Hearing Research 268 (2010) 213–226

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Evidence for modality-specific but not frequency-specific modulation of human primary auditory cortex by attention

P.E. Gander, D.J. Bosnyak, L.E. Roberts*

Department of Psychology, Neuroscience, and Behaviour McMaster University, Hamilton, Ontario, Canada L8S 4K1

ARTICLE INFO

Article history:
Received 8 March 2010
Received in revised form
20 May 2010
Accepted 7 June 2010
Available online 12 June 2010

ABSTRACT

We used the stimulus-driven 40-Hz auditory steady-state response (ASSR) that localizes tonotopically to the region of primary auditory cortex (A1) to study modulation of this region by top–down attention. Experiment 1 presented amplitude modulated (AM) auditory and visual stimuli simultaneously (AM at 40 Hz and 16 Hz, respectively) while participants responded to targets in one modality or the other. ASSR amplitude increased from an unattended passive baseline during auditory but not visual attention demonstrating *modality-specific* auditory attention, when attention was required for brief (1 s) but not long (2 min) time intervals. Modality-specific visual attention occurred at both time intervals. Experiment 2 asked whether attention directed to one or the other of two simultaneous auditory streams (carrier frequencies of 250 and 4100 Hz AM at 37 and 41 Hz respectively, counterbalanced) increased ASSR amplitude for the attended stream (*frequency-specific* auditory attention). Behaviour was strongly controlled by carrier frequency (overall target rate 1.7 Hz), and the cortical sources of the two carriers were resolved by inverse modeling. Despite these conditions favourable to frequency specificity, frequency-specific modulation of ASSR amplitude was not found at either time interval. Frequency-specific modulation of A1 may require re-entrant feedback to the auditory core from auditory percepts that possess distinct spectral attributes and are attended in higher regions of the auditory system.

1. Introduction

Functional imaging studies have established that neural activity is increased in several regions of the auditory cortex when human participants listen to and process auditory stimuli. Increased activity has been reported in the region of primary auditory cortex (A1) in Heschl's gyrus (Jäncke et al., 1999) where the spectrotemporal features of sounds are represented, as well as in lateral and posterior nonprimary regions of the superior temporal gyrus (STG) that are believed to be involved in processes underlying auditory object identification (Alho et al., 1999; Grady et al., 1997; Woodruff et al., 1996). Increased neural activity during active listening could reflect (i) sensory activations evoked by the attended sounds, or (ii) activations induced by attention, or (iii) both factors. Functional imaging studies distinguishing these effects have reported that while sensory activations are distributed throughout A1 and nonprimary auditory regions (Hall et al., 2000; Johnson and Zatorre, 2005; Petkov et al., 2004; Tzourio et al., 1997), attentional modulations are more prominent in (or in some cases confined to) nonprimary areas of the STG where more complex forms of auditory processing are believed to take place (Johnson and Zatorre, 2005, 2006; Petkov et al., 2004). It has been suggested that attention directed to sounds consisting of simple spectrotemporal features (for example, a consonant vowel combination) may be more likely to augment activity in A1 than is attention directed to sounds that require relational processing of sound features (for example, a novel melody) or reference to auditory working memory (Johnson and Zatorre, 2005). While this hypothesis is broadly consistent with functional imaging data (Jäncke et al., 2003), the conditions that favor the expression of effects of attention in A1 have not been firmly identified.

Abbreviations: 2IFC, two interval forced choice procedure; A1, primary auditory cortex; AEF, auditory evoked field; AEP, auditory evoked potential; AM, amplitude modulated; ANOVA, analysis of variance; ASSR, auditory steady-state response; BESA, Brain Electrical Source Analysis; BF, basal forebrain; EEG, electroencephalography; FFT, fast fourier transform; fMRI, functional magnetic resonance imaging; HG, Heschl's gyrus; ITI, intertrial interval; LED, light-emitting diode; MEG, magnetoencephalography; MLRs, middle latency responses; PET, positron emission tomography; SPL, sound pressure level; STG, superior temporal gyrus; TH, threshold; V1, primary visual cortex; VSSR, visual steady-state response.

* Corresponding author. Department of Psychology, Neuroscience, and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. Tel.: +1 905 525 9140x24052; fax: +1 905 529 6225.

E-mail addresses: pgander@mcmaster.ca (P.E. Gander), bosnyak@mcmaster.ca (D.J. Bosnyak), roberts@mcmaster.ca (L.E. Roberts).

0378-5955/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2010.06.003

An alternative approach to investigating the effects of top–down attention in A1 is to study auditory evoked potentials (AEPs) that are known to localize to generators in the region of the auditory core. One such response is the stimulus-driven *auditory steady-state response* (ASSR). This response reaches its maximum amplitude when sounds are amplitude-modulated (AM) at rates near 40 Hz and localizes in intracerebral recordings to posteriormedial Heschl's gyrus (HG) where A1 is found (Bidet-Caulet et al., 2007; Brugge et al., 2009). ASSR sources determined by inverse modeling from MEG data (Pantev et al., 1996; Wienbruch et al., 2006) show a high-frequency medial, low-frequency lateral tonotopic ordering in this region that is consistent with tonotopic organization described in several fMRI studies of tonotopy (Formisano et al., 2003; Petkov et al., 2004; Talavage et al., 2004) and with the common low frequency border situated laterally in tonotopic maps described by Kaas and Hackett (2000) for the macaque monkey. The source waveform of the ASSR extracted by deconvolution of ASSRs recorded at different AM rates (Bosnyak and Roberts, 2001; Gutschalk et al., 1999) resembles the oscillatory Na/Pa/Nb/Pb waveform of auditory "middle latency" responses (MLRs, ~19–55 ms post-stimulus) that also localize to posteriormedial HG (Godey et al., 2001) and represent the earliest stages of sound processing in the auditory cortex. Although the precise relation of middle latency and ASSR responses remains to be clarified (Galambos et al., 1981; Ross et al., 2005; Santarelli et al., 1995), linear summation of the ASSR source waveform (which, like the MLR waveform, has a wave period of ~25 ms) gives a reasonable account of ASSRs recorded at widely different AM rates including its amplitude maximum at rates near 40 Hz (Bosnyak and Roberts, 2001; Gutschalk et al., 1999). The ASSR is of interest for studies of attention, because it gives a picture of events taking place in the region of A1 at a time resolution exceeding that of other functional brain imaging methods.

Until recently few studies have investigated the effects of attention on the ASSR, and their results have been equivocal. Early investigations were mainly exploratory in scope and yielded a mixture of positive and negative findings (Hari et al., 1989; Makeig and Galambos, 1982). A more extensive series of studies by Linden et al. (1987) found no effect on ASSR amplitude when participants were required to switch their attention between streams of 500 Hz and 1000 Hz stimulation presented separately to each ear at different AM rates. However, because ASSRs evoked by two 40-Hz AM sounds presented passively to the same ear suppress another when their carrier frequencies are $\leq$3 octaves apart (Ross et al., 2003), it is possible that competitive binaural interactions among the closely spaced stimuli used by Linden et al. may have obstructed the expression of attention effects.

More recent research using single AM sounds has yielded more positive results. Ross et al. (2004) presented a 40-Hz AM tone to the right ear (standard stimulus) and required participants to press a button on trials on which the AM rate changed commencing 100 ms after stimulus onset (deviant stimulus). An increase in ASSR amplitude was found in the time interval ~180 and 500 ms contralateral to the stimulated ear, compared to a control condition in which the same participants viewed and silently categorized pictures into landscapes, animals, or humans while the AM sounds played in the background (these results were replicated by Saupe et al., 2009 for binaural stimulation). These findings are consistent with the conclusion that ASSR amplitude is increased by auditory attention implying more A1 neurons depolarizing synchronously under this condition. However, the alternative possibility exists that ASSR amplitude may have been suppressed by visual attention during the categorization task. Suppression of both primary and nonprimary auditory areas by visual processing has been reported in fMRI studies of bimodal attention (Johnson and Zatorre, 2005, 2006), and anatomical connections exist between auditory and visual areas that may mediate interactions between these modalities (Eckert et al., 2008; Falchier et al., 2002). Gander et al. (2007, submitted for publication) controlled for visual suppression in an auditory training study using the ASSR. Participants in an experimental group attended to auditory signals and pressed a button whenever they detected a single 40-Hz AM pulse of increased amplitude that occurred randomly in one of two tones in a two-interval forced choice (2IFC) task. ASSR amplitude increased under auditory attention compared to a preceding passive baseline in which the same stimuli were presented as unattended background signals while the participants watched a subtitled video. ASSR amplitude did not change from an identical baseline in an unattended control group when the video was switched off, demonstrating that its increase during auditory attention in the experimental group was caused by the auditory task and not by its release from suppression by visual processing. However, because the auditory task of Gander et al. (submitted for publication) was designed as a training procedure, it included additional elements (a requirement for behavioural responding at trial offset and knowledge of results) that may have contributed to ASSR enhancement, in addition to auditory attention. It should be noted that preparation of a button-press response was also required during auditory but not visual attention in the experiments of Ross et al. (2004) and Saupe et al. (2009). In summary, these studies point strongly to an effect of attention on ASSR amplitude, but the contribution of other task features has yet to be disambiguated.

Here we report two experiments designed to evaluate effects of top–down attention on the ASSR when variables other than auditory attention were controlled. Experiment 1 investigated the *modality-specificity* of auditory attention. Is ASSR amplitude increased specifically by attention to the auditory modality, or will attention to the visual modality have the same effect? To answer this question, we presented amplitude-modulated auditory and visual stimuli simultaneously, each AM at a different rate, while participants attended to targets embedded in one stream or the other. Auditory (ASSR) and visual (VSSR) steady-state responses were separated in the frequency domain and compared with a preceding baseline condition in which identical stimuli were presented as unattended background stimuli. This procedure allowed us to determine whether ASSR amplitude was enhanced preferentially from an unattended baseline by auditory compared to visual attention (modality-specific auditory attention), and also whether the VSSR was similarly enhanced from baseline compared to auditory attention (modality-specific visual attention), when other task features were constant between the conditions. Because attention effects were assessed from a preceding baseline, suppression of brain responses in one modality by attention to the other modality could also be evaluated.

Experiment 2 investigated the *frequency-specificity* of auditory attention. Can top–down attention selectively modulate specific-frequency regions in A1, enhancing ASSRs recorded from those regions compared to ASSRs evoked by unattended frequencies? To address this question we presented carrier frequencies of 250 Hz and 4100 Hz simultaneously, each AM at a different rate (37 and 41 Hz, counterbalanced), while participants attended to targets in one carrier frequency or the other carrier frequency. We chose carrier frequencies that were not harmonically related and separated by more than 4 octaves, in order to minimize interactions among the carriers and increase the likelihood that selective modulation of distinct tonotopic regions in A1 might be found. The AM rates associated with each carrier frequency, while sufficiently different to distinguish the two brain responses by Fourier analysis, were perceptually confusable and counterbalanced with respect to carrier frequency, to ensure that participants solved the task by attending to carrier frequency. Because the

same sounds were presented to both ears, selective modulation of ASSR amplitude could only be based on attention to carrier frequency and not to ear of presentation.

A third question investigated by both experiments concerned the time course of attention. The ASSR modulations reported by Ross et al. (2004) and Gander et al. (submitted for publication) occurred in the specific time intervals in which acoustic targets were presented, suggesting that auditory attention is sensitive to the timing of task events. In the present experiments auditory attention was contrasted between groups in which targets were presented randomly in time on trials of either 1 s or 2 min duration. The results were contrasted to those for visual attention deployed for the same durations, to examine possible differences in the properties of attention between the auditory and visual systems.

2. Experiment 1

2.1. Materials and methods

2.1.1. Participants

Thirty-four students at McMaster University aged 17–37 years (mean 21.2 years, 7 left-handed, 20 female) signed a consent form approved by the Research Ethics Board of McMaster University and received course credit or $20 for participation. Participants sat in a chair in front of a computer monitor in a sound attenuated (ambient noise level 16 dBA SPL), electrically shielded booth. Normal hearing status to 8 kHz was confirmed by measurements taken for each participant. All participants had normal or corrected to normal vision.

2.2. Auditory/visual stimuli

Auditory and visual stimuli were presented simultaneously (see Fig. 1a). The auditory stimulus was a 2 kHz pure tone presented at 60 dB SPL and AM by a sinusoid at 40.96 Hz (100% modulation depth). Tones were presented binaurally via ear inserts (Etymotic Research ER-2). The visual stimulus was an 8×8 array of light-emitting diodes (LEDs) mounted on a black panel 32 mm^2 covering the middle of the computer monitor placed 1.4 m from the participant. The LED array was AM by a sinusoid at 16 Hz (100% modulation depth) and coloured green (570 nm). The 16 Hz AM rate was near the peak of the mid-frequency AM transfer function for the visual system described by Regan (1989, p. 381). AM was done by a digital signal processor (Tucker Davis RP2.1). In pilot measurements with ten additional participants, the peak amplitude of the visual stimulus was adjusted until it subjectively matched the auditory stimulus at 60 dB SPL. Peak intensity determined by this procedure (13 cd/m^2) was adopted for the main experiment.

Embedded in each dual auditory/visual stream were "targets" consisting of single pulses of 40-Hz auditory or 16-Hz visual stimulation (Fig. 1b). Auditory targets consisted of a single 40-Hz AM pulse of variable increased sound intensity; visual targets consisted of a single 16-Hz AM pulse of variable increased luminance of the full 8×8 LED array. Participants signalled their detection of these targets (either auditory or visual, according to task instructions) by pressing a mouse key, as described below. At the outset of the experiment a staircase procedure determined the amplitude increases that were to be used as auditory and visual targets. The staircase commenced with an amplitude augmentation of 200% (auditory, from 60 dB SPL) or 100% (visual, from 13 cd/m^2) that all participants detected with 100% accuracy. Target intensity decreased from this level until the target was not detected, at which point the target increased again until it was detected, and then decreased until it was not (1 up/1 down procedure). After 80 trials, the amplitude corresponding to an accuracy of 50% (average of the last 20 trials) was defined as threshold (TH). From this measurement two sets of five targets of differing intensity (one set auditory and the other visual) were created, consisting of TH+10%, +20%, +30%, +40%, and +50%. These targets covered a range commencing just above threshold and were chosen to ensure that most would be detected. The purpose was to deliver detectable target events at a rate that maintained the participants' concentration on the attended modality.

2.2.1. Design and procedure

Four stages were delivered in a single session, in the order Passive 1 (P1), attend Auditory (A), attend Visual (V), and Passive 2 (P2) (see Fig. 1c). The order of the A and V stages was counterbalanced across participants. Stages were ~10 min duration and were separated by a brief pause of ~2 min. At the outset of the P1 stage participants were told that brain responses would be measured to auditory and visual stimuli. They were asked to sit passively and maintain their fixation on a white cross displayed on a black background 4 cm above the LED array (Fig. 1a). Instructions for stages A and V were given after the P1 stage had been completed. The P2 stage was identical to the P1 stage. During all stages participants were monitored by video camera for correct visual fixation.

Participants were randomly assigned to one of two groups that received different stimulus timing conditions (Group 1SEC, $n = 21$; Group 2MIN, $n = 13$). In Group 1SEC a "trial" consisted of 1 s of simultaneous 40 Hz auditory and 16 Hz visual stimulation. In each stage of the experiment 200 trials were delivered, grouped into five blocks of 40 each. Blocks were separated by a rest period under the control of the participant, which was typically brief (10 s). Two-thirds of the trials determined randomly contained a single auditory or single visual target, or one each of both targets (Fig. 1b illustrates a trial with both targets), distributed equally across the five intensity levels. Auditory targets were programmed to occur randomly between 488 and 830 ms (pulses 20–34) and visual targets randomly between 500 and 875 ms (pulses 8–14) from stimulus onset. There were no constraints on target timing so that they could overlap if a target was present in both modalities. After each trial in the attention stages (A and V) participants were prompted by text on the computer screen "Did you hear a target?" (stage A), or "Did you see a target?" (stage V). Participants signalled their answer with a mouse click (left for yes, right for no); the next trial commenced between 1.4 and 1.6 s later. In stages P1 and P2 auditory/visual stimuli identical to those presented in the attend stages were presented while participants maintained their focus on the fixation cross. The intertrial interval (ITI) varied between 1900 and 2100 ms in the passive stages, which approximated the intertrial interval in the A and V stages that depended on the time required for behavioural responses.

Group 2MIN received a similar procedure, except that each "trial" now consisted of 2 min of continuous auditory/visual stimulation. Five trials of 2-min each were presented in each stage, followed by a brief rest interval. Targets occurred randomly at a rate of 0.85 Hz within each stream, with the constraint that targets within a stream were separated by at least 200 ms. Thus overall target rate averaged 1.7 Hz throughout the 2-min trial, which required that participants maintain their focus on the attended stream. The five target intensity values were each randomly assigned to one of the five trials in each stage, such that each 2-min trial had a similar percentage increase for the auditory and visual targets. In stages A and V participants were instructed to press a button after each detected target in the attended stream, and were reminded which stream to attend to by the words "Look" or "Listen" presented at the top of the computer monitor. In Stages P1 and P2 participants maintained focus on the fixation cross while identical auditory and visual stimuli were presented passively as background stimuli.

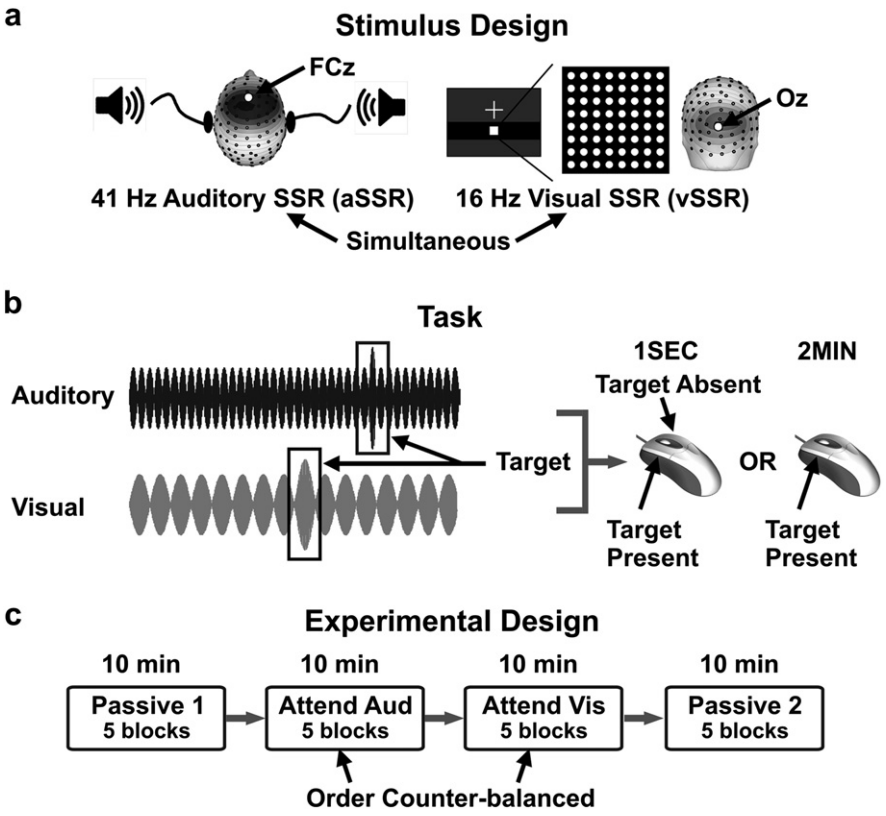

Fig. 1. Procedure for Experiment 1. (a) Stimulus delivery method and grand average voltage map are shown for the 41-Hz ASSR and 16-Hz VSSR. (b) Task design. Diagrams of dual auditory and visual AM stimulation are shown with a box around a target in each. Response requirements are indicated for each of the 1SEC and 2MIN groups. (c) Experimental design. For the two middle attention stages the order was counterbalanced across participants. [The procedures for Experiment 2 were analogous to those shown panels a–c, except that two auditory stimuli were substituted for auditory and visual stimuli.].

2.2.2. Electrophysiological recording

The electroencephalogram (EEG) was sampled at 2048 Hz (DC to 417 Hz) using a 128-channel Biosemi ActiveTwo amplifier (Cortech Solutions, Wilmington, NC). The electrode array was digitized for each participant (Polhemus Fastrak) prior to recording. EEG data were stored as continuous data files referenced to the vertex electrode.

2.2.3. Analysis of behavioural data

For the 1SEC group, behavioural performance was evaluated during the two attention stages for each participant by calculating the probability of reporting a target on trials containing a target (hits) for each of the five target values, and for trials containing no target (0% enhancement, false alarms). The probability of a hit was averaged across target values and compared with chance performance of $p = 0.50$ for each participant. In the 2MIN group, button presses during the active stages were tabulated in 5 ms bins throughout the stimulus. Response latency was then calculated for responses in each bin in two ways, first, with respect to the preceding target in the attended stream, and then with respect to the preceding target in the unattended stream. Distributions of response latency determined this way were then compared for each participant. Correct responding was signalled by a distribution containing a distinct peak corresponding to the latency of responses to attended targets, whereas the distribution of response latencies for the same responses calculated with respect to targets in the unattended stream was expected to be flat (random). Correct performance was assessed by comparing response rates between these distributions at the peak latency found for responses to attended targets.

2.2.4. Signal processing of EEG data

In Group 1SEC the EEG (128 channels) was epoched for each trial including 200 ms pre- and post-stimulus baselines. Epochs containing amplitude shifts of >100 μV (artifacts) were rejected from the analysis and the remaining epochs (90% of trials) averaged and converted to the average reference. To extract the ASSR the data were filtered 37–45 Hz (zero phase) and to extract the VSSR the data were filtered 12–20 Hz (zero phase). The 128-channel data for each participant were then collapsed into a two-pulse wide waveform over the interval 244–976 ms for the ASSR and from 250 to 1000 ms for the VSSR (for an example of the 2-pulse waveforms, see Fig. 3a later). In Group 2MIN the continuous EEG was epoched into consecutive 1-s segments. Epochs retained after artefact rejection (90%) were averaged, converted to the average reference, filtered to extract the ASSR and VSSR as in Group 1SEC, and collapsed into two-pulse wide waveforms. In both groups amplitude and phase were extracted from the 2 pulse wide waveforms by Fourier analysis, for electrode FCz for the ASSR and for electrode Oz for the VSSR, where the corresponding evoked field patterns reached their peaks in the grand average (see Fig. 1a).

Source analysis was also performed by fitting regional sources (each source containing three orthogonal dipoles) to the collapsed ASSR and VSSR field patterns (128 electrodes) for each participant. Single sources were seeded in the auditory and visual cortices of each hemisphere and were constrained to be symmetric between the hemispheres. However, the 3D location of the symmetrical regional sources was determined only by the inverse solutions generated by BESA. Regional sources accounted for an average of 87.2% of the variance in the field patterns for the ASSR and 86.9% of the variance for the VSSR in the P1 baseline. Response amplitude

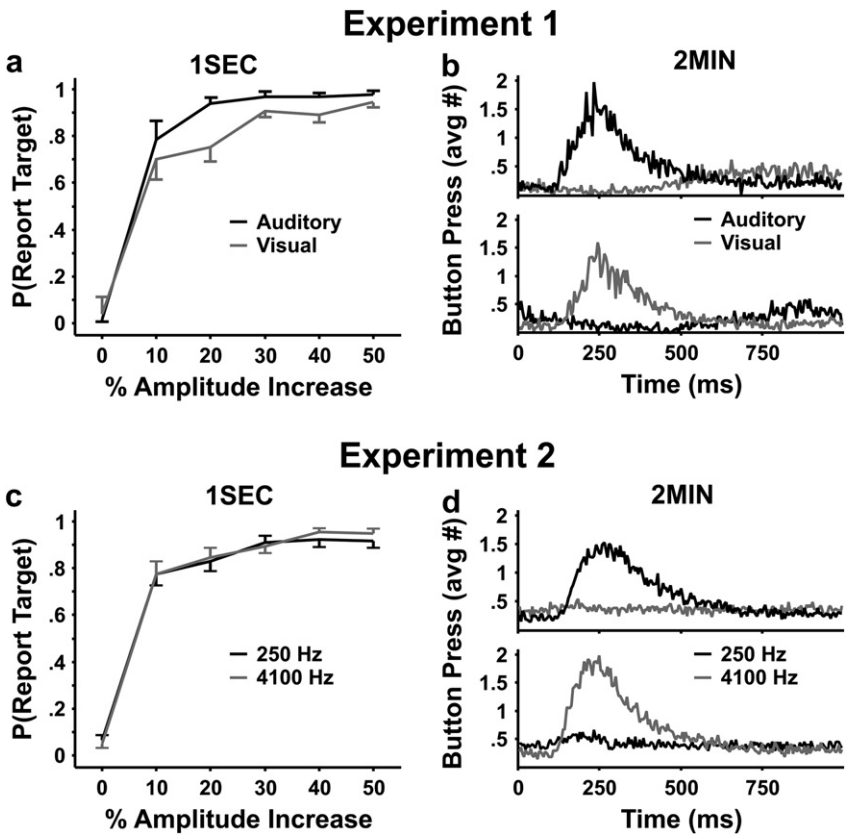

Fig. 2. Grand average behavioural responses in Experiments 1 (top panel) and 2 (bottom panel). (a) Group 1SEC, Experiment 1. The mean probability of reporting a target is shown across the five target values (hits) and when no target was present (false alarms, 0% amplitude increase), for the auditory (black) and visual (grey) tasks. Bars are 1 SE. (b) Group 2MIN, Experiment 1. Response latency distributions (number of button presses per 5 ms bin) are shown, calculated with respect to targets in the attended and unattended modalities (attend auditory black, attend visual grey). (c) Group 1SEC, Experiment 2. The mean probability of reporting a target is shown across the five target values (hits) and when no target was present (false alarms, 0% amplitude increase), for the attend 250 Hz (black) and attend 4100 Hz (grey) tasks. Bars are 1 SE. (d) Group 2MIN, Experiment 2. Response latency distributions (number of button presses per 5 ms bin) are shown, calculated with respect to targets in the attended and unattended frequency streams (attend 250 Hz black, attend 4100 Hz grey).

was measured for each steady-state response, hemisphere, and condition of attention (P1, attend A, attend V) as the dipole power associated with largest orthogonal vector in the regional source (this vector typically oriented toward FCz for the ASSR and Oz for the VSSR). Source analysis was conducted to assess hemispheric effects and to determine whether the 3D location of ASSR and VSSR generators changed from the passive baseline when attention was required. The analysis also provided a check on the extent to which attention effects that were observed at the field maxima for the ASSR and VSSR (electrodes FCz and Oz, respectively) appeared as well when field patterns were modeled using 128 electrodes. It should be noted that while a single regional source in each hemisphere is typically adequate to describe the ASSR, it is not a preferred model for the VSSR which exhibits a more complex time domain waveform with upper harmonics (Fig. 3a; see Di Russo et al., 2007). Notwithstanding this limitation, the approach permitted a comparison of the behaviour of the two sensory modalities when the same source model was used for each.

2.2.5. Statistical evaluation

Behavioural data were analyzed with nonparametric statistics (Group 1SEC) and by *t*-tests comparing response rates between attended and unattended distributions of response latency as described above (Group 2MIN). ASSRs and VSSRs were assessed for changes in amplitude (μV or nAm) and phase (time delay between the AM stimulus and response waveforms, in radians) across stages and between groups by ANOVA using Statistica (version 6.0). In infrequent cases where phase differences between stages for an individual participant were greater than π, phases were unwrapped by adding or subtracting 2π to minimize the phase differences, so that linear statistics could be employed. Greenhouse–Geisser corrections were applied to repeated measures with more than two levels. Significant main effects and interactions were described by pre-planned *t*-tests. Significance level was set at 0.05 (two-tailed).

3. Results

3.1. 1SEC group

3.1.1. Behavioral responses

Group averaged behavioural data are shown in Fig. 2a. The probability of a hit (P[H]) averaged 0.913 ($SD = 0.113$) when collapsed across targets in the auditory task and 0.884 ($SD = 0.089$) when similarly averaged for the visual task, and exceeded chance ($P = 0.5$) for all participants on both tasks ($P < 0.00001$, sign test). False alarm rates (P[FA], 0% amplitude increase) were low and averaged 0.012 ($SD = 0.005$) and 0.076 ($SD = 0.037$) for the auditory and visual tasks, respectively. These results indicate that while the auditory task was slightly easier than the visual task, performance on the tasks was similar, and participants attended to the auditory and visual stimuli as instructed. No change in performance was observed across the five blocks in either of the attend stages, and no differences were found for the order of the attend stages.

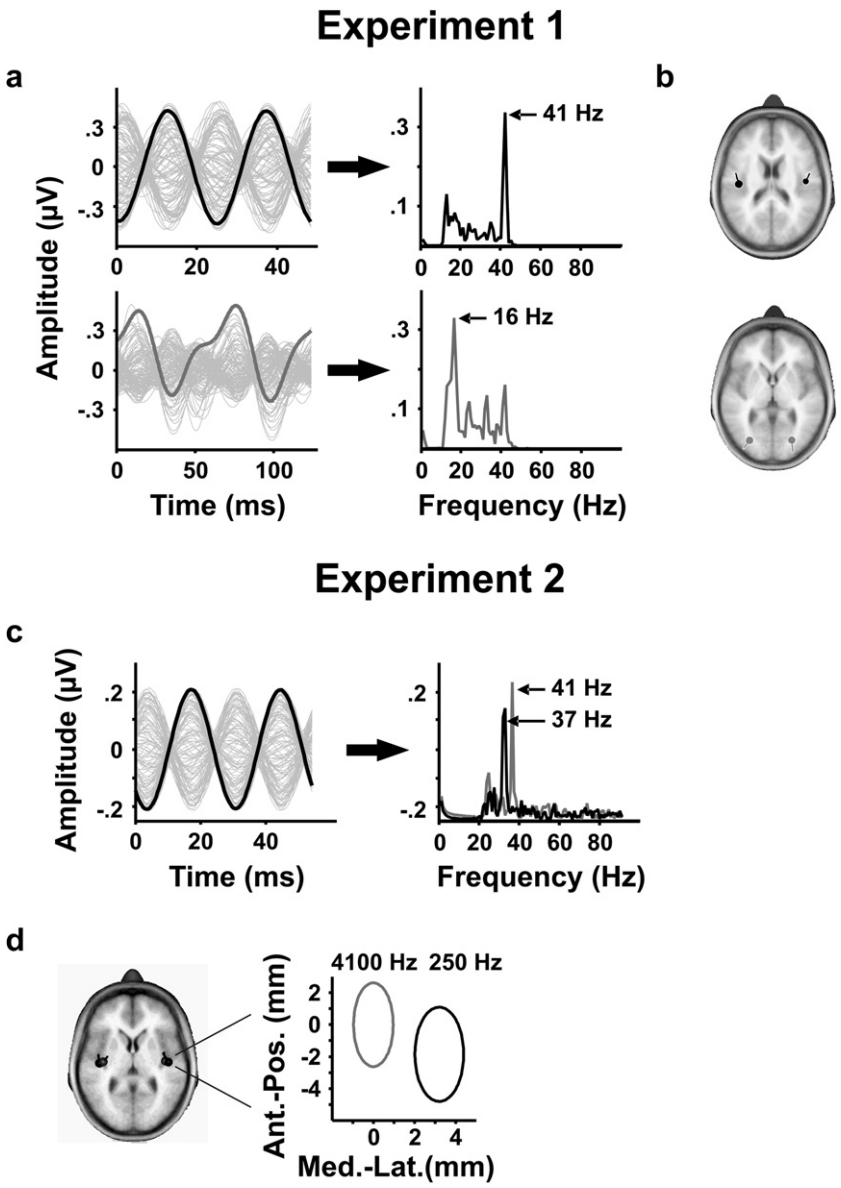

Fig. 3. Signal processing of steady-state responses in Experiments 1 (top panel) and 2 (bottom panel). (a) The left panel shows the time domain ASSR (top) and VSSR (bottom) collapsed to two pulses for a representative participant in the 1SEC group of Experiment 1 (128 electrodes). Electrode FCz is shown in thick black (ASSR) and electrode Oz in thick grey (VSSR). The right panel gives the amplitude spectra showing separable peaks at each modulation rate (16 and 41 Hz, from the highlighted electrodes) for this participant. (b) Regional sources fit to the two-pulse ASSR and VSSR waveforms for Experiment 1 (1SEC group, grand averaged coordinates). ASSR sources (top) localize medially to the region of primary auditory cortex, and the VSSR sources to the region of visual cortex. (The same analyses conducted for the 2MIN group were highly similar and are not shown.) (c) The left panel shows the 4100 Hz ASSR for a representative participant in the 1SEC Group of Experiment 2 (128 electrodes, FCz in thick black). The right panel gives the amplitude spectrum from the Fourier decomposition of the 4100 Hz waveform for this participant, showing a peak at this modulation rate (37 Hz, FCz electrode). The spectrum of the 250 Hz waveform AM at 41 Hz is overlayed on this spectrum. (Analyses of the 2MIN data gave similar results and are not shown.) (d) Source analysis of the 250 Hz and 4100 Hz ASSRs in Experiment 2. The locations of the sources for each ASSR are shown in the axial plane (average of solutions determined for each individual participant). The expansion shows the mean position of the 250 Hz source (black) relative to the mean 4100 Hz source (grey) in the axial plane (the cortical sources of the 250 Hz ASSR are lateral to those of the 4100 Hz ASSR). The ellipses are drawn through ±1 SE in the anterior-posterior and medial-lateral planes.

3.1.2. Steady-state responses

EEG data for a representative participant in Group 1SEC are presented in Fig. 3a. Modulation cycles for the ASSR (top left panel) and VSSR (lower left panel) were collapsed into two-pulse wide waveforms for each participant. Fourier decomposition of the collapsed data revealed distinctive signals (Fig. 3a, right panels) corresponding to the modulation rate for the ASSR and VSSR (40 Hz and 16 Hz respectively). The group averaged scalp voltage topography given in Fig. 1a showed strong foci consistent with the known distribution of bilateral sources in auditory and visual cortices, peaking near FCz for the ASSR and Oz for the VSSR.

For a group analysis, amplitude and phase were calculated from the 40-Hz and 16-Hz FFTs at electrode FCz for the ASSR and Oz for the VSSR. Analyses of ASSR and VSSR amplitude found no differences between the baseline P1 stage and the second passive stage (P2), or between the two presentation orders (A or V first). ASSR and VSSR amplitude during P1 was therefore subtracted from the corresponding measures in the attend stages A and V for each participant, to reference the responses to the P1 baseline and set attention effects into relief.

The results are shown in Fig. 4a. Main effects were found for Response (ASSR vs. VSSR, $F_{1,20} = 18.69$, $P = 0.0003$) and for

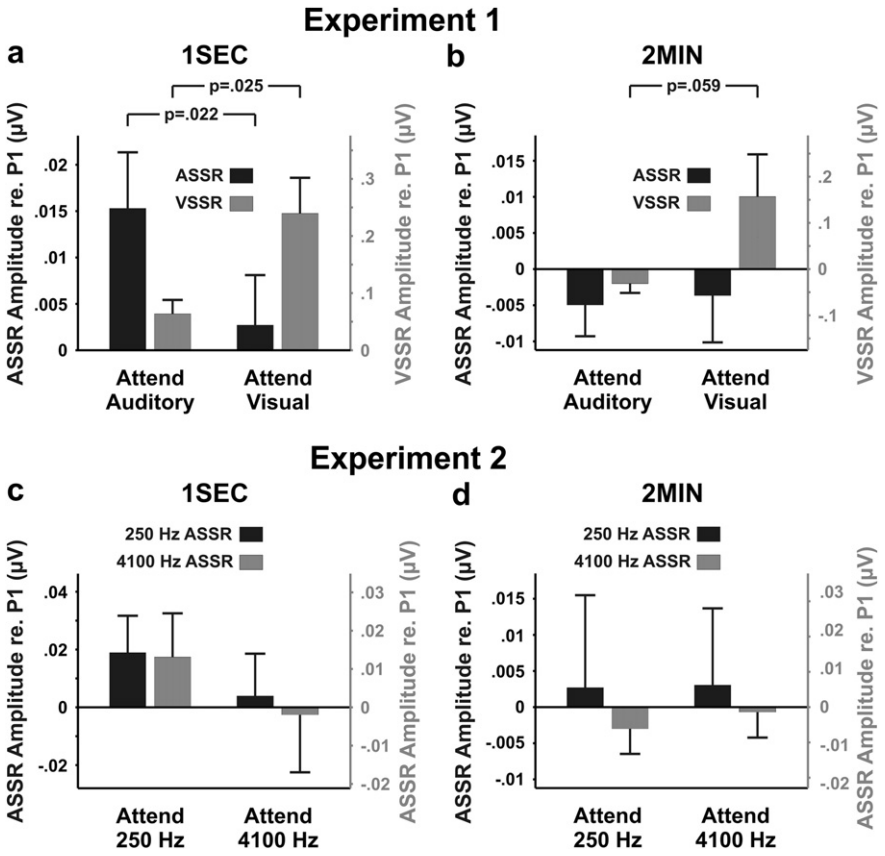

Fig. 4. Modality specificity (a, b; Experiment 1) and frequency specificity (c, d; Experiment 2) of the ASSR. (a) Group 1SEC, Experiment 1. ASSR and VSSR amplitude are shown as changes from the P1 passive baseline for each response, for the attend auditory (black) and attend visual (grey) stages. Error bars are 1 SE (within-subject). (b) Group 2MIN, Experiment 1. Same as in (a). (c) Group 1SEC, Experiment 2. 250 Hz ASSR amplitude (black) and 4100 Hz ASSR amplitude (grey) are shown as differences from the P1 baseline for each ASSR, for the attend 250 Hz (black) and attend 4100 Hz (grey) conditions. (d) Group 2MIN, Experiment 2. Same as in (c).

Attention (attend auditory or visual, $F_{1,20} = 5.0$, $P = 0.037$), which were attributable to changes in VSSR amplitude from the P1 baseline being larger overall than changes in ASSR amplitude (the ordinates were therefore scaled differently for the two responses in Fig. 4a). More importantly, an interaction of Attention with Response was found ($F_{1,20} = 6.60$, $P = 0.018$), revealing effects of modality-specific attention. ASSR amplitude increased from the P1 baseline when auditory attention was required ($P = 0.017$) and VSSR amplitude when visual attention was required ($P = 0.001$); in addition, auditory attention increased VSSR amplitude from the P1 baseline ($P = 0.016$, a modality non-specific effect) but ASSR amplitude did not change from baseline during visual attention ($P = 0.615$). Pre-planned contrasts applied to each response separately confirmed that response amplitude was larger for a given response (ASSR or VSSR) when the modality of that response was attended than when it was not, confirming the presence of both modality-specific auditory ($P = 0.022$) and modality-specific visual ($P = 0.025$) attention. These analyses were repeated for ASSR and VSSR phase. No main effects or interactions involving Stage (P1 vs P2), Order (A or V first), or Attention (A versus V) were found.

Results obtained at electrodes FCz and Oz were corroborated when ASSRs and VSSRs were modeled in source space. Source modeling confirmed generators localizing to auditory cortex for the 40-Hz ASSR and to visual cortex for the 16-Hz VSSR (see Fig. 3b). The 3D location of regional sources determined for each response and participant in the P1 baseline did not change in any coordinate (medial-lateral x, anterior-posterior y, or superior-inferior z) when attention to either modality was required. Dipole power for the P1 baseline was subtracted from dipole power in the attended stages, to set the effects of attention on ASSR and VSSR amplitude into relief. Main effects of Response (ASSR/VSSR, $F_{1,20} = 7.54$, $P < 0.012$) and Attention (attend auditory or visual, $F_{1,20} = 14.82$, $P < 0.001$) were found, which reflected larger VSSR responses overall occurring to attention. The interaction of Response with Attention was also significant ($F_{1,32} = 17.3$, $P < 0.0005$), pointing to modality-specific effects of attention. ASSR dipole power increased from baseline when auditory attention was required ($P < 0.007$) and VSSR power when visual attention was required ($P < 0.0003$), but attention to the other modality did not yield significant changes from baseline for either response. Pre-planned contrasts confirmed that VSSR power differed between the attend visual and attend-auditory conditions ($P < 0.0006$). This contrast did not reach significance for the ASSR, but it was significant ($P = 0.036$) when assessed only for participants ($n = 11$) for whom the variance explained by the regional source model exceeded the group average of 87.2%. Main effects and interactions involving hemisphere were not significant for ASSR or VSSR responses to attention, nor were any effects of attention or hemisphere found on ASSR or VSSR phase.

3.2. 2MIN group

3.2.1. Behavioral responses

Group averaged behavioural performance is shown for the 2MIN group in Fig. 2b. Response latency calculated with respect to targets in the attended modality gave orderly distributions, with peaks at 273 ms ($SD = 70.4$ ms) when auditory targets were attended and

280 ms ($SD = 51.9$ ms) when visual targets were attended. When response latency was calculated with respect to targets in the unattended modality, the distributions were random (flat). These results reveal strong modality-specific control of behavioural responding. Button presses measured at the peak latency during the attend stages were more numerous when a stream was attended than when it was not (attend auditory, $t_{12} = 6.42$, $P < 0.0001$; attend visual, $t_{12} = 6.30$, $P < 0.0001$). The overall rate of button pressing did not differ between A and V stages, and was not affected by presentation order (A or V first). Averaged over participants the mean ratio of total button presses in the latency window 150–500 ms to the total number of targets was 320/510 (62.7%) when attending to auditory targets 265/510 (52.0%)when attending to visual targets, which suggests that overall up to 57% targets were detected but 43% were missed.

3.2.2. Steady-state responses

ASSR and VSSR amplitude were determined for each subject by Fourier analysis of the collapsed two-pulse waveforms, as described in Fig. 3a for the 1SEC group (the group averaged waveforms were highly similar between the two timing conditions). No differences were found in ASSR or VSSR amplitude (measured at FCz and Oz respectively) between the baseline P1 and P2 stages or between the two presentation orders (attend A or V first). The P1 baseline was therefore subtracted from each response, to reference the responses to the P1 baseline and set effects of attention into relief.

The group-averaged results are presented in Fig. 4b. The main effect of Response was not significant in this analysis ($P > 0.19$). However, the main effect of Attention (attend auditory or visual, $F_{1,12} = 4.38$, $P = 0.058$) and the interaction of Attention with Response ($F_{1,12} = 4.27$, $P = 0.061$) were close to significance. When the responses were tested separately VSSR amplitude increased from the P1 baseline when visual attention was required ($P = 0.024$), but no other response changed from baseline during auditory or visual attention. A pre-planned contrast of VSSR amplitude between visual and auditory attention pointed to a modality-specific effect for vision ($P = 0.059$), but this contrast was not significant for the ASSR ($P > 0.81$). Analysis of ASSR and VSSR phase found no changes from baseline to the attention conditions, and no effect of modality-specific attention on this measure.

The results obtained when ASSR and VSSR amplitude were measured as dipole power followed those shown in Fig. 4b but were more robust for visual attention. The interaction of Response and Attention reached significance ($F_{1,12} = 7.65$, $P = 0.017$), reflecting enhanced VSSR power under visual attention ($P = 0.024$), but no other effect of attention was observed on the VSSR or ASSR. A contrast between visual and auditory attention was significant for the VSSR ($F_{1,12} = 7.86$, $P = 0.016$) confirming modality specificity for vision, while the same contrast for the ASSR was not significant ($P > 0.58$). The results for the ASSR did not change when only participants were considered for whom the variance explained by the source model exceeded the group average. No effects of attention or hemisphere were found on ASSR or VSSR phase.

3.3. Group comparisons

To directly assess the effects of timing requirements, we contrasted modality-specific attention between Groups 1SEC and 2MIN using brain responses recorded at FCz for the ASSR and Oz for the VSSR. ASSR amplitude during visual attention was subtracted from ASSR amplitude during auditory attention in each of the timing groups, and the effects compared between groups. This comparison approached significance for the ASSR ($t_{32} = 1.78$, $P = 0.084$), which reflected descriptively greater modality-specific auditory attention when stimulus duration was comparatively brief (1 s). The same calculation for VSSR amplitude was not close to significance ($P = 0.920$), indicating that stimulus duration had less influence on modality-specific visual attention.

To verify the presence of brief auditory-specific attention, a supplementary analysis was performed on the data of the 2MIN group. If participants in the 2MIN group were able to focus their auditory attention only for brief segments of time, this would presumably be most likely to occur coincident with detected auditory targets. An analysis was therefore performed with a new averaging epoch. A detected target was counted if a button press occurred within 150–500 ms after a target in the attended stream (A or V). This time window for counting detected targets was based on the averaged response latency distributions for the attended streams (Fig. 2b). The new averaging epoch was defined as the preceding 21 pulses for the ASSR (40-Hz = 513 ms) and the preceding 8 pulses for the VSSR (16-Hz = 500 ms) from the detected auditory or visual target. A control averaging epoch for the unattended stream was generated using the same time periods as the detected target epochs, but were slightly shifted in time (a few ms) in order to be aligned to the nearest onset of an AM pulse in the unattended stream. This generated two sets of ~500 ms averages, one for the ASSR and the other for the VSSR, for each of the attend auditory and attend-visual stages. These averages were collapsed to two pulses, and the amplitude and phase for each 16-Hz signal at Oz and 41-Hz signal at FCz were extracted by Fourier transform. Modality-specific effects on the ASSR were studied by contrasting ASSR amplitude when attention was directed to the auditory stream with ASSR amplitude when attention was directed away from it (to vision). A modality-specific effect on ASSR amplitude (ASSR amplitude during attend auditory larger than during attend visual) was now found ($t_{13} = 2.59$, $P < 0.05$), indicating that auditory attention of brief duration was detected over the 2-min trial in Group 2MIN. Evidence for modality-specific visual attention was also found for the VSSR (attend visual compared to attend auditory, $t_{13} = 2.15$, $P = 0.05$) which was similar in magnitude to modality-specific visual attention expressed over the entire 2-min trial in Group 2MIN. It should be noted that both of these effects reflected modulation of ASSR and VSSR amplitude by modality-specific attention and not by motor activity related to behavioural responding, because all steady-state responses preceded a motor act. Modality-specific effects on phase were not found for either the ASSR or VSSR.

4. Discussion

Participants in Experiment 1 were presented with simultaneous streams of 40-Hz AM auditory and 16-Hz AM visual stimulation and successfully complied with instructions to attend to targets presented in either the auditory or visual modalities. Because response requirements and the conditions of sensory stimulation were identical in the two attention conditions, differences between the conditions necessarily reflected the effects of top–down attention on auditory and visual brain responses.

Under these conditions ASSR amplitude increased from a preceding bimodal passive baseline when auditory attention was required. In addition, ASSR amplitude did not change from baseline during visual attention, indicating that behavioural responding had no effect on ASSR amplitude unless auditory targets were attended. These results indicate that auditory attention increased neural activity in the region of A1 where the generators of the ASSR are found, and that this effect was specific to the auditory modality. VSSR amplitude also increased from baseline when visual attention was required, more so than during auditory attention, revealing

modality-specific effects of visual attention on the cortical sources of this response. Because both of these modality-specific effects appeared as increases in ASSR or VSSR amplitude from a preceding unattended baseline, modulation of the responses could not be attributed alternatively to suppression of brain activity in one modality when the other was attended.

Experiment 1 also found evidence for an enhanced VSSR of small magnitude during auditory attention, but only in the unmodeled data, and when visual attention was required for 1 s not 2 min. This effect suggests a weak non-specific effect of auditory attention on VSSR amplitude. Compared to visual attention, modality-specific auditory attention was more evident when trials were 1 s long than when trials were 2 min long. However, modality-specific auditory attention was found in the 2MIN group when 500 ms epochs preceding button presses were analyzed.

Because the cortical sources of the ASSR localize to the region of A1, our findings imply that attention can modulate global neural activity in this region, at least for brief time intervals on the order of 500 ms. The neural mechanisms responsible for ASSR amplitude enhancement by top–down attention are at present unknown. However, a mechanism that could perform this function is the basal forebrain (BF) cholinergic system that projects to the auditory and visual cortices in a broadly tuned corticotopic arrangement (Bigl et al., 1982) and integrates bottom-up sensory input with top–down signals from the prefrontal cortex where task requirements are represented (Sarter et al., 2005). Cholinergic projections are paralleled by GABAergic projections targeting inhibitory synapses (Freund and Meskenaite, 1992) which may alter the balance of excitation and inhibition giving a synergistic effect. In Experiment 2 we asked whether top–down attention mediated by the BF attention system (or some other system) can also be *frequency specific*, facilitating neural activity in one tonotopic region of A1 and not another tonotopic region, when other task requirements are held constant.

5. Experiment 2

The design of Experiment 2 followed that of Experiment 1, except that participants directed their attention to one or the other of two streams of auditory stimulation that were distinguished only by carrier frequency, rather than between streams of auditory and visual stimulation.

5.1. Materials and methods

Unless otherwise stated, the methods for Experiment 2 were identical to those of Experiment 1. Only differences between the two studies are described here.

5.1.1. Participants

Thirty-nine McMaster University students aged 17–30 years (mean 22.0 years, 5 left-handed, 22 female) received course credit or $20 for participation. None had participated in Experiment 1.

5.1.2. Auditory stimuli

Two streams of auditory stimulation, consisting of two carrier frequencies (250 Hz and 4100 Hz) AM at either 36.57 Hz or 40.96 Hz (100% modulation depth, AM rate counterbalanced between participants across carrier frequency) were presented simultaneously to both ears. The two AM rates were chosen to be similar and perceptually confusable, whereas the two carrier frequencies (end points of the frequency-amplitude characteristic reported by Ross et al., 2000) were chosen to be highly distinct and distant from one another in the cortical tonotopic representation. These features ensured that subjects solved the task by attending to carrier frequency, and were intended to enhance the likelihood that frequency-specific attention could be detected. The 250 Hz carrier frequency was presented at a moderate intensity of 64 dB SPL and the 4100 Hz carrier frequency at 61 dB SPL, to ensure that the amplitude of each ASSR was near the center of the dynamic range of ASSRs recorded for these frequencies (Ross et al., 2000). Based on a matching study conducted with 13 other participants, 3 dB was added to the 250 Hz carrier at each AM rate in order equate subjective loudness between the carriers.

Auditory targets, consisting of an increase in the amplitude of single AM pulses in either the 250 Hz or 4100 Hz streams, were embedded within each stimulus. A staircase procedure was administered after an initial passive stage (P1) to determine the amplitude increase corresponding to the threshold of detection (TH) in each auditory stream. Two sets of targets (one set 250 Hz and the other 4100 Hz) were created (TH+10%, +20%, +30%, +40%, and +50%) most of which participants were expected to detect. On average targets corresponded to a TH of +43% in the A250 and A4100 stages. This level was similar to the TH of +45% which was used in the P1 and P2 stages (the latter determined from the matching study).

5.1.3. Design and procedure

Four stages were administered in a single session, in the order Passive 1 (P1), attend 250 Hz (A250), attend 4100 Hz (A4100), and Passive 2 (P2). Order of the middle attentive stages was counterbalanced across participants. As in Experiment 1, participants were randomly assigned to two groups that received different trial durations, Group 1SEC ($n = 15$) or Group 2MIN ($n = 24$).

In Group 1SEC two-thirds of the trials (determined randomly) contained a single 250-Hz or single 4100-Hz target, or one each of both targets. Targets in 37-Hz AM stimuli occurred randomly between 492 and 820 ms (pulses 18 to 30) and 41-Hz targets between 488 and 830 ms (pulses 20 to 34), with no constraints such that targets could overlap if present in both streams. Following each trial in the A250 and A4100 stages, participants were prompted by text on the computer screen "Did you hear a target?", and gave their answer with a mouse click (left for yes, right for no). The next trial commenced between 1.4 and 1.6 s later. In Stages P1 and P2 the same auditory stimuli were presented as unattended background sounds while participants focussed their vision on a fixation point (Fig. 1a) on the computer monitor. The ITI varied between 1900 and 2100 ms, which approximated the intertrial interval in the attend stages where behavioural responses were required.

Group 2MIN received a similar procedure, except that each "trial" now consisted of 2 min of continuous auditory stimulation. Five 2-min trials were presented in each stage, with a brief rest between trials. In each trial targets were presented continuously with a mean rate of 0.85 Hz in each of the 250-Hz and 4100-Hz streams, with no constraints on relative timing. The combined target rate of 1.7 Hz was designed to require a constant focus of attention on the designated stream (A250 or A4100). Participants were instructed to press a button after each target in the attended stream, and were reminded which stimulus to attend to with the text "Listen Low" or "Listen High" presented at the top of the screen.

5.1.4. Signal processing of EEG data

Averages were generated using the same procedure used in Experiment 1, with the addition that the raw data were high-pass filtered at 27 Hz for the 2MIN group to remove low frequency noise. Subject averages were filtered from 32 to 46 Hz (zero phase) to extract the 37-Hz and 41-Hz ASSRs. A two-cycle collapsed average was generated for each modulation rate over the interval 270–984 ms for 37-Hz AM rate, and 244–976 ms for 41-Hz AM rate, for Group 1SEC. In Group 2MIN the two-cycle average was

collapsed over the 2-min trial. These averages were analyzed by FFT to extract the amplitude and phase of the 250 Hz and 4100 HZ ASSRs corresponding to the 37-Hz and 41-Hz signals, at electrode FCz, for each stage of the experiment. Because the P2 stage was added in course, some participants did not receive this stage. Comparisons between P1 and P2 that are reported below were based on participants that received both stages ($n = 13$ in Group 1SEC and $n = 15$ in Group 2MIN).

Source analysis was employed to assess whether the 250 Hz and 4100 Hz carrier frequencies activated different generators in the region of A1. Two symmetrical regional sources were fit to each participant's average collapsed data, separately for stages P1, A250, and A4100, for the 250 Hz and 4100 Hz carrier frequencies. The 3D locations of these fits were recorded for the medial-lateral (x), anterior-posterior (y), and inferior-superior (z) coordinates for each participant, carrier frequency, and stage. Goodness of fit averaged 80.39% overall. Dipole power was also analyzed and compared to the results obtained when ASSR amplitude was measured at FCz.

6. Results

6.1. 1SEC group

6.1.1. Behavioural responses

Behavioural performance for the 1SEC group is shown in Fig. 2c. P[H] scores averaged 0.870 ($SD = 0.122$) across target intensities in the A250 stage and 0.884 ($SD = 0.111$) in the A4100 stage, with low false alarm rates (+0% amplitude increase) of 0.064 ($SD = 0.083$) and 0.050 ($SD = 0.070$), respectively. All participants had hit rates averaged across the five target values exceeding chance ($P < 0.00001$, sign test) indicating that the designated carrier frequency was attended as instructed. Performance was nearly identical between the 250 Hz and 4100 Hz streams and did not change across the five blocks in either attend stage (A250 or A4100) or differ between the two presentation orders (A250 or A4100 first).

6.1.2. Steady-state responses

Each subject's EEG data were collapsed into a two-pulse wide waveform and decomposed into 37 Hz and 41 Hz AM signals. The results for a representative subject in Group 1SEC are shown in Fig. 3c. Amplitude and phase were calculated for the 250 Hz and 4100 Hz ASSRs from their corresponding AM peaks in the FFT (electrode FCz) and grouped according to the A250 and A4100 conditions. No differences were found in ASSR amplitude or phase between the two presentation orders (A250 or A4100 first) or between stages P1 and P2.

The mean amplitude of the 250 Hz and 4100 Hz ASSRs was normalized for each subject by subtracting the corresponding P1 baseline amplitude, in order to set the effects of attention into relief. The results are reported in Fig. 4c. ASSR amplitude did not differ from zero in any condition, indicating no increase from the P1 baseline when attention to either auditory stream was required. Correspondingly, ASSR amplitude for a given stream (250 Hz or 4100 Hz) was not different when that stream was attended, than when it was not attended (the maximum t_{14} was 1.63, $P = 0.126$, reflecting a larger 4100 Hz ASSR when attending to 250 Hz). These findings did not change when this analysis was repeated only for participants whose mean hit rate (averaged cross the five target values) exceeded the group median. The same analyses performed for ASSR phase showed no changes from baseline to the attend conditions, and no effect of frequency-specific attention on this measure. Results from source analyses did not differ from those at electrode FCz.

6.2. 2MIN group

6.2.1. Behavioural responses

Behavioural performance for the 2MIN group is shown in Fig. 2d. Response latency calculated with respect to targets in the attended frequency gave an orderly distribution of latencies peaking at 284 ms ($SD = 73.2$ ms) for the A250 condition, and 250 ms ($SD = 60.4$ ms) for the A4100 condition, averaged over participants. Distributions of response latency calculated with respect to targets in the unattended frequency were random (flat). These results reveal a strong modulation of behavioural responding by the attended carrier frequency. Button presses measured at the peak latency during the attend stages were more numerous when the stream was attended than when it was not (A250, $t_{23} = 10.20$, $P < 0.0001$, A4100, $t_{23} = 17.80$, $P < 0.0001$). Button pressing was more frequent in the A4100 condition than in the A250 condition ($t_{23} = 3.30$, $P = 0.003$, measured at the peak latency during the attend stages) suggesting that 4100 Hz targets were somewhat easier to detect, but the differences between the conditions were not marked. Averaged over participants the mean ratio of total button presses in the latency window 150–500 ms to the total number of targets was 370/510 in the A250 condition and 415/510 in the A4100 condition, which suggests that overall up to 77% of the targets were detected but 23% were missed.

6.2.2. Steady-state responses

The continuous EEG data were collapsed into a two-pulse waveform and the amplitude and phase of the two ASSRs were extracted from their corresponding AM peaks in the FFT as described for Group 1SEC. Grand averaged ASSR amplitude calculated by subtracting the P1 baseline is shown in Fig. 4d for the A250 and A4100 stages. ASSR amplitude did not differ from zero in any stage, and did not vary as a function of the frequency-specific attention requirement (A250 or A4100). The frequency-specific contrast closest to significance was obtained for the 4100 Hz ASSR ($t_{23} = 1.30$, $P = 0.20$) which was suppressed below baseline when 250 Hz was attended. No effect of frequency-specific attention was found when these analyses were repeated using response amplitude unreferenced to the P1 baseline, or only for participants whose difference in the number of responses between attended and unattended streams at the peak latency of the attended stream exceeded the group median. Analyses performed for ASSR phase showed no changes from baseline to the attend conditions, and no effect of frequency-specific attention on this measure. Results from source analyses were similar to those at electrode FCz.

If participants in the 2MIN group were able to attend only for brief segments of time, this would presumably be most likely to occur coincident with detected targets. Although the results for Group 1SEC do not encourage this possibility, an additional analysis was performed following the method reported for Group 2MIN in Experiment 1. A detected target (hit) was counted when a button press occurred within 150–500 ms after a target in the attended stream, based on the group averaged response latency distributions for the attended streams (see Fig. 2d). An averaging epoch was then defined consisting of 21 pulses preceding the detected target (37-Hz 574 ms duration, 41-Hz 513 ms duration). Control epochs for the unattended stream were similarly generated using the same time periods as the detected target epochs, but were slightly shifted in time in order to be aligned to the nearest AM pulse in the unattended stream. This generated two sets of 540 ms averages, one from the attended stream and the other not, for each of the A250 and A4100 conditions. These averages were collapsed into two pulses and the amplitude and phase for the 250 Hz and 4100 Hz ASSRs were extracted from their corresponding AM rates at electrode FCz by FFT. No differences were found when attended and

unattended epochs were compared between the A250 and A4100 conditions, in either ASSR amplitude or phase.

6.2.3. Supplementary analyses

Supplementary analyses compared the properties of ASSRs recorded for the 250 Hz and 4100 Hz carrier frequencies with those reported in previous literature. These analyses were confined to the P1 stage in each of groups 1SEC and 2MIN.

When 40-Hz AM tones consisting of a single carrier frequency are presented for durations on the order of minutes, the phase of the resulting ASSR is known to vary with carrier frequency, such that the phase delay between the 40-Hz stimulus and response waveforms is shorter for high-frequency ASSRs than for low frequency ASSRs. This relation is attributed to traveling delays on the basilar membrane (Ross et al., 2000) which, when measured in auditory nerve fibers, are shorter for high-frequency pure tones transduced near the base of the membrane than for low frequency pure tones transduced near the apex (see Greenberg et al., 1998, for a review). Similarly, in our study ASSR phase was shorter for the 4100 Hz carrier than for the 250 Hz carrier by 95.7° ($t_{38} = 10.10$, $P < 0.0001$), which aligns with published data (Ross et al., 2000). The amplitude of the 250 Hz ASSR was also larger than the amplitude of the 4100 Hz ASSR ($t_{38} = 2.51$, $P = 0.016$), which also agrees with results reported by Ross et al. (2000) even though sound intensity was adjusted in our study to equate the perceived loudness of the two carriers. These relationships of ASSR phase and amplitude to carrier frequency were found in both the 1SEC and 2MIN groups, with no difference between the groups.

A final analysis compared the 3D location of the cortical generators of the 250 Hz and 4100 Hz ASSRs, which have been found to display a high-frequency medial, low-frequency lateral tonotopic organization in MEG studies (Pantev et al., 1996; Romani et al., 1982; Wienbruch et al., 2006). In these MEG studies carrier frequencies were presented one at a time, compared to simultaneous presentation of 250 Hz and 4100 Hz carriers at different AM rates in our EEG experiment. We performed source analyses on the collapsed pulse data for each subject in each group in our study, by fitting two symmetric regional sources for each frequency in the stages P1, A250, and A4100. In both groups a tonotopic ordering was found in each stage (see Fig. 3d), with the cortical sources of the 4100 Hz carrier frequency localizing medially with respect to those of the 250 Hz carrier frequency (main effect of ASSR, $F_{1,37} = 8.28$, $P = 0.007$, x-coordinate), in agreement with MEG data for carrier frequencies presented singly. The 3D location of ASSR sources did not change between P1 and the A250 or A4100 Hz stages, or differ between the groups, in any coordinate.

7. General discussion

The experiments of this paper utilized the 40-Hz ASSR known to localize tonotopically to the region of A1, to examine the properties of top–down attention in this area of the brain. Experiment 1 found that attention directed to auditory targets in a simultaneous stream of 40-Hz AM auditory and 16-Hz visual stimulation increased ASSR amplitude from a bimodal passive baseline, but that ASSR amplitude did not change from baseline when visual targets were detected, demonstrating *modality-specific* auditory attention. Experiment 2 assessed whether attention directed to one or the other of two simultaneous streams of auditory stimulation differing in carrier frequency could be *frequency specific,* selectively enhancing the ASSR evoked by an attended carrier frequency and its tonotopic sources in A1 compared to ASSRs evoked by an unattended carrier frequency. Although discriminative behaviour was strongly controlled by carrier frequency and the cortical sources of ASSRs evoked by the two carrier frequencies were resolved by inverse modeling, no evidence for frequency-specific modulation of ASSR amplitude was found.

In the following sections we discuss modality-specific auditory attention first and frequency-specific auditory attention second. In a concluding section we summarise principles suggested by the present and other findings that may describe how attention modulates activity in the region of A1.

7.1. Modality-specific auditory attention

In order to attribute modulation of a brain response to top–down attention, a comparison must be made between conditions of task-directed attention when all other aspects of the procedure (sensory stimulation, behavioural responses, and feedback) are held constant. This requirement was met by the procedures of Experiment 1, which not only revealed effects of modality-specific auditory attention on ASSR amplitude, but also found that behavioural responding had no effect on ASSR amplitude when the participants responded to attended visual rather than to auditory targets. The latter result is noteworthy, because it suggests that auditory attention was likely a key factor contributing to the ASSR enhancements reported in studies by Ross et al. (2004), Saupe et al. (2009), and Gander et al. (submitted for publication) in which participants performed behavioural responses when auditory attention was required but did not respond behaviourally in unattended control conditions. Our findings are also congruent with effects of attention on MLRs that are evoked 20–50 ms after stimulus onset and appear to contribute to ASSR generation. These responses increased in amplitude when attention was directed selectively to auditory targets that were presented dichotically (Woldorf et al., 1993) or embedded in bimodal auditory and visual stimuli (Poghosyan and Ioannides, 2008). In the latter study the cortical sources of MLRs evoked by different carrier frequencies localized tonotopically to the region of A1, showing the same high-frequency medial low-frequency lateral organization as ASSR sources in this region with which they likely overlap. These converging results appear sufficient to conclude that the earliest stages of auditory processing in A1 are modulated by top–down attention directed to auditory stimuli.

In Experiment 1 modality-specific auditory attention was observed on trials of 1 s duration (Group 1SEC) and on trials of 2 min duration (Group 2MIN), although in the latter group it was evident only when the ASSR was analyzed for brief time intervals (500ms) preceding target detection. In Group 2MIN button presses occurred within a 150–500 ms window after a target in the attended stream for up to 62% of the targets that were presented, suggesting that auditory attention was operating in a discrete fashion and that once a target was detected attention was reset and search recommenced. Modality-specific visual attention expressed in VSSR amplitude also appeared to operate discretely in the 2MIN group where 52% of the targets were detected, but unlike auditory attention, enhancement of VSSR amplitude was sustained over the 2-min trial. This difference between the modalities could reflect differences in the cortical sources contributing to 40-Hz ASSRs and 16-Hz VSSRs. While the 40-Hz AM rate chosen for auditory stimulation likely selected out cortical sources in A1 that responded with a latency of ~20–30 ms post-stimulus (brain responses in this latency range surviving time-locked averaging at 40 Hz), the 16 Hz AM rate chosen for visual stimulation most likely selected out an early visual AEP (the N75, latency ~65–75 ms) that appears to arise from two major cortical sources, one situated medially in primary visual cortex (V1) and the other more laterally in extrastriate regions (area V5) that may be attention sensitive (Di Russo et al., 2007). The latter source contributing to the VSSR from

nonprimary visual areas was not likely paralleled by nonprimary sources contributing to the ASSR.

In Experiment 1, but not in previous ASSR studies of modality-specific attention, ASSR amplitude during auditory attention was assessed with respect to a passive bimodal baseline administered before participants were informed of the auditory and visual attention tasks. ASSR amplitude increased from this baseline during auditory attention, and did not change from baseline during visual attention. These results rule out the alternative interpretation that increased ASSR amplitude during auditory attention reflected, not its enhancement by auditory attention, but its release from suppression during visual processing. The absence of evidence of auditory suppression by visual attention in our data differs from fMRI studies (Johnson and Zatorre, 2005, 2006; also see Laurienti et al., 2002) in which attention not only increased cortical activity in the attended auditory or visual modality compared to baseline (a result we observed in both modalities), but also suppressed activity in the unattended modality relative to baseline (a result we did not observe in either modality). One explanation of these results is that the ASSR reflects neural activity occurring in the region of A1, whereas in fMRI and PET studies the auditory areas affected most by auditory attention have been found in regions of auditory cortex (the planum temporale, planum polare, and lateral aspects of the STG) that are considered nonprimary areas that support more complex forms of auditory processing. Activity in nonprimary areas may be more easily modulated by intermodal competition than is neural activity in the primary sensory zones. It is also possible that suppression of ASSR amplitude by visual processing may be demonstrable under conditions that we did not test. In particular, had we measured ASSR amplitude under passive conditions when only auditory (and not bimodal auditory and visual) stimulation was present, and referenced attention effects to this baseline, we might have found evidence for suppression of ASSR amplitude by visual processing that was obscured by visual suppression during the bimodal baseline of the present study. Identification of possible reciprocal inhibitory effects expressed in these responses and the conditions under which they occur will require further research.

Electrophysiological studies (Bidet-Caulet et al., 2007; Brugge et al., 2009) and source localization data (Pantev et al., 1993, 1996) concur that the principal generators of the ASSR are found in the posteriormedial region of HG where A1 is located. However, source localization results reported by Gutschalk et al. (1999) gave evidence for a weak second source in ~17/30 hemispheres that was situated ~1 cm lateral to, and activated ~5 ms after, a dominant medial generator that was about two times larger and present in all hemispheres. Evidence for a second source in anterolateral HG (which might correspond to area RT of the auditory core in the model of Kaas and Hackett, 2000) can be also seen in intracortical recordings obtained by Bidet-Caulet et al. (2007) where attention was required, whereas Brugge et al. (2009) reported little activation of this area when 40-Hz ASSRs were recorded without task attention. Because participants in the Gutschalk et al. (1999) study were required to press button press when the AM rate changed, Gutschalk et al. suggested that the second weaker source they detected in their study may have been activated by attention. In this connection it is noteworthy that in Experiment 1 neither ASSR phase nor the 3D location of ASSR generators determined by source modeling changed from baseline when auditory attention was required, although ASSR amplitude increased during attention. Had a second source in anterolateral HG been activated by auditory attention, a change in ASSR phase (a lag) and a lateral shift in 3D location would have been expected with respect to baseline indicating a change in source configuration between these conditions, but no evidence for a change in these measures was found. Our results indicating that attention modulates ASSR amplitude with no effect on ASSR phase corroborates results from auditory training experiments reported by Gander et al. (submitted for publication). In those experiments ASSR amplitude and phase were modulated by attention and acoustic experience, respectively, suggesting their dependence on different underlying mechanisms.

7.2. Frequency-specific auditory attention

Experiment 2 tested for frequency-specific modulation of tonotopic regions in A1 following the method of Experiment 1, substituting simultaneous auditory/auditory stimulation for auditory/visual stimulation. Several conditions favourable to a test for frequency-specific auditory attention appear to have been met. The carrier frequencies (250 Hz and 4100 Hz) were separated by more than four octaves, in order to activate spatially distinct tonotopic regions in A1. This separation was also intended to avoid interactions among ASSRs that have been reported when carrier frequencies $\leq$3 octaves apart are presented simultaneously under passive conditions (Ross et al., 2003). The different AM rates (37 Hz and 41 Hz) associated with the carrier frequencies were perceptually confusable and were counterbalanced across participants with respect to carrier frequency, to ensure that group discrimination was based on carrier frequency and not AM rate. The duration of attention was manipulated between groups and included a level of this variable (1 s) found in Experiment 1 to characterize modality-specific auditory attention. The cortical sources of the 4100 Hz and 250 Hz ASSRs were spatially distinguished from one another in the medial-lateral axis, revealing separable sources in agreement with MEG investigations of the frequency (tonotopic) organization of the ASSR (Pantev et al., 1996; Romani et al., 1982; Wienbruch et al., 2006). Under these conditions we observed strong control of behavioural responding by carrier frequency, confirming that top–down attention directed to the two carrier frequencies was achieved. However, attentional modulation was not expressed in ASSRs localizing to the region of A1, but appears instead to have exercised its control at later stages of auditory processing.

In another recent study of selective attention in the auditory core region, Bidet-Caulet et al. (2007) investigated the effect of attending to one or the other of two ASSRs that were evoked by distinctive AM rates. ASSRs were recorded in ten epileptic patients by indwelling electrodes with multiple contacts placed in HG, the planum temporale, and the planum polare. Two AM sounds, one AM at 21 Hz and the other at 29 Hz, were presented simultaneously to both ears. The 21-Hz AM sound was comprised of two carrier frequencies separated by two octaves that varied randomly over trials, while the 29-Hz AM sound added a single carrier frequency intermediate to those of the 21-Hz AM sound. To enhance the distinctiveness of the two sounds and foster the perception of two auditory streams, the 21-Hz AM sound commenced first and the 29-Hz AM sound 810 ms later. Both sounds continued for an additional ~1 s, at which time on 50% of the trials one or both AM streams shifted their spatial position to the left or right ear through manipulation of interaural time delays. The subject's task was to attend to one of the two streams (AM 21 Hz or AM 29 Hz) and indicate with a joystick the direction of the shift (left or right) of changes in the attended stream at trial offset. ASSRs were observed bilaterally at focal sites predominantly in posteriormedial HG in both hemispheres. Of the ten participants tested, two were able to correctly indicate directional shifts in the 21 Hz and 29 Hz AM streams. These two participants showed larger ASSRs to the 29-Hz AM stream than to the 21-Hz stream when the 29-Hz AM stream was attended, in both hemispheres. Although the task is difficult, these results suggest that attention directed to auditory streams defined in part by distinctive AM rates can selectively activate neurons in the region of A1 that code for this sound feature.

Although the study of Bidet-Caulet et al. (2007) and Experiment 2 of this paper both investigated selective effects of top–down attention in the region of the auditory core, it should be noted that the results do not necessarily conflict with one another. On the contrary, they may be informative when taken together. Bidet-Caulet et al. (2007) defined their two auditory streams in terms of distinctive AM rates and differing temporal onsets, in order to foster the perception of two auditory objects. The two auditory objects (percepts) were further differentiated by the addition of a third carrier frequency with the onset of the 29 Hz AM sound, which produced a pitch change. This procedure was not designed to reveal selective modulation of different specific tonotopic regions in A1 by attention, because multiple carrier frequencies were associated with the two AM rates and these carrier frequencies varied randomly over trials. However, this does not mean that attentional facilitation of specific-frequency representations was not occurring in the procedure of Bidet-Caulet et al. In particular, selective attention to the two auditory streams may have fostered spectrotemporal integration of neurons coding AM rate with those coding for the carrier frequencies contained in the AM sound, thereby forging distinct auditory percepts in auditory association areas. Re-entrant feedback to the auditory core from auditory percepts attended at higher levels may have augmented the specific-frequency representations in A1 that were contributing to those percepts (Enns and Di Lollo, 2000). Re-entrant processing as a mechanism for selective attention in the auditory core could target A1 neurons tuned to frequencies in the percept without requiring highly specific projections to tonotopic regions from an attention system controlled independently by task demands. Re-entrant processing could also explain the findings of Paltoglou et al. (2009), who observed frequency-specific modulation of high and low frequency regions of HG in an fMRI study where participants attended selectively to high and low pitched sounds that were embedded in auditory streams distinguished by distinctive temporal patterns. In our study (Experiment 2) auditory streaming may have been difficult because the two auditory signals were distinguished only by carrier frequency and no other feature. Under these conditions inimical to auditory object formation, and in the absence of a top–down attention system targeting specific tonotopic zones, frequency-specific modulation of ASSR amplitude could not occur. Instead, behavioural performance may have been reliant on change detection mechanisms operating automatically in early auditory pathways that conveyed bottom-up information about target occurrence to higher brain regions where task attention directed behavioural responses. Frequency specific-auditory attention was clearly at work at some level of the projection pathway (our subjects performed well), but it did not modulate neural activity in specific tonotopic regions of A1 where the generators of the ASSR are found.

In Experiment 2 we presented the two auditory streams simultaneously to both ears, to ensure that any effect of attention would necessarily imply selective attention to carrier frequency and not to ear of presentation. Following a different approach, Müller et al. (2009) presented a 20 Hz AM sound to one ear and a 45 Hz AM sound to the other ear (AM rate counterbalanced between ears), both sounds consisting of a carrier frequency of 655 Hz. Participants were instructed to attend to one ear or the other ear and press a button when the AM rate changed within a 800 ms stimulus epoch (the probability of a change was 0.10). Subjects detected 74% of the changes independently of ear of presentation, indicating that they could perform the task. When the 20 Hz and 45 Hz AM ASSRs were separated in the frequency domain, ASSR amplitude in the left hemisphere was increased by attention to the right ear and suppressed by attention to the left ear, but modulation by attention to ear was found only for ASSRs evoked by the 20 Hz AM rate, and only in the left hemisphere. These findings suggest that ASSRs recorded from the region of A1 in the left hemisphere can be selectively modulated by attention directed selectively to one or the other ear, for AM rates near 20 Hz.

7.3. Summary and conclusion

The experiments discussed above suggest that top–down auditory attention can modulate A1 in a modality-specific fashion (Experiment 1 of this paper) and (under some conditions) in a specific hemisphere (Müller et al., 2009). Experiments by Bidet-Caulet et al. (2007) and Paltoglou et al. (2009) further suggest that frequency-specific modulation of A1 by attention may be achieved by re-entrant feedback into this region from spectrotemporal representations that are attended at higher levels of auditory processing. However, Experiment 2 of the current study suggests that top–down auditory attention does not directly modulate specific tonotopic regions in A1 in the absence of auditory object formation.

An additional unexpected result from Experiment 2 suggests a further general principle. In this experiment the amplitude of ASSRs evoked by the 250 Hz and 4100 Hz carrier frequencies did not increase from the passive baseline in our 39 participants (Groups 1SEC and 2MIN combined, $P > 0.75$) when attention to one or the other carrier frequencies was required. This contrasts with Experiment 1 of the present paper where increases from baseline were observed ($N = 21$, $P = 0.017$, Group 1SEC) when a single 2-kHz 40-Hz AM sound was used, and with two experiments reported by Gander et al. (submitted for publication, $N = 49$, $P < 0.00001$) which used the same single AM sound to assess effects of task attention. The ASSR amplitudes observed in each of these four studies were appropriate for their carrier frequencies and were evoked by moderate levels of sound intensity, suggesting that amplitude saturation was not likely a factor causing the different result in Experiment 2.

Although design constraints meant that the carrier frequencies used in these experiments were not always the same, the more critical difference may have been the addition of a second AM sound in Experiment 2. Evidence from animal studies (Kilgard et al., 2001) suggests that competitive interactions among frequency representations may normalize cortical activity in A1 under conditions of attention, even when the representations are distantly separated in the cochleotopic frequency map. This suggests that the effect of top–down auditory attention may be, not to increase global neural activity in A1 (although that may happen) or to operate selectively on complex maps found in this region (Cheung et al., 2001; Schreiner and Winer, 2007), but to expand the receptive fields of excitatory and inhibitory neurons in A1 (Fritz et al., 2007), fostering interactions among stimulus-driven signals and enabling integration of their features by attention-sensitive auditory association areas where complex auditory processing is performed (Johnson and Zatorre, 2005, 2006; Rauschecker and Tian, 2000). According to this view, re-entrant feedback from auditory percepts facilitated by attention at these higher levels may sculpt neural activity in A1, but attention may not operate selectively at this level, which is more directly concerned with forming neural representations for sounds that are present in the individual's environment.

Acknowledgements

This research was supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada (L.E.R.). The authors contributed equally to this work. P.E.G. is now at the University of Nottingham (UK).

References

Alho, K., Medvedev, S.V., Pakhomov, S.V., Roudas, M.S., Tervaniemi, M., Reinikainen, K., Zeffiro, T., Näätänen, R., 1999. Selective tuning of the left and right auditory cortices during spatially directed attention. Cogn. Brain Res. 7, 335–341.

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.

Bigl, V., Woolf, N.J., Butcher, L.L., 1982. Cholinergic projections from the basal forebrain to frontal, parietal, temporal, occipital, and cingulate cortices: a combined fluorescent tracer and acetylcholinesterase analysis. Brain Res. Bull. 8, 727–749.

Bosnyak, D.J., Roberts, L.E., 2001. Analysis of Swept-rate Auditory Stimulation Using Temporal Deconvolution Reveals No Departure from Linear Summation Hypothesis. Society for Neuroscience, San Diego.

Brugge, J.F., Nourski, K.V., Oya, H., Reale, R.A., Kawasaki, H., Steinschneider, M., Howard, M.A., 2009. Coding of repetitive transients by auditory cortex on Heschl's gyrus. J. Neurophysiol. 102, 2358–2374.

Cheung, S.W., Bedenbaugh, P.H., Nagarajan, S.S., Schreiner, C.E., 2001. Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. J. Neurophysiol. 85, 1732–1749.

Di Russo, F., Pitzalis, S., Aprile, T., Spitoni, G., Patria, F., Stella, A., Spinelli, D., Hillyard, S.S., 2007. Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. Hum. Brain Map. 28, 323–334.

Eckert, M.A., Kamdar, N.V., Chang, C.E., Beckmann, C.F., Greicius, M.D., Menon, V., 2008. A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. Hum. Brain Map. 29, 848–857.

Enns, J.T., Di Lollo, V., 2000. What's new in visual masking? Trends Cogn. Sci. 9, 345–352.

Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical evidence of multimodal integration in primate striate cortex. J. Neurosci. 22, 5749–5759.

Formisano, E., Kim, D.S., Di Salle, F., van de Moortele, P.F., Ugurbil, K., Goebel, R., 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40, 859–869.

Freund, T.F., Meskenaite, V., 1992. γ-aminobutyric acid-containing basal forebrain neurons innervate inhibitory interneurons in the neocortex. Proc. Natl. Acad. Sci. USA 89, 738–742.

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. 22, 186–203.

Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. Proc. Natl. Acad. Sci. USA 78, 2643–2647.

Gander, P.E., Bosnyak, D.J., Wolek, R., Roberts, L.E., 2007. Modulation of the 40-Hz auditory steady state response by attention during acoustic training. In: Cheyne, D., Ross, B., Stroink, G., Weinberg, H. (Eds.), New Frontiers in Biomagnetism. International Congress Series. Elsevier, Amsterdam, pp. 37–40.

Gander, P.E., Bosnyak, D.J., Roberts, L.E., under review. Acoustic experience but not attention modifies neural population phase in human primary auditory cortex. submitted for publication.

Godey, B., Schwartz, D., de Graaf, J.B., Chauvel, P., Liégeois-Chauvel, C., 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. Clin. Neurophysiol. 112, 1850–1859.

Grady, C.L., Van Meter, J.W., Maisog, J.M., Pietrini, P., Krasuski, J., Rauschecker, J.P., 1997. Attention-related modulation of activity in primary and secondary auditory cortex. Neuroreport 8, 2511–2516.

Greenberg, S., Poeppel, D., Roberts, T., 1998. A space-time theory of pitch and timbre based on cortical expansion of the cochlear traveling wave delay. In: Palmer, A., Rees, A., Summerfield, Q., Meddis, R. (Eds.), Psychophysical and Physiological Advances in Hearing. Whurr, London, pp. 293–300.

Gutschalk, A., Mase, R., Roth, R., Ille, N., Rupp, A., Hähnel, S., Picton, T.W., Scherg, M., 1999. Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clin. Neurophysiol. 110, 856–868.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Summerfield, Q., Palmer, A.R., Elliott, M.R., Bowtell, R.W., 2000. Modulation and task effects in auditory processing measured using fMRI. Hum. Brain Map. 10, 107–119.

Hari, R., Hämäläinen, M., Kaukoranta, E., Mäkelä, J., Joutsiniemi, S.L., Tiihonen, J., 1989. Selective listening modifies activity of the human auditory cortex. Exp. Brain Res. 74, 463–470.

Jäncke, L., Mirzazade, S., Shah, N.J., 1999. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. Neurosci. Lett. 266, 125–128.

Jäncke, L., Specht, K., Shah, N.J., Hugdahl, K., 2003. Focused attention in a simple dichotic listening task: an fMRI Experiment. Cogn. Brain Res. 16, 257–266.

Johnson, J.A., Zatorre, R.J., 2005. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. Cereb. Cortex 15, 1609–1620.

Johnson, J.A., Zatorre, R.J., 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. Neuroimage 31, 1673–1681.

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proc. Natl. Acad. Sci. USA 97, 11793–11799.

Kilgard, M.P., Pandya, P.K., Vazquez, J., Gehi, A., Schreiner, C.E., Merzenich, M.M., 2001. Sensory input directs spatial and temporal plasticity in primary auditory cortex. J. Neurophysiol. 86, 326–338.

Laurienti, P.J., Burdette, J.H., Wallace, M.T., Yen, Y.-F., Field, A.S., Stein, B.E., 2002. Deactivation of sensory-specific cortex by cross-modal stimuli. J. Cogn. Neurosci. 14, 420–429.

Linden, R.D., Picton, T.W., Hamel, G., Campbell, K.B., 1987. Human auditory steady-state evoked potentials during selective attention. Electroencephalogr. Clin. Neurophysiol. 66, 145–159.

Makeig, S.D., Galambos, R., 1982. Attending or discriminating 40 Hz tones induces phase-locked subharmonic resonances in EEG. Psychopharmacol. Bull. 18, 55–58.

Müller, N., Schlee, W., Hartmann, T., Lorenz, I., Weisz, N., 2009. Top–down modulation of the auditory steady-state response in a task-switch paradigm. Front. Hum. Neurosci. 3, 1–9.

Pantev, C., Elbert, T., Makeig, S., Hampson, S., Eulitz, C., Hoke, M., 1993. Relationship of transient and steady-state auditory evoked fields. Electroencephalogr. Clin. Neurophysiol. 88, 389–396.

Pantev, C., Roberts, L.E., Elbert, T., Ross, B., Wienbruch, C., 1996. Tonotopic organization of the sources of human auditory steady-state responses. Hear. Res. 101, 62–74.

Paltoglou, A.E., Sumner, C.J., Hall, D.A., 2009. Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention. Hear. Res. 257, 106–118.

Petkov, C.I., Kang, X., Alho, K., Bertrand, O., Yund, E.W., Woods, D.L., 2004. Attentional modulation of human auditory cortex. Nat. Neurosci. 7, 658–663.

Poghosyan, V., Ioannides, A.A., 2008. Attention modulates earliest responses in the primary auditory and visual cortices. Neuron 58, 802–813.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. USA 97, 11800–11806.

Regan, D., 1989. Human Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier, New York.

Romani, G.L., Williamson, S.J., Kaufman, L., 1982. Tonotopic organization of the human auditory cortex. Science 216, 1339–1340.

Ross, B., Borgmann, C., Draganova, R., Roberts, L.E., Pantev, C., 2000. A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J. Acoust. Soc. Am. 108, 679–691.

Ross, B., Draganova, R., Picton, T.W., Pantev, C., 2003. Frequency specificity of 40-Hz auditory steady-state responses. Hear. Res. 186, 57–68.

Ross, B., Picton, T.W., Herdman, A.T., Pantev, C., 2004. The effect of attention on the auditory steady-state response. Neurol. Clin. Neurophysiol. 2004, 22.

Ross, B., Herdman, A.T., Pantev, C., 2005. Stimulus induced desynchronization of human auditory 40-Hz steady-state responses. J. Neurophysiol. 94, 4082–4093.

Santarelli, R., Maurizi, M., Conti, G., Ottaviani, F., Paludetti, G., Pettorossi, V.E., 1995. Generation of human auditory steady-state responses (SSRs). II: addition of responses to individual stimuli. Hear. Res. 83, 9–18.

Sarter, M., Hasselmo, M.E., Bruno, J.P., Givens, B., 2005. Unraveling the attentional functions of cortical cholinergic inputs: interactions between signal-driven and cognitive modulation of signal detection. Brain Res. Rev. 48, 98–111.

Saupe, K., Widmann, A., Bendixen, A., Müller, M.M., Schröger, E., 2009. Effects of intermodal attention on the auditory steady-state response and the event-related potential. Psychophysiology 46, 321–327.

Schreiner, C.E., Winer, J.A., 2007. Auditory cortex mapmaking: principles, projections, and plasticity. Neuron 56, 356–365.

Talavage, T.M., Sereno, M.I., Melcher, J.R., Ledden, P.J., Rosen, B.R., Dale, A.M., 2004. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. J. Neurophysiol. 91, 1282–1296.

Tzourio, N., Massioui, F.E., Crivello, F., Joliot, M., Renault, B., Mazoyer, B., 1997. Functional anatomy of human auditory attention studied with PET. Neuroimage 5, 63–77.

Wienbruch, C., Paul, I., Weisz, N., Elbert, T., Roberts, L.E., 2006. Frequency organization of the 40-Hz auditory steady-state response in normal hearing and in tinnitus. Neuroimage 33, 180–194.

Woldorf, M.G., Gallen, C.C., Hampson, S.A., Hillyard, S.A., Pantev, C., Sobel, D., Bloom, F.A., 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc. Natl. Acad. Sci. USA 90, 8722–8726.

Woodruff, P.W.R., Benson, R.R., Bandettini, P.A., Kwong, K.K., Howard, R.J., Talavage, T., Belliveau, J., Rosen, B.R., 1996. Modulation of auditory and visual cortex by selective attention is modality-dependent. NeuroReport 7, 1909–1913.