

# Endogenous auditory frequency-based attention modulates electroencephalogram-based measures of obligatory sensory activity in humans

Caroline M. Sheedy<sup>a</sup>, Alan J. Power<sup>b,e</sup>, Richard B. Reilly<sup>a,b,c,d</sup>,  
Michael J. Crosse<sup>a,b,c</sup>, Gerard M. Loughnane<sup>a,b,c</sup> and Edmund C. Lalor<sup>a,b,c</sup>

Auditory selective attention is the ability to enhance the processing of a single sound source, while simultaneously suppressing the processing of other competing sound sources. Recent research has addressed a long-running debate by showing that endogenous attention produces effects on obligatory sensory responses to continuous and competing auditory stimuli. However, until now, this result has only been shown under conditions where the competing stimuli differed in both their frequency characteristics and, importantly, their spatial location.

Thus, it is unknown whether endogenous selective attention based only on nonspatial features modulates obligatory sensory processing. Here, we investigate this issue using a diotic paradigm, such that competing auditory stimuli differ in frequency, but had no separation in space. We find a significant effect of attention on electroencephalogram-based measures of obligatory sensory processing at several poststimulus latencies.

## Introduction

Auditory selective attention refers to the ability to enhance the processing of information received from one sound source while suppressing the processing of information from other, competing sound sources [1]. Selective auditory attention typically involves focusing on a subset of stimuli within an auditory scene on the basis of specific auditory features such as pitch, location, and intensity [1,2]. Research on visual attentional mechanisms suggests the presence of a goal-directed top-down signaling mechanism that recruits frontoparietal cortical regions for spatial attention [3], and that, although similar cortical regions are used for feature-based attention, the precise amount of overlap is unclear. More recent work has addressed this issue in the auditory modality by showing that attention to both spatial and frequency-based features is controlled by an overlapping (left dominated) frontoparietal attention network [4], albeit with significant differences in the patterns of activity between spatial and feature-based attention. Thus, the neural mechanisms underpinning spatial attention cannot be assumed to be precisely the same as those governing feature-based attention.

To investigate such issues, researchers often utilize the electroencephalogram (EEG) [5,6]. Typically, these studies involve presenting discrete stimuli under different

We discuss these results in terms of previous research on feature-based attention and by comparing our findings with the previous work using stimuli that differed both in terms of spatial and frequency-based characteristics. *NeuroReport* 00:000–000 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins.

*NeuroReport* 2013, 00:000–000

**Keywords:** auditory evoked potential, competing sound sources, diotic listening, endogenous attention

<sup>a</sup>Trinity College Institute of Neuroscience, <sup>b</sup>School of Engineering, <sup>c</sup>Trinity Centre for Bioengineering, <sup>d</sup>School of Medicine, Trinity College Dublin, Dublin 2, Ireland and <sup>e</sup>Department of Psychology, Centre for Neuroscience in Education, University of Cambridge, Cambridge, UK

Correspondence to Edmund C. Lalor, PhD, Trinity College Dublin, Dublin 2, Ireland  
Tel: +353 1 8961743; fax: +353 1 6772442; e-mail: edlador@tcd.ie

Received 27 August 2013 accepted 24 September 2013

attentional conditions and assessing task effects on time-locked averaged event-related potentials (ERPs). Studies showing attention effects on auditory ERPs [i.e. auditory evoked potentials (AEPs)] led to a long-lasting debate as to the mechanisms underpinning endogenous attention. Hillyard *et al.* [5] showed that, a major component of the AEP, the N1, was considerably larger in AEPs to attended tones than to unattended tones. This enhancement was proposed to be a result of increased sensory processing of the attended stimulus. However, Näätänen [6] contested that the AEP effects did not actually represent a true enhancement of sensory processing, but that it only appeared to be so because of the presence of a separate attentionally generated negativity that overlapped the N1 in time. It was proposed that this negativity had a different cerebral source than that of the obligatory response and that it may arise as a result of a matching process between the sensory input and an attentional trace [7].

That a separate endogenously generated negativity overlaps the obligatory sensory ERP remains a real possibility. However, recent research has provided strong evidence that enhancement of obligatory auditory processing activity by endogenous spatial attention does occur [8]. That study utilized a method for obtaining temporally detailed EEG responses to continuous stimuli, which

facilitated investigation of obligatory sensory processing without the risk of complication from any endogenously generated activity time-locked to the presentation of a discrete stimulus. The method, which is known as the auditory-evoked spread spectrum analysis (AESPA), also allows for the derivation of separate responses to each of several simultaneous stimuli [9]. It produces a measure indexing the obligatory sensory cortical activity to the amplitude modulation of a time-varying continuous auditory stimulus. By presenting two such stimuli simultaneously in a dichotic manner, robust endogenous attention effects were shown on these AESPA responses at  $\sim 136$  ms that were localized to the auditory cortex [8]. The method has also been used to identify attention effects to natural speech [10].

Although both of these studies provide evidence for attentional enhancement of obligatory sensory processing, it is still unknown whether selective attention on the basis of a feature other than space will produce such an effect. This study aims to investigate this using frequency-based selective attention in a paradigm that is identical to that used in [8], except for the presentation of stimuli in a diotic (same sound to both ears) rather than dichotic (different stimulus to each ear) manner.

## Methods

### Participants and data acquisition

Seventeen individuals aged 21–33 (mean  $\pm$  SD =  $24.5 \pm 3.4$  years; 15 men) participated in the study. The experiment was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Psychology at Trinity College Dublin. Each participant provided written informed consent. Participants reported no history of hearing impairment or neurological disorder. EEG data were recorded from 128 scalp electrode positions, rereferenced to the average of two mastoid electrodes, filtered over the range 0–134 Hz, and digitized at 512 Hz using a BioSemi Active Two system (BioSemi BV, Amsterdam, the Netherlands). EEG data were then digitally filtered off-line between 2 and 35 Hz.

### Stimuli

For all trials, two auditory stimuli were presented simultaneously and diotically. These stimuli were amplitude-modulated narrowband noise. Specifically, two distinct carrier waves were amplitude modulated by independent Gaussian noise signals with uniform power between 0 and 30 Hz. The carrier waves were root mean square normalized band-pass noise signals of bandwidth 1 kHz centered at 1 kHz (Low stream) and 5 kHz (High stream), respectively. Given their separation in frequency and the fact that they were modulated with different Gaussian signals, these noise signals could be perceived separately.

Because of the logarithmic relationship between auditory stimulus intensity and perception, the modulating signals,  $x$ , were mapped to the amplitude of the auditory stimulus,  $x'$ , according to the exponential relationship

$$x' = 10^{2x}$$

and normalized to between 0 and 1. Transitions between levels were smoothed using a 5-ms ramp consisting of half a period of a 100-Hz sine wave. The modulation rate of each signal was set to 60 Hz (see Lalor *et al.* [9]).

### Procedure

Testing was carried out in a dark room. Each participant undertook 10 120-s trials, where they attended the High stream, and 10 120-s trials, where they attended the Low stream. The sequence of conditions was randomized for each participant. Stimuli were presented at an intensity level deemed comfortable by the participant before beginning the experiment. Participants were instructed to fixate on a small cross presented at the center of a computer monitor and to minimize blinking and other motor activity during each trial.

To monitor behavioral performance, target and distracter events were inserted into each stream (Fig. 1d and e). These events consisted of specific patterns of amplitude modulation imposed on the Low and High streams. Targets consisted of a modulation level of  $-2.5$  dBfs for 25.5 ms, followed by  $-12$  dBfs for 16 ms, followed by  $-2.5$  dBfs for 25.5 ms, yielding a total length of 67 ms, whereas distracters consisted of a flat modulation of  $-6$  dBfs for 67 ms. dBfs refers to decibels full scale and represents a dB value relative to the maximum modulation level for each participant. Although the events are embedded in the stimulus, they are still distinguishable from the ongoing amplitude modulations. Participants were directed to click a mouse button only when a target in the attended stream was heard. Each trial contained 24 events; however, the proportion of targets and distracters was randomized for each trial, ranging from eight targets (and 16 distracters) to 16 targets (and eight distracters). On average, 48.75% of events across trials were targets (51.25% distracters). An event (in either stream) could not occur within 1 s of another event and also could not occur later than 9 s after another event. Responses to both High and Low streams were extracted from the EEG for each attention condition.

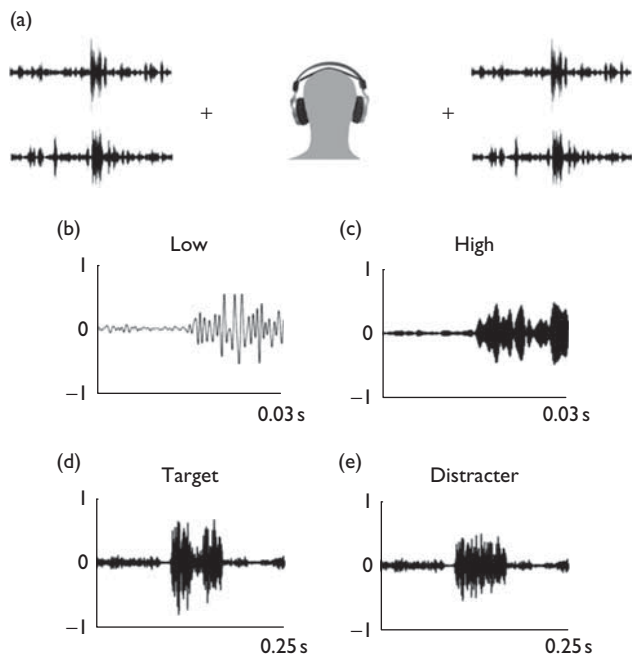
### Signal processing

The AESPA is an estimate of the impulse response function of the auditory system and is determined by performing a linear least squares fit of the response model

$$y(t) = w(\tau) * x(t) + \text{noise},$$

where  $y(t)$  is the EEG response,  $x(t)$  is the amplitude modulation signal, the symbol  $*$  indicates convolution,  $w(\tau)$  is the unknown AESPA response, and the noise is assumed to be Gaussian [9]. The AESPA  $w(\tau)$  can be

Fig. 1



Stimulation setup (a) Low-frequency and High-frequency streams were presented to each participant dichotically. A segment of the (b) Low-frequency and (c) High-frequency carrier stimuli, respectively. Examples of (d) target and (e) distracter events.

compared with a filter that describes how the continuous auditory stimulus is transformed by the brain in to the continuous EEG. It is important to note that the AESPA does not correspond precisely to a typical ERP time axis. Each point on the time axis of the AESPA represents the relative time between the continuous EEG and the continuous input intensity signal.

After deriving the AESPA for both the Low and the High stimulus for each attention condition, run, and participant, the AESPA responses were run through an algorithm to detect flat or bad channels. A channel was defined as ‘bad’ if its SD was greater than three times the SD of the surrounding channels. AESPAs on those channels were then recalculated by spline interpolating the AESPAs on the surrounding good channels.

### Data analysis

To determine task performance, we considered any response occurring within a 1-s period after an event to be a response to that event. The percentages of correct responses to targets in the attended stream, to distracters in the attended stream, and to events in the unattended stream were calculated (Table 1). Because of the difficulty of the discrimination task, participants responded equally often to targets and distracters in the attended stream (see the Results section below). As such,

Table 1 Behavioral results

	Low	High
p(T)±SD (%)	76.64±14.26	76.83±10.43
p(D)±SD (%)	74.35±30.92	71.52±34.28
p(H)±SD (%)	75.08±24.38	73.93±23.12
p(U)±SD (%)	1.61±3.47	1.54±2.75

Mean task performance across all participants and runs. Data are the percentage of particular events where a response was given±SD (%). p(T) is that for targets, p(D) is that for distracters in the attended stream, p(H) is that for hits (i.e. responses to any event in the attended stream), and p(U) is that for events in the unattended stream.

we will henceforth refer to a response to any event (target or distracter) in the attended stream as a ‘hit’ (relative to responses to events in the unattended stream). A  $2 \times 2$  repeated measures analysis of variance (ANOVA) was carried out using factors of stimulus (Low vs. High) and response (hits vs. responses to events in the unattended stream).

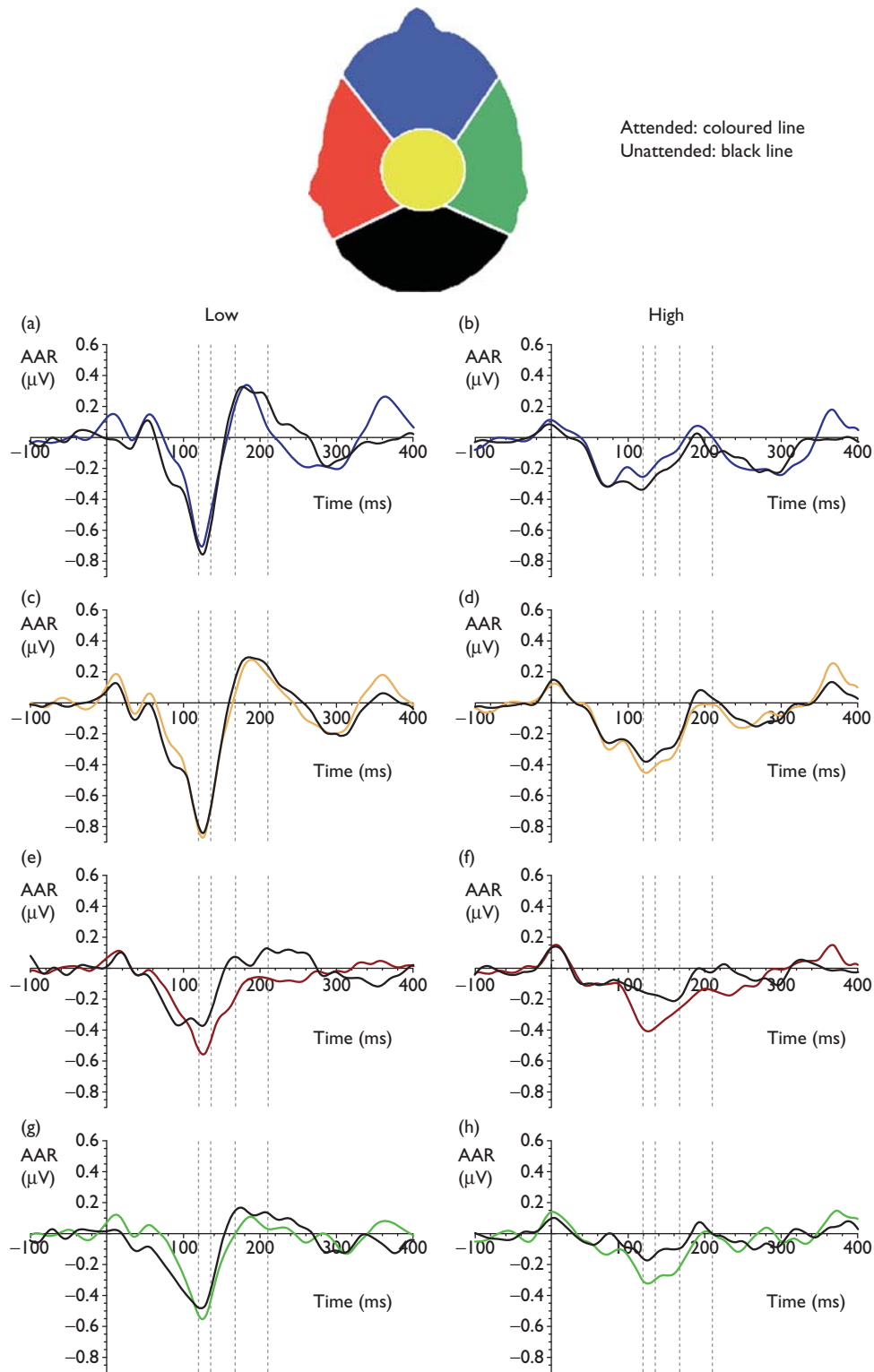
We examined the two most robust AESPA components for any possible attention effects. We did this by first defining their amplitudes as being the root mean square amplitudes for each in the following intervals (negative component: 119–135 ms and positive component: 167–210 ms), which we determined on the basis of the grand average AESPA responses. We carried out a 4-way  $2 \times 2 \times 4 \times 2$  repeated-measures ANOVA using factors of stimulation (Low vs. High stream), attention (attended vs. unattended), scalp region (frontal, central, left temporal, right temporal; Fig. 2), and component (negative vs. positive). Our independent measures were the average component amplitudes across all electrodes in a defined scalp region (frontal, central, left temporal and right temporal) for each stimulus and condition. Parietal and occipital regions were not included in this analysis as previous studies have shown the AESPA not to be strongly represented in these regions [9]. In cases where the repeated measures factors violated the sphericity assumption, Greenhouse–Geisser corrected *P* values and degrees of freedom are reported.

## Results

### Behavioral results

The task of discriminating between targets and distracters in the attended stream while ignoring events in the opposite stream was purposely made difficult to maximize the participants’ attentional efforts. Although this was previously found to be manageable when the Low and High streams were presented dichotically [8], the lack of spatial separation in the current study caused a considerable increase in task difficulty. The participants responded equally often to target and distracter events in the attended stream (Table 1). As mentioned previously, we then combined all of these responses into a measure called ‘hits’. When comparing this with the very low percentage of events responded to in the unattended

Fig. 2



Average AESPA responses (AAR) across electrodes in each of the four regions – (a, b) frontal, (c, d) central, (e, f) left temporal, (g, h) right temporal for Low (first column) and High (second column) stimuli when attended and unattended. AESPA, auditory-evoked spread spectrum analysis. Dashed lines indicate intervals used to define the amplitudes of the negative and positive component.

stream, it is clear that participants were successful at attending to the cued stream while ignoring the other. This was confirmed by our  $2 \times 2$  repeated measures ANOVA, which showed a significant main effect of attention ( $P < 0.001$ ), with no main effect of stimulus or stimulus  $\times$  response interaction ( $P > 0.05$ ), indicating that task performance did not differ whether it was the High or the Low stream that was being attended to.

### Auditory-evoked spread spectrum analysis results

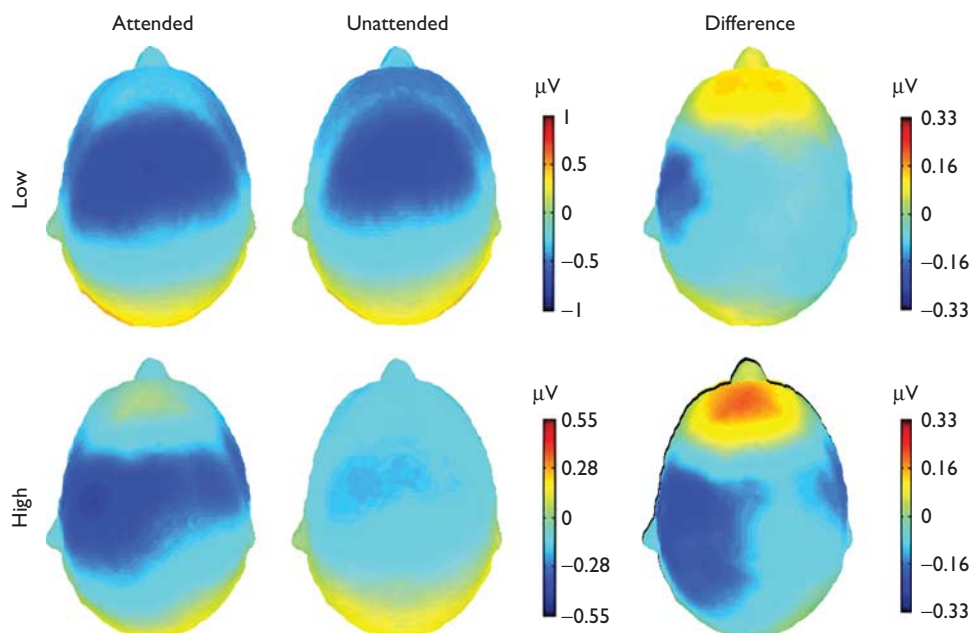
Figure 2 shows the AESPA responses to the Low and High streams, averaged across electrodes within all four regions, when they were attended and unattended. Differences between the attended and unattended responses are clear, in particular, for the left and right temporal regions. The primary result of our study is that our four-way repeated-measures ANOVA yielded a main effect of attention ( $F_{1,16} = 6.035$ ,  $P < 0.05$ ). In addition, there was a significant region  $\times$  attention interaction ( $F_{3,48} = 3.088$ ,  $P < 0.05$ ). Post-hoc  $t$ -tests show that this interaction was driven by attentional effects in left (two-tailed  $t$ -test;  $t_{16} = 3.924$ ,  $P < 0.001$ ) and right (two-tailed  $t$ -test;  $t_{16} = 2.471$ ,  $P < 0.05$ ) regions only, although this latter result did not survive Bonferroni's correction for multiple comparisons. There was no significant attention  $\times$  stimulus interaction ( $P = 0.092$ ), indicating that attentional effects were consistent for Low and High stimuli. There was also a main effect of stimulus (repeated measures ANOVA;  $F_{1,16} = 38.828$ ,  $P < 0.0001$ ), driven by the fact that Low

responses are larger than High responses, a result that has been noted in previous work [8]. Finally, there was no significant attention  $\times$  component interaction ( $P = 0.618$ ), indicating that attentional effects were not different between the two components.

Figure 3 provides another view of the data in the form of topographic scalp maps for the negative component when attended and unattended, alongside a plot of the difference topographies. Maps for the positive component are not included as our statistical results showed no attention  $\times$  component interaction. Responses consistent with activity in auditory cortex are visible in both the attended and the unattended conditions. The difference maps highlight the left and (weaker) right temporal attention effects reported in our ANOVA analysis above.

The possibility that increased activity during the time interval of the negative component may be because of an engagement of additional nonobligatory generators and not increased activity of the sensory activity was investigated using the (paired) TANOVA method [11]. This method assesses whether two topographies are statistically different using a nonparametric randomization procedure. Topographies in attended and unattended conditions were found not to be statistically different for either the Low ( $P = 0.92$ ) or the High ( $P = 0.90$ ) response. Although finding no significant difference is not equivalent to proving that they are definitively the

Fig. 3



Topographic maps of the negative component of Low and High AESPA responses when attended and unattended as well as difference topographies shown in the right panel. Size and polarity of responses are indicated by color bars.

same, we take this result as supporting our contention that the same generators are involved in both attended and unattended conditions.

## Discussion

Following our previous study, which showed that obligatory sensory processing is modulated by endogenous auditory attention in a dichotic paradigm [8], here, we have shown that the same is true for endogenous attention without spatial cues. It is important to note that in our previous study, the two stimuli were separated both in space and in frequency [8]. Thus, it is possible that the attention effects we observed in that study were driven by frequency-based attention and not space. However, on the basis that target/distracter discrimination was so much better in our previous study compared with the present one, we contend that spatial attention played an important role in our previous work, and that our present study provides unequivocal evidence for an enhancement in sensory processing based purely on frequency-based attention. As such, we maintain that this finding makes a further contribution to the debate on the neural mechanisms underpinning endogenous attention [5,6].

Although it is difficult to directly compare the findings from the present work with that of our previous paper because of the aforementioned differences in the performance on the discrimination task, we note that the results from the two studies are largely similar, albeit not identical. Both studies showed clear attention effects on both the early negative and the later positive AESPA components; however, the attention effects in the dichotic study were found to be more widely distributed on the scalp (left, right, and central regions) than in the present study (just left and right temporal regions). Although it is tempting to suggest that this may be because of greater activation in the dorsal precentral sulcus during spatial attention [4], it is perhaps more likely that it is simply driven by weaker attention effects in the present study because of lower attentional engagement in our overly difficult discrimination task. Indeed, on the basis of our method, it is actually difficult for us to comment at all on the differential engagement of attention networks in controlling spatial versus frequency-based endogenous attention. This is because the AESPA response measures obligatory sensory processing of the temporal variations in our stimulus and, as such, it is likely not to be directly sensitive to tonic changes in the EEG that index the involvement of networks controlling top-down attentional deployment [3,4,12–14]. As such, what we see in our data are simply the effects that these networks have on sensory processing. Future work combining our complex stimuli with cued preparatory periods may yield more insights.

Although the results of the present study have provided further evidence of an enhancement of auditory sensory

processing by endogenous attention, we cannot rule out the existence of an attention-related negativity that temporally overlaps the N1 and that is generated by a region of cortex distinct from that generating the AESPA [7]. Because such a component has been suggested to index a matching process between the neural representation of a discrete stimulus and an ‘attentional trace’, it would be unlikely to be well synchronized to our complex stimulus fluctuations and, as such, it would not be reflected in the AESPA [8].

Previous work has shown that attention operates to select relevant information at the stage of processing that is most overloaded by a particular stimulus–task combination [15,16]. This theory was supported by recent AESPA work on selective attention that showed temporally specific attention effects at around 200 ms when participants were asked to attend to one of two competing speech streams with a view to subsequently answering questions on that stream [10]. In the present study, as in our previous paper [8], we observed temporally broad attention effects on our AESPA responses. This may be because our attention task involved distinguishing particular patterns of amplitude modulation. Because amplitude modulation is the very stimulus property that produces the entire AESPA response, it seems unsurprising that attention to this very property led to broad effects across the response.

## Conclusion

We have found that endogenous auditory attention without the use of spatial cues produces significant effects on obligatory sensory processing that are similar to those effects observed when spatial cues are available.

## Acknowledgements

This study was supported by a grant from Science Foundation Ireland (09-RFP-NES2382) and by the Irish Research Council for Science, Engineering & Technology.

## Conflicts of interest

There are no conflicts of interest.

## References

- 1 Coch D, Sanders LD, Neville HJ. An event-related potential study of selective auditory attention in children and adults. *J Cogn Neurosci* 2005; **17**:605–622.
- 2 Maddox RK, Shinn-Cunningham BG. Influence of task-relevant and task-irrelevant feature continuity on selective auditory attention. *J Assoc Res Otolaryngol* 2012; **13**:119–129.
- 3 Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 2002; **3**:201–215.
- 4 Hill KT, Miller LM. Auditory attentional control and selection during cocktail party listening. *Cereb Cortex* 2010; **20**:583–590.
- 5 Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. *Science* 1973; **182**:177–180.
- 6 Näätänen R. Selective attention and evoked potentials in humans – a critical review. *Biol Psychol* 1975; **2**:237–307.
- 7 Näätänen R. Processing negativity: an evoked-potential reflection of selective attention. *Psychol Bull* 1982; **92**:605–640.

- 8 Power AJ, Lalor EC, Reilly RB. Endogenous auditory spatial attention modulates obligatory sensory activity in auditory cortex. *Cereb Cortex* 2011; **21**:1223–1230.
- 9 Lalor EC, Power AJ, Reilly RB, Foxe JJ. Resolving precise temporal processing properties of the auditory system using continuous stimuli. *J Neurophysiol* 2009; **102**:349–359.
- 10 Power AJ, Foxe JJ, Forde EJ, Reilly RB, Lalor EC. At what time is the cocktail party? A late locus of selective attention to natural speech. *Eur J Neurosci* 2012; **35**:1497–1503.
- 11 Murray MM, Brunet D, Michel CM. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr* 2008; **20**:249–264.
- 12 Giesbrecht B, Woldorff MG, Song AW, Mangun GR. Neural mechanisms of top-down control during spatial and feature attention. *Neuroimage* 2003; **19**:496–512.
- 13 Schulze K, Gaab N, Schlaug G. Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neurosci* 2009; **10**:106.
- 14 Lee AK, Rajaram S, Xia J, Bharadwaj H, Larson E, Hamalainen MS, *et al.* Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Front Neurosci* 2012; **6**:190.
- 15 Vogel EK, Woodman GF, Luck SJ. Pushing around the locus of selection: evidence for the flexible-selection hypothesis. *J Cogn Neurosci* 2005; **17**:1907–1922.
- 16 Kelly SP, Gomez-Ramirez M, Foxe JJ. Spatial attention modulates initial afferent activity in human primary visual cortex. *Cereb Cortex* 2008; **18**:2629–2636.