side/polarity) per subject and condition; PLVs were computed from these resampled pools. As in Experiment 1, the bootstrapping procedure was repeated 240 times to obtain final estimates of stimulus induced PLV as well as the noise floor distribution.

Statistical analysis was performed in R using the same procedures previously described for Experiment 1.

4.2.3. *Cortical alpha band analysis*

Recordings from the 12 subjects for whom EFR PLVs were computed were used in this analysis. For these subjects, we considered the 36 trials in which they either attended to a 113 Hz stimulus in the dichotic condition, attended to a 113 Hz stimulus when presented monaurally, or ignored the 113 Hz stimulus while performing the visual task.

Data from electrode locations CP1, CP2, CP5, CP6, P3, Pz, P4, PO3, PO4, O1, Oz, and O2 (in 10–20 notation) were considered for this analysis. Channels from this set were visually inspected after the application of a 1–40 Hz 1st order Butterworth bandpass filter. This resulted in three channels from a single subject being removed from further analyses (electrode locations Oz, O1, and O2 in Subject 12, all of which appeared to be making poor contact with the scalp).

Alpha activity was isolated from the raw signals by applying an 8–13 Hz bandpass FIR filter on the raw data and referencing the data against the average of the signal at the two earlobes; the filtered signals were then downsampled to 64 Hz. The envelope of this signal was extracted by full-wave rectifying the signal and low-pass filtering with a 31 tap FIR moving average filter (a Gaussian window in the time domain). The constant group delay introduced with the application of each FIR filter was compensated by time shifting after filtering. The amplitude of this envelope was converted to decibels (dB) relative to each subject's mean alpha activity in the 3.5 s prior to the start of stimulus presentation. When all processing was complete, approximately 64.2 s of data (including the 3.5 s in the baseline) were considered per trial.

ANOVAs and t-tests for these data were conducted in R.

Acknowledgments

This work supported by an NSSEFF grant to Barbara G. Shinn Cunningham, as well as by CELEST, a National Science Foundation Science of Learning Center (NSF SMA-0835976). The authors wish to thank Scott Bressler for his assistance in designing the stimuli used in this experiment. The authors would also like to thank Lorraine Delhorne for performing hearing screenings on participants in this study. We thank two anonymous reviewers for their comments and suggestions on an earlier version of this manuscript.

REFERENCES

Abdi, H., 2007. Signal detection theory (SDT). In: Salkind, N. (Ed.), Encyclopedia of Measurement and Statistics. SAGE, Thousand Oaks, CA, pp. 886–889.

Aiken, S.J., Picton, T.W., 2008. Envelope and spectral frequency-following responses to vowel sounds. Hear. Res. 245, 35–47, http://dx.doi.org/10.1016/j.heares.2008.08.004.

Ananthanarayan, A.K., Durrant, J.D., 1992. The frequency-following response and the onset response: evaluation of frequency specificity using a forward-masking paradigm. Ear Hear. 13, 228–232, http://dx.doi.org/10.1097/00003446-199208000-00003.

Bajo, V.M., Nodal, F.R., Moore, D.R., King, A.J., 2010. The descending corticocollicular pathway mediates learning-induced auditory plasticity. Nat. Neurosci. 13, 253–260, http://dx.doi.org/10.1038/nn.2466.

Bakeman, R., 2005. Recommended effect size statistics for repeated measures designs. Behav. Res. Methods 37, 379–384, http://dx.doi.org/10.3758/bf03192707.

Bauer, L.O., Bayles, R.L., 1990. Precortical filtering and selective attention: an evoked potential analysis. Biol. Psychol. 30, 21–33, http://dx.doi.org/10.1016/0301-0511(90)90088-e.

Bharadwaj, H., Shinn-Cunningham, B., 2014. Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. Clin. Neurophysiol. 125, 1878–1888, http://dx.doi.org/10.1016/j.clinph.2014.01.011.

Bharadwaj, H.M., Lee, A.K.C., Shinn-Cunningham, B.G., 2014. Measuring auditory selective attention using frequency tagging. Front. Integr. Neurosci. 8, 6, http://dx.doi.org/10.3389/fnint.2014.00006.

Bharadwaj, H.M., Masud, S., Mehraei, G., Verhulst, S., Shinn-Cunningham, B.G., 2015. Individual differences reveal correlates of hidden hearing deficits. J. Neurosci. 35, 2161–2172, http://dx.doi.org/10.1523/jneurosci.3915-14.2015.

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P.E., Giard, M.H., Bertrand, O., 2007. Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. J. Neurosci. 27, 9252–9261, http://dx.doi.org/10.1523/jneurosci.1402-07.2007.

Bokil, H., Purpura, K., Schoffelen, J.-M., Thomson, D., Mitra, P., 2007. Comparing spectra and coherences for groups of unequal size. J. Neurosci. Methods 159, 337–345, http://dx.doi.org/10.1016/j.jneumeth.2006.07.011.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436, http://dx.doi.org/10.1163/156856897x00357.

Brix, R., 1984. The influence of attention on the auditory brain stem evoked responses. preliminary report. Acta Otolaryngol. 98, 89–92, http://dx.doi.org/10.3109/00016488409107538.

Burkard, R.F., McNeary, K., 2009. Introduction to auditory evoked potentials. In: Katz, J., Medwetsky, L., Burkard, R., Hood, L. (Eds.), Handbook of Clinical Audiology 6th ed. Lippincott Williams & Wilkins, Baltimore, MD, pp. 222–241.

Carcagno, S., Plack, C.J., Portron, A., Semal, C., Demany, L., 2014. The auditory enhancement effect is not reflected in the 80-Hz auditory steady-state response. J. Assoc. Res. Otolaryngol. 15, 621–630, http://dx.doi.org/10.1007/s10162-014-0455-y.

Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47, 236–246, http://dx.doi.org/10.1111/j.1469-8986.2009.00928.x.

Chang, W., 2015. Plotting Means and Error Bars (ggplot2). Cookbook for R. URL: ⟨http://www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2)/⟩ (accessed 10.06.15).

Choi, I., Rajaram, S., Varghese, L.A., Shinn-Cunningham, B.G., 2013. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. Front. Hum. Neurosci. 7, 115, http://dx.doi.org/10.3389/fnhum.2013.00115.

Cohen, L.T., Rickards, F.W., Clark, G.M., 1991. A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. J. Acoust. Soc. Am. 90, 2467–2479, http://dx.doi.org/10.1121/1.402050.

Connolly, J.F., Aubry, K., McGillivary, N., Scott, D.W., 1989. Human brainstem auditory evoked potentials fail to provide evidence of efferent modulation of auditory input during attentional tasks. Psychophysiology 26, 292–303, http://dx.doi.org/10.1111/j.1469-8986.1989.tb01920.x.

Dai, L., Shinn-Cunningham, B., 2014. Simultaneous measurement of cortical responses, sub-cortical responses, and behavior performance in an auditory attention task. Assoc. Res. Otolaryngol. Abs. 37, 34–35.

Delano, P.H., Elgueda, D., Hamame, C.M., Robles, L., 2007. Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas. J. Neurosci. 27, 4146–4153, http://dx.doi.org/10.1523/jneurosci.3702-06.2007.

Dolphin, W.F., Mountain, D.C., 1992. The envelope following response: scalp potentials elicited in the mongolian gerbil using sinusoidally am acoustic signals. Hear. Res. 58, 70–78, http://dx.doi.org/10.1016/0378-5955(92)90010-k.

Foxe, J.J., Simpson, G.V., Ahlfors, S.P., 1998. Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. Neuroreport 9, 3929–3933, http://dx.doi.org/10.1097/00001756-199812010-00030.

Foxe, J.J., Snyder, A.C., 2011. The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. Front. Psychol. 2, 154, http://dx.doi.org/10.3389/fpsyg.2011.00154.

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A., 2007. Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1?. Hear. Res. 229, 186–203, http://dx.doi.org/10.1016/j.heares.2007.01.009.

Frye, R.E., Fisher, J.M., Coty, A., Zarella, M., Liederman, J., Halgren, E., 2007. Linear coding of voice onset time. J. Cogn. Neurosci. 19, 1476–1487, http://dx.doi.org/10.1162/jocn.2007.19.9.1476.

Galbraith, G.C., Arroyo, C., 1993. Selective attention and brainstem frequency-following responses. Biol. Psychol. 37, 3–22, http://dx.doi.org/10.1016/0301-0511(93)90024-3.

Galbraith, G.C., Bhuta, S.M., Choate, A.K., Kitahara, J.M., Mullen, T. A., 1998. Brain stem frequency-following response to dichotic vowels during attention. Neuroreport 9, 1889–1893, http://dx.doi.org/10.1097/00001756-199806010-00041.

Galbraith, G.C., Doan, B.Q., 1995. Brainstem frequency-following and behavioral responses during selective attention to pure tone and missing fundamental stimuli. Int. J. Psychophysiol. 19, 203–214, http://dx.doi.org/10.1016/0167-8760(95)00008-g.

Galbraith, G.C., Kane, J.M., 1993. Brainstem frequency-following responses and cortical event-related potentials during attention. Percept. Mot. Skills 76, 1231–1241, http://dx.doi.org/10.2466/pms.1993.76.3c.1231.

Galbraith, G.C., Olfman, D.M., Huffman, T.M., 2003. Selective attention affects human brain stem frequency-following response. Neuroreport 14, 735–738, http://dx.doi.org/10.1097/01.wnr.0000064983.96259.49.

Giard, M.H., Collet, L., Bouchet, P., Pernier, J., 1994. Auditory selective attention in the human cochlea. Brain Res. 633, 353–356, http://dx.doi.org/10.1016/0006-8993(94)91561-x.

Gregory, S.D., Heath, J.A., Rosenberg, M.E., 1989. Does selective attention influence the brain-stem auditory evoked potential?. Electroencephalogr. Clin. Neurophysiol. 73, 557–560, http://dx.doi.org/10.1016/0013-4694(89)90266-6.

Gutschalk, A., Micheyl, C., Oxenham, A.J., 2008. Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol. 6, e138, http://dx.doi.org/10.1371/journal.pbio.0060138.

Hackley, S.A., Woldorff, M., Hillyard, S.A., 1990. Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. Psychophysiology 27, 195–208, http://dx.doi.org/10.1111/j.1469-8986.1990.tb00370.x.

Hairston, W.D., Letowski, T.R., McDowell, K., 2013. Task-related suppression of the brainstem frequency following response. PLoS One 8, e55215, http://dx.doi.org/10.1371/journal.pone.0055215.

Hansen, J.C., Dickstein, P.W., Berka, C., Hillyard, S.A., 1983. Event-related potentials during selective attention to speech sounds. Biol. Psychol. 16, 211–224, http://dx.doi.org/10.1016/0301-0511(83)90025-x.

Hautus, M.J., 1995. Corrections for extreme proportions and their biasing effects on estimated values of d'. Behav. Res. Methods Instrum. Comput. 27, 46–51, http://dx.doi.org/10.3758/bf03203619.

Hernandez-Peon, R., Scherrer, H., Jouvet, M., 1956. Modification of electric activity in cochlear nucleus during attention in unanesthetized cats. Science 123, 331–332, http://dx.doi.org/10.1126/science.123.3191.331.

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. Science 182, 177–180, http://dx.doi.org/10.1126/science.182.4108.177.

Hillyard, S.A., Picton, T.W., 1979. Event-related brain potentials and selective information processing in man. In: Desmedt, J.E. (Ed.), Cognitive Components in Cerebral Event-related Potentials and Selective Attention: Progress in Clinical Neurophysiology, vol. 6. Karger, Basel, pp. 1–52.

Jewett, D.L., Williston, J.S., 1971. Auditory-evoked far fields averaged from the scalp of humans. Brain 94, 681–696, http://dx.doi.org/10.1093/brain/94.4.681.

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795, http://dx.doi.org/10.2307/2291091.

Kerlin, J.R., Shahin, A.J., Miller, L.M., 2010. Attentional gain control of ongoing cortical speech representations in a "cocktail party. J. Neurosci. 30, 620–628, http://dx.doi.org/10.1523/jneurosci.3631-09.2010.

Keselman, H.J., Algina, J., Kowalchuk, R.K., 2001. The analysis of repeated measures designs: a review. Br. J. Math. Stat. Psychol. 54, 1–20, http://dx.doi.org/10.1348/000711001159357.

Kleiner M, P.D., Brainard, D., 2007. What's new in Psychtoolbox-3?. Perception 36 ECVP Abs. Suppl., 14.

Krishnan, A., 2002. Human frequency-following responses: representation of steady-state synthetic vowels. Hear. Res. 166, 192–201, http://dx.doi.org/10.1016/s0378-5955(02)00327-1.

Kuwada, S., Anderson, J.S., Batra, R., Fitzpatrick, D.C., Teissier, N., D'Angelo, W.R., 2002. Sources of the scalp-recorded amplitude-modulation following response. J. Am. Acad. Audiol. 13, 188–204.

Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208, http://dx.doi.org/10.1002/(sici)1097-0193(1999)8:4<194::aid-hbm4>3.0.co;2-c.

Lalor, E.C., Power, A.J., Reilly, R.B., Foxe, J.J., 2009. Resolving precise temporal processing properties of the auditory system using continuous stimuli. J. Neurophysiol. 102, 349–359, http://dx.doi.org/10.1152/jn.90896.2008.

Lawrence, M.A., 2015. Package 'ez'. URL ⟨http://cran.r-project.org/web/packages/ez/ez.pdf⟩ (accessed 10.06.15).

Lehmann, A., Schönwiesner, M., 2014. Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. PLoS One 9, e85442, http://dx.doi.org/10.1371/journal.pone.0085442.

Lins, O.G., Picton, T.W., 1995. Auditory steady-state responses to multiple simultaneous stimuli. Electroencephalogr. Clin. Neurophysiol. 96, 420–432, http://dx.doi.org/10.1016/0168-5597(95)00048-w.

Lins, O.G., Picton, T.W., Boucher, B.L., Durieux-Smith, A., Champagne, S.C., Moran, L.M., Perez-Abalo, M.C., Martin, C.V., Savio, G., 1996. Frequency-specific audiometry using steady-state responses. Ear Hear. 17, 81–96, http://dx.doi.org/10.1097/00003446-199604000-00001.

Lukas, J.H., 1980. Human auditory attention: the olivocochlear bundle may function as a peripheral filter. Psychophysiology 17, 444–452, http://dx.doi.org/10.1111/j.1469-8986.1980.tb00181.x.

Lukas, J.H., 1981. The role of efferent inhibition in human auditory attention: an examination of the auditory brainstem potentials. Int. J. Neurosci. 12, 137–145, http://dx.doi.org/10.3109/00207458108985796.

Macmillan, N.A., Creelman, D.C., 2005. In: Detection Theory: A User's Guide 2nd ed. Lawrence Erlbaum Associates, Mahwah, NJ.

Macmillan, N.A., Kaplan, H.L., 1985. Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. Psychol. Bull. 98, 185–199, http://dx.doi.org/10.1037/0033-2909.98.1.185.

Maison, S., Micheyl, C., Collet, L., 2001. Influence of focused auditory attention on cochlear activity in humans. Psychophysiology 38, 35–40, http://dx.doi.org/10.1111/1469-8986.3810035.

Melcher, J.R., Kiang, N.Y., 1996. Generators of the brainstem auditory evoked potential in cat. III: Identified cell populations. Hear. Res. 93, 52–71, http://dx.doi.org/10.1016/0378-5955(95)00200-6.

Morey, R.D., 2008. Confidence intervals from normalized data: a correction to Cousineau (2005). Tutor Quant. Methods Psychol. 4, 61–64.

Morey, R.D., Rouder, J.N., Jamil, T., 2015. Package 'BayesFactor'. URL ⟨http://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf⟩ (accessed 10.06.15).

Oatman, L.C., 1971. Role of visual attention on auditory evoked potentials in unanesthetized cats. Exp. Neurol. 32, 341–356, http://dx.doi.org/10.1016/0014-4886(71)90003-3.

Oatman, L.C., 1976. Effects of visual attention on the intensity of auditory evoked potentials. Exp. Neurol. 51, 41–53, http://dx.doi.org/10.1016/0014-4886(76)90052-2.

Oatman, L.C., Anderson, B.W., 1977. Effects of visual attention on tone burst evoked auditory potentials. Exp. Neurol. 57, 200–211, http://dx.doi.org/10.1016/0014-4886(77)90057-7.

Oatman, L.C., Anderson, B.W., 1980. Suppression of the auditory frequency following response during visual attention. Electroencephalogr. Clin. Neurophysiol. 49, 314–322, http://dx.doi.org/10.1016/0013-4694(80)90225-4.

Obleser, J., Weisz, N., 2012. Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. Cereb. Cortex 22, 2466–2477, http://dx.doi.org/10.1093/cercor/bhr325.

Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., Maess, B., 2012. Adverse listening conditions and memory load drive a common alpha oscillatory network. J. Neurosci. 32, 12376–12383, http://dx.doi.org/10.1523/jneurosci.4908-11.2012.

Okada, Y.C., Wu, J., Kyuhou, S., 1997. Genesis of MEG signals in a mammalian CNS structure. Electroencephalogr. Clin. Neurophysiol. 103, 474–485, http://dx.doi.org/10.1016/s0013-4694(97)00043-6.

Olejnik, S., Algina, J., 2003. Generalized eta and omega squared statistics: measures of effect size for some common research designs. Psychol. Methods 8, 434–447, http://dx.doi.org/10.1037/1082-989X.8.4.434.

Palombi, P.S., Backoff, P.M., Caspary, D.M., 2001. Responses of young and aged rat inferior colliculus neurons to sinusoidally amplitude modulated stimuli. Hear. Res. 153, 174–180, http://dx.doi.org/10.1016/S0378-5955(00)00264-1.

Picton, T.W., Hillyard, S.A., 1974. Human auditory evoked potentials II: effects of attention. Electroencephalogr. Clin. Neurophysiol. 36, 191–199, http://dx.doi.org/10.1016/0013-4694(74)90156-4.

Picton, T.W., John, M.S., Dimitrijevic, A., Purcell, D., 2003. Human auditory steady-state responses. Int. J. Audiol. 42, 177–219.

Picton, T.W., Stapells, D.R., Campbell, K.B., 1981. Auditory evoked potentials from the human cochlea and brainstem. J. Otolaryngol. Suppl. 9, 1–41.

Purcell, D.W., John, S.M., Schneider, B.A., Picton, T.W., 2004. Human temporal auditory acuity as assessed by envelope following responses. J. Acoust. Soc. Am. 116, 3581–3593, http://dx.doi.org/10.1121/1.1798354.

Core Team, R., 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria (accessed 10.06.15).

Rinne, T., Balk, M.H., Koistinen, S., Autti, T., Alho, K., Sams, M., 2008. Auditory selective attention modulates activation of human inferior colliculus. J. Neurophysiol. 100, 3323–3327, http://dx.doi.org/10.1152/jn.90607.2008.

Rouder, J.N., Morey, R.D., Speckman, P.L., Province, J.M., 2012. Default Bayes factors for ANOVA designs. J. Math. Psychol. 56, 356–374, http://dx.doi.org/10.1016/j.jmp.2012.08.001.

Ruggles, D., Bharadwaj, H., Shinn-Cunningham, B.G., 2012. Why middle-aged listeners have trouble hearing in everyday settings. Curr. Biol. 22, 1417–1422, http://dx.doi.org/10.1016/j.cub.2012.05.025.

Ruggles, D., Shamma, S., Oxenham, A.J., 2014. Sub-cortical phase locking to attended and unattended streams of resolved and unresolved harmonic complex tones. Assoc. Res. Otolaryngol. Abs. 37, 34.

Scherg, M., von Cramon, D., 1985. A new interpretation of the generators of BAEP waves I-V: results of a spatio-temporal dipole model. Electroencephalogr. Clin. Neurophysiol. 62, 290–299, http://dx.doi.org/10.1016/0168-5597(85)90006-1.

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038

Scherg, M., Vajsar, J., W, P.T., 1989. A source analysis of the late human auditory evoked potentials. J. Cogn. Neurosci. 5, 363–370, http://dx.doi.org/10.1162/jocn.1989.1.4.336.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270, 303–304, http://dx.doi.org/10.1126/science.270.5234.303.

Sharma, A., Marsh, C.M., Dorman, M.F., 2000. Relationship between N1 evoked potential morphology and the perception of voicing. J. Acoust. Soc. Am. 108, 3030–3035, http://dx.doi.org/10.1121/1.1320474.

Shinn-Cunningham, B., Ruggles, D.R., Bharadwaj, H., 2013. How early aging and environment interact in everyday listening: from brainstem to behavior through modeling. In: Moore, B.C. J., Patterson, R.D., Winter, I.M., Carlyon, R.P., Gockel, H.E. (Eds.), Basic Aspects of Hearing. Springer Science+Business Media, New York, pp. 501–510, http://dx.doi.org/10.1007/978-1-4614-1590-9\s\do5(5)5.

Shinn-Cunningham, B.G., 2008. Object-based auditory and visual attention. Trends Cogn. Sci. 12, 182–186, http://dx.doi.org/10.1016/j.tics.2008.02.003.

Shinn-Cunningham, B.G., Best, V., 2008. Selective attention in normal and impaired hearing. Trends Amplif. 12, 283–299, http://dx.doi.org/10.1177/1084713808325306.

Sininger, Y.S., Hyde, M.L., 2009. Auditory brainstem response in audiometric threshold prediction. In: Katz, J., Medwetsky, L., Burkard, R., Hood, L. (Eds.), Handbook of Clinical Audiology 6th ed Lippincott Williams & Wilkins, Baltimore, MD, pp. 293–321.

Skoe, E., Kraus, N., 2010a. Auditory brain stem response to complex sounds: a tutorial. Ear Hear. 31, 302–324, http://dx.doi.org/10.1097/aud.0b013e3181cdb272.

Skoe, E., Kraus, N., 2010b. Hearing it again and again: on-line subcortical plasticity in humans. PLoS One, 5, http://dx.doi.org/10.1371/journal.pone.0013645.

Slee, S., David, S., 2015. Rapid task-related plasticity of spectrotemporal receptive fields in the auditory midbrain. Assoc. Res. Otolaryngol. Abs. 38, 238.

Slepian, D., 1978. Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty—V: The Discrete Case. Bell System Technical Journal 57, 1371–1430.

Sohmer, H., Feinmesser, M., 1967. Cochlear action potentials recorded from the external ear in man. Ann. Otol. Rhinol. Laryngol. 76, 427–435, http://dx.doi.org/10.1177/000348946707600211.

Strauss, A., W"ostmann, M., Obleser, J., 2014. Cortical alpha oscillations as a tool for auditory selective inhibition. Front. Hum. Neurosci. 8, 350, http://dx.doi.org/10.3389/fnhum.2014.00350.

Sussman, E.S., Bregman, A.S., Wang, W.J., Khan, F.J., 2005. Attentional modulation of electrophysiological activity in auditory cortex for unattended sounds within multistream auditory environments. Cogn. Affect. Behav. Neurosci. 5, 93–110.

Winer, J.A., 2006. Decoding the auditory corticofugal systems. Hear. Res. 212, 1–8, http://dx.doi.org/10.1016/j.heares.2005.06.014.

Wittekindt, A., Kaiser, J., Abel, C., 2014. Attentional modulation of the inner ear: a combined otoacoustic emission and EEG study. J. Neurosci. 34, 9995–10002, http://dx.doi.org/10.1523/jneurosci.4861-13.2014.

Woldorff, M., Hansen, J.C., Hillyard, S.A., 1987. Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. Electroencephalogr. Clin. Neurophysiol. Suppl. 40, 146–154.

Woods, D.L., Elmasian, R., 1986. The habituation of event-related potentials to speech sounds and tones. Electroencephalogr. Clin. Neurophysiol. 65, 447–459, http://dx.doi.org/10.1016/0168-5597(86)90024-9.

Worden, F.G., Marsh, J.T., 1968. Frequency-following (microphonic-like) neural responses evoked by sound. Electroencephalogr. Clin. Neurophysiol. 25, 42–52, http://dx.doi.org/10.1016/0013-4694(68)90085-0.

Xiao, Z., Suga, N., 2002. Modulation of cochlear hair cells by the auditory cortex in the mustached bat. Nat. Neurosci. 5, 57–63, http://dx.doi.org/10.1038/nn786.

Yan, J., Suga, N., 1996. Corticofugal modulation of time-domain processing of biosonar information in bats. Science 273, 1100–1103, http://dx.doi.org/10.1126/science.273.5278.1100.

Yan, J., Suga, N., 1999. Corticofugal amplification of facilitative auditory responses of subcortical combination-sensitive neurons in the mustached bat. J. Neurophysiol. 81, 817–824.

Zhu, L., Bharadwaj, H., Xia, J., Shinn-Cunningham, B., 2013. A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. J. Acoust. Soc. Am. 134, 384–395, http://dx.doi.org/10.1121/1.4807498.

Please cite this article as: Varghese, L., et al., Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Research (2015), http://dx.doi.org/10.1016/j.brainres.2015.06.038