

SCIENTIFIC REPORTS



OPEN

Neural correlates of distraction and conflict resolution for nonverbal auditory events

Hannah J. Stewart^{1,2,3,4}, Sygal Amitay^{2,5} & Claude Alain¹

In everyday situations auditory selective attention requires listeners to suppress task-irrelevant stimuli and to resolve conflicting information in order to make appropriate goal-directed decisions. Traditionally, these two processes (i.e. distractor suppression and conflict resolution) have been studied separately. In the present study we measured neuroelectric activity while participants performed a new paradigm in which both processes are quantified. In separate block of trials, participants indicate whether two sequential tones share the same pitch or location depending on the block's instruction. For the *distraction* measure, a positive component peaking at ~250 ms was found – a distraction positivity. Brain electrical source analysis of this component suggests different generators when listeners attended to frequency and location, with the distraction by location more posterior than the distraction by frequency, providing support for the dual-pathway theory. For the *conflict resolution* measure, a negative frontocentral component (270–450 ms) was found, which showed similarities with that of prior studies on auditory and visual conflict resolution tasks. The timing and distribution are consistent with two distinct neural processes with suppression of task-irrelevant information occurring before conflict resolution. This new paradigm may prove useful in clinical populations to assess impairments in filtering out task-irrelevant information and/or resolving conflicting information.

It is well accepted that attention is not a unitary phenomenon, but rather involves processes that may vary as a function of the task at hand^{1,2}. In auditory attention, internal goal-directed actions enable listeners to prioritize and select task-relevant sounds for further processing^{3,4}. In addition to enhancing processing of task-relevant stimuli, selective attention may also involve top-down suppression of task-irrelevant information to stay on task^{5–8}, which would prevent information overload. Auditory attention also often requires processing incongruency in relation to the task demand (i.e. conflict resolution). For instance, prior behavioural studies have shown that listeners take longer to respond to the word 'high' when presented at a low pitch⁹. In most everyday listening situations, these two processes may be required to occur concurrently for successful goal-directed action. Yet prior research has typically evaluated these two skills separately. For instance, neuroimaging studies using dichotic listening paradigm focuses primarily on early selection mechanisms¹⁰ while those using an auditory analogue of the Stroop task focus on late selection process and conflict resolution^{11,12}. In the present study, we use electroencephalography (EEG) to investigate the neural correlates supporting suppression of distracting irrelevant information and conflict resolution using a new auditory paradigm, the Test of Attention in Listening (TAiL)¹³.

Within TAiL the listener has to decide if the trial's two sequential tones were same or different regarding their features of frequency (attend-frequency task) or location (attend-location task). Through different combinations of the task-relevant and -irrelevant stimulus features (frequency and location) changing, and/or staying constant, the listener's auditory *distraction* by the task-irrelevant stimulus feature can be calculated, along with the listener's ability to deal with sensory *conflict* (i.e. when one of the stimuli features changes compared to both staying the same or changing)¹³.

Prior research using scalp-recording of visual event-related potentials (ERPs) have revealed a distracter positivity (Pd) modulation. This ERP modulation (i.e. component) has been found using mainly visual search

¹Rotman Research Institute, Baycrest Hospital & Department of Psychology, University of Toronto, Toronto, Canada. ²Medical Research Council Institute of Hearing Research, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK. ³School of Psychology, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK. ⁴Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, 45229, USA. ⁵Nottingham University Hospitals NHS Trust, Nottingham, UK. Correspondence and requests for materials should be addressed to C.A. (email: calain@research.baycrest.org)

paradigms where the target differs from the distracters in one feature (e.g. shape¹⁴; orientation¹⁵; colour¹⁶). It is thought to reflect the suppression of, and possibly the orientation to, irrelevant but potentially distracting lateralized stimulus^{14–16}. The Pd has an onset ranging from 250–300 ms post-stimulus with a positive parietal placement, contralateral to the presentation of the distracter¹⁴. The Pd appears to be similar to another ERP modulation, the PTC (i.e. a positive modulation over temporal scalp region contralateral to the distracter), which is also thought to index the suppression of distracter stimulus. The PTC peaks at about 290–370 ms after stimulus onset and is largest over the temporal scalp region contralateral to the lateralized distracter¹⁷. The small differences in latency and amplitude distribution between these two ERP modulations could be attributed to the salience of the distracter stimulus^{15, 18}. Further to these components, the rejection positivity (RP) has been used to describe a component in auditory-based studies akin to the Pd and PTC – with a frontocentral distribution occurring at 200–250 ms post-stimulus^{6, 7, 19}.

A popular method to assess conflict resolution has been the Stroop paradigm. Prior visual Stroop studies have revealed a frontocentral negative difference wave component, the N450^{20–23}. This component typically has an onset of 350 to 400 ms continuing to 500 ms with an origin in the medial dorsal area, specifically the anterior cingulate cortex²⁰. Comparatively, very few EEG studies have explored auditory Stroop tasks, at this time we are only aware of two such studies^{11, 12}. Both of these studies used verbal stimuli (e.g. “High” and “Low” or “Left” and “Right”) and found difference wave components similar in topography to visual Stroop tasks (frontocentral negativity). However, the ERPs presented earlier with an onset of 200 to 250 ms lasting until 500 ms and peaking at about 300 ms.

A later posterior sustained positivity (SP) was also identified at 500–800 ms for visual and auditory Stroop tasks^{11, 20–23}. It has been proposed that this component, SP, reflects additional supramodal processing, such as word meaning²⁰, or the control required to accurately respond to the trial²⁴. Therefore we expect to find this component coinciding with the *conflict resolution* measure.

Whilst the behavioural outcomes of TAI_L have consistently been shown across independent populations^{13, 25}, the underlying neural underpinnings have not yet been investigated. Understanding TAI_L's neural correlates may help break down where in auditory processing task instruction modulates performance. As currently, behavioural results suggest that by simply changing the TAI_L task instructions auditory attention processes change, despite the stimuli remaining constant²⁵.

TAI_L has been compared to other tests of attention using auditory and visual stimuli, namely the visual and auditory Attention Network Tests and the Test of Everyday Attention. Results indicate that TAI_L's attend-location task provides supramodal measures of selective attention, whilst the attend-frequency task provides auditory-specific measures²⁵. This follows the dual-pathway theory of processing spatial and non-spatial stimuli features in the dorsal and ventral streams, respectively (see review: refs 26, 27). Within this theory it is suggested that non-spatial features are processed in a modality-specific fashion, and spatial features in a supramodal fashion^{28–30}.

Results

Behavioural results. Repeated 2×2 ANOVAs (frequency: same, different; location: same, different) for each TAI_L task showed significant effects for the measures of *distraction* and *conflict resolution*. For comparison with prior research, conflict resolution and distraction measures were examined for attend frequency and attend location, separately.

Conflict resolution. In both tasks (Fig. 1A,B; Table 1), participants were faster when both stimulus features stayed the same or changed than when only one of the stimulus features changed (frequency \times location interaction: attend-frequency task: $F(1, 15) = 17.62, p < 0.001, \eta_p^2 = 0.54$; attend location task: $F(1, 15) = 28.16, p < 0.001, \eta_p^2 = 0.65$). Post-hoc analysis of this interaction shows slower response times (RTs) when the task-irrelevant feature changed than when it stayed constant in trials where the task-relevant feature also stayed constant (attend-frequency: $t(15) = -7.23, p < 0.001$; attend-location: $t(15) = -7.43, p < 0.001$), but not when the task-relevant feature changed (attend-frequency: $t(15) = -1.68, p = 0.11$; attend-location: $t(15) = -0.08, p = 0.94$).

Similarly for the accuracy data (Fig. 2A,B; Table 2), participants were more accurate when both features remained constant or changes than when only one feature changed (attend frequency: $F(1, 15) = 8.17, p = 0.012, \eta_p^2 = 0.35$; attend location: $F(1, 15) = 13.02, p = 0.003, \eta_p^2 = 0.47$). Post-hoc analysis showed that listeners were more accurate when the task-irrelevant feature changed than when it stayed constant in trials where the task-relevant feature also stayed constant (attend-frequency: $t(15) = -3.74, p = 0.002$; attend-location: $t(15) = -4.08, p = 0.001$). In the attend-location task it was found that listeners were more accurate when the task-irrelevant feature stayed constant than when it changed in trials where the task-relevant feature changed ($t(15) = 2.34, p = 0.034$). This effect was not found in the attend-frequency task ($t(15) = 0.23, p = 0.82$).

Distraction. In the attend-frequency task (Fig. 1A; Table 1), RTs were significantly slower for trials where the irrelevant location of the two tones changed than when they stayed constant ($F(1, 15) = 28.44, p < 0.001, \eta_p^2 = 0.66$). Participants were also less accurate (Fig. 2A; Table 2) when the location of the two tones differed ($F(1, 15) = 12.38, p = 0.003, \eta_p^2 = 0.45$). Similarly, in the attend location task, participants were slower (Fig. 1B; Table 1) ($F(1, 15) = 35.30, p < 0.001, \eta_p^2 = 0.70$) and less accurate (Fig. 2B; Table 2) ($F(1, 15) = 6.35, p = 0.024, \eta_p^2 = 0.30$) for trials where the irrelevant-feature of tone frequency changed between the two tones.

In summary, we found behavioral evidence of *conflict resolution* and *distraction* in both the attend-frequency and attend-location tasks (Figs 1C,D and 2C,D).

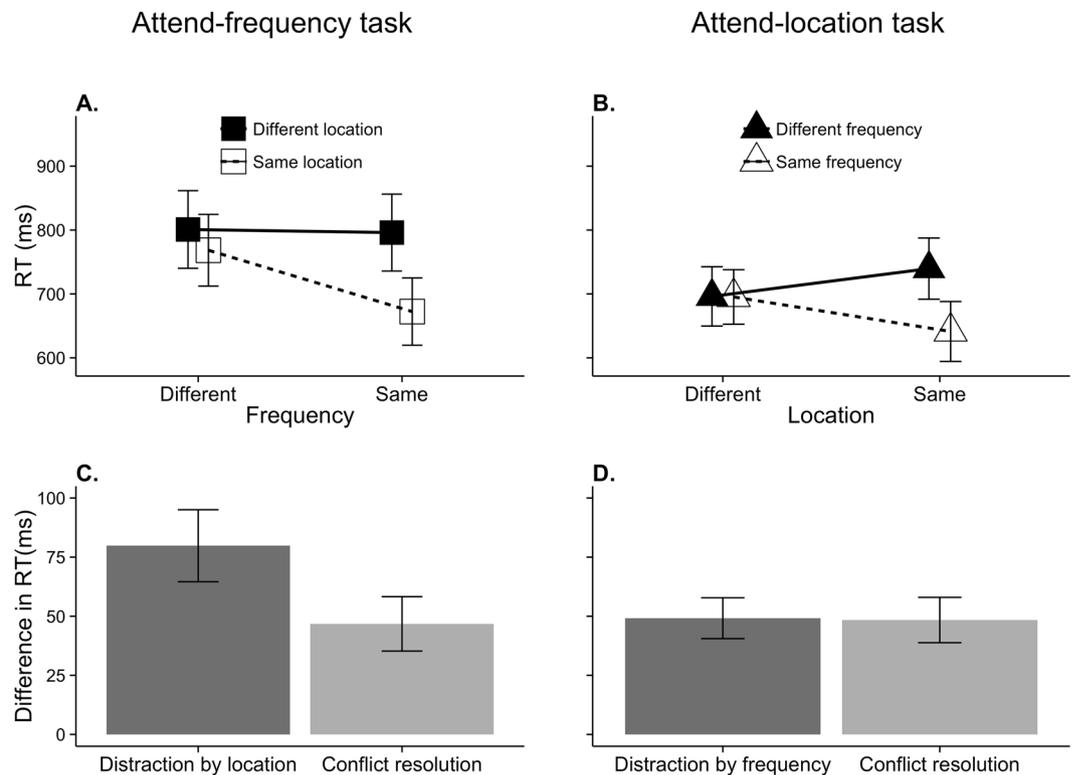


Figure 1. Group mean response times for the (A) attend-frequency and (B) attend-location TAI tasks. Group mean measures of distraction and conflict resolution for the (C) attend-frequency and (D) attend-location TAI tasks. The error bars represent the standard error of the mean.

	Attend-frequency task RT (ms)		Attend-location task RT (ms)	
	Mean	SEM	Mean	SEM
Different irrelevant feature	798.46	42.05	717.90	33.05
Same irrelevant feature	720.40	38.87	668.26	31.57
Incongruent	736.64	41.28	668.69	32.84
Congruent	782.22	40.55	717.46	31.61

Table 1. Group mean and standard error of the mean response times from both TAI tasks for different and same irrelevant feature changes, and incongruent and congruent feature changes.

Electrophysiological data. Figure 3 shows butterfly plots for the *baseline* measure from the attend-frequency and attend-location tasks. In both tasks, the tone pair generated N1 and P2 waves that were time-locked on sound onset. The iso-contour maps show that the N1 wave has a frontocentral scalp distribution and inverted in polarity at mastoid sites, consistent with generators in the superior temporal gyrus along the Sylvian fissure. The P2 wave has a more centro-parietal distribution whereas the slow wave showed a more frontal distribution, with a peak anterior to that of the N1 wave. The auditory evoked responses elicited by the first tone of the pair were little affected by task instructions. The effect of task began after the second tone was presented. All future ERP timings are with reference to the onset of tone 2 (T2) at 400 ms.

Conflict resolution. Figure 4 shows the conflict resolution, which is highlighted by contrasting ERPs for congruent trials (where both tone features stay constant or change) with those elicited by incongruent trials (where just one of the tone features change). In the attend frequency task, there was a significant difference, with incongruent trials generating more negative ERPs between 270 and 450 ms after T2 at frontocentral sites (Fig. 4A). In the attend-location task, the incongruent trials generated significant increased in negativity between 270 and 450 ms and again between 475 and 1100 ms over the right frontocentral sites (Fig. 4B).

Post-hoc analysis of this *conflict resolution* effect showed that trials where the task-irrelevant feature changed generated more positive ERPs than when it stayed constant. This effect was found when the task-relevant feature stayed constant in the attend-frequency until 995 ms after T2, peaking at 250 ms, and in the attend-location task until 1025 ms after T2, peaking at 260 ms (Supplementary Material Fig. 1A,B). The effect was also found when the

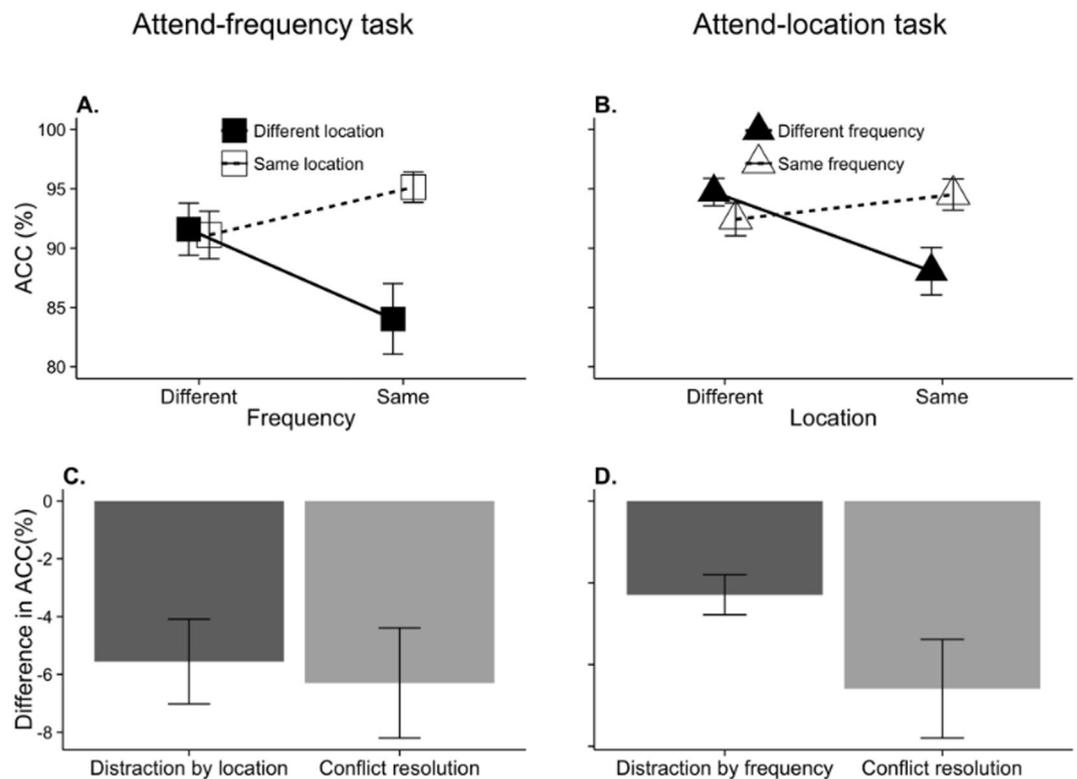


Figure 2. Group mean accuracy for the (A) attend-frequency and (B) attend-location TAI tasks. Group mean measures of distraction and conflict resolution for the (C) attend-frequency and (D) attend-location TAI tasks. The error bars represent the standard error of the mean.

	Attend-frequency task ACC (%)		Attend-location task ACC (%)	
	Mean	SEM	Mean	SEM
Different irrelevant feature	87.82	1.95	91.39	1.30
Same irrelevant feature	93.13	1.23	93.47	0.96
Incongruent	87.57	1.88	90.24	1.27
Congruent	93.37	1.30	94.62	0.86

Table 2. Group mean and standard error of the mean accuracy levels from both TAI tasks for different and same irrelevant feature changes, and incongruent and congruent feature changes.

task-relevant feature changed within the trial: between 170–300 ms (peaking at 250 ms) and 535–930 ms after T2 in the attend-frequency task, and between 190 and 680 ms (peaking at 280 ms) after T2 in the attend-location task (Supplementary Material Fig. 2A,B).

We compared the source activity from the frequency and location tasks using Brain Electrical Source Analysis (BESA) software (version 5.3). The analysis was conducted on the *conflict resolution* difference waves. For each participant and condition, we began with a regional source in the left and right temporal lobes near Heschl's gyrus. Then we optimize the location of both sources simultaneously. The ANOVA on source coordinates was not significant. This was expected given that the congruent trials included a mixture of pitch and location feature changes.

Distraction. Figure 5 shows the ERPs elicited when the task-irrelevant feature changed versus when it stayed constant. In both tasks, we found a significant increase in positivity at frontocentral sites ($p < 0.001$; Fig. 5). The polarity of this effect inverted at posterior sites (not shown). In both tasks, the difference wave shows an ERP ranging from around 200–300 ms after the onset of the second tone (T2), peaking at about 250 ms, with the attend-location task ever so slightly later than the attend-frequency task. Evidence from prior research suggests that processing sound identity and sound location engage to a greater extent the ventral and dorsal pathways, respectively²⁶. As for *conflict resolution* we compared the source activity from the frequency and location tasks using discrete source modeling. The analysis was conducted on the *distraction* difference waves. The group mean coordinates (Fig. 6) for each experimental condition were then compared using t-test. At a group level

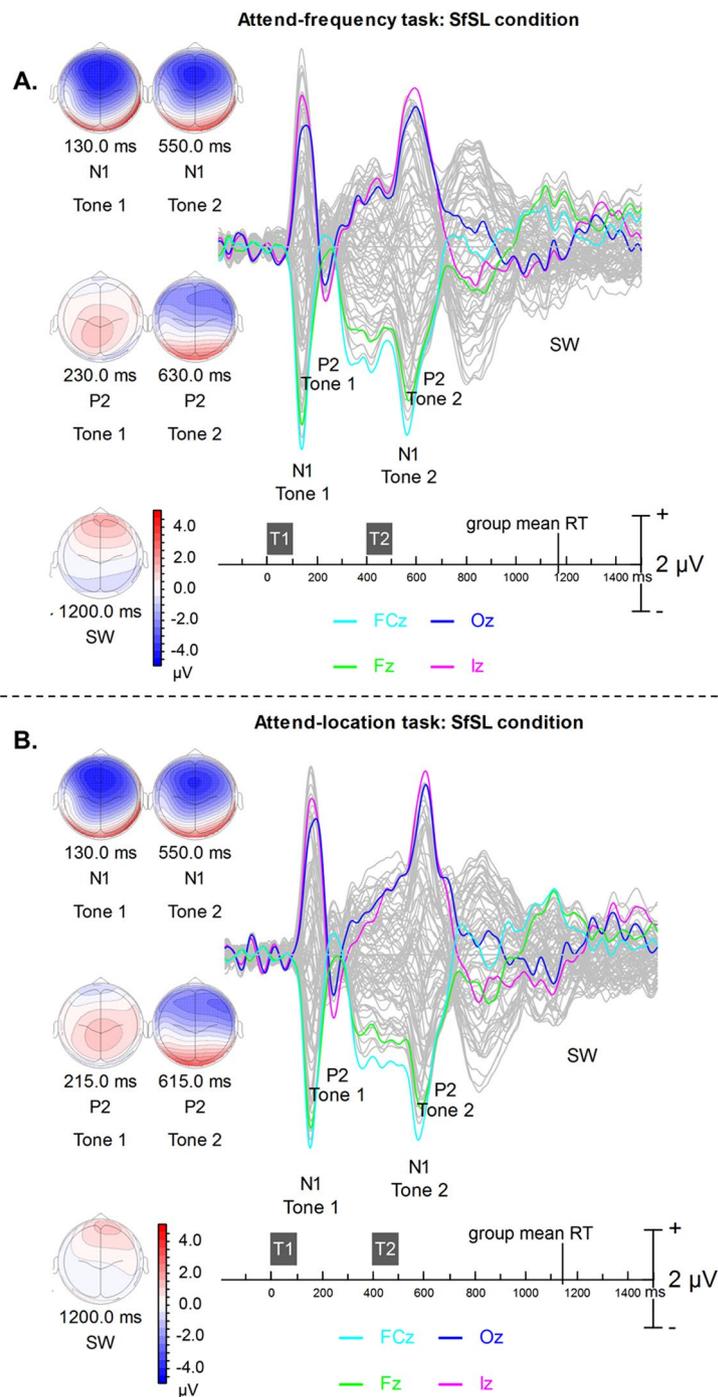


Figure 3. Group mean event-related brain potentials (for 65 channels) averaged over all SfSL trials from (A) attend-frequency and (B) attend-location TAIL tasks. The FCz electrode is shown in turquoise, Fz in green, Oz in purple and Iz in pink, all other channels are shown in gray. T1 = first tone; T2 = second tone. Isopotential contour maps are shown for the N1 (115 ms and 100 ms, respectively), P2 (200 ms and 200 ms, respectively) for each tone and slow wave (SW). Group mean RT for Same frequency Same location trials was 772 ms for the attend-frequency task and 741 ms for the attend-location task, time-locked to the onset of the second tone.

the source location for *distraction* by location (in the attend-frequency task) (mean \pm sem: $x = 33.60 \pm 5.08$; $y = -9.93 \pm 7.34$; and $z = 44.66 \pm 4.77$) was more medial and posterior than the source location for *distraction* by frequency (in the attend-location task) (mean \pm sem: $x = 3.86 \pm 9.67$; $y = 12.84 \pm 7.10$; $z = 47.63 \pm 6.43$). Paired t-tests on the x -, and y - were significantly different between the frequency and location tasks, but not for the z -axis (x : $t(14) = -2.27$, $p = 0.040$; y : $t(14) = -2.96$, $p = 0.010$; z : $t(14) = 0.58$, $p = 0.57$).

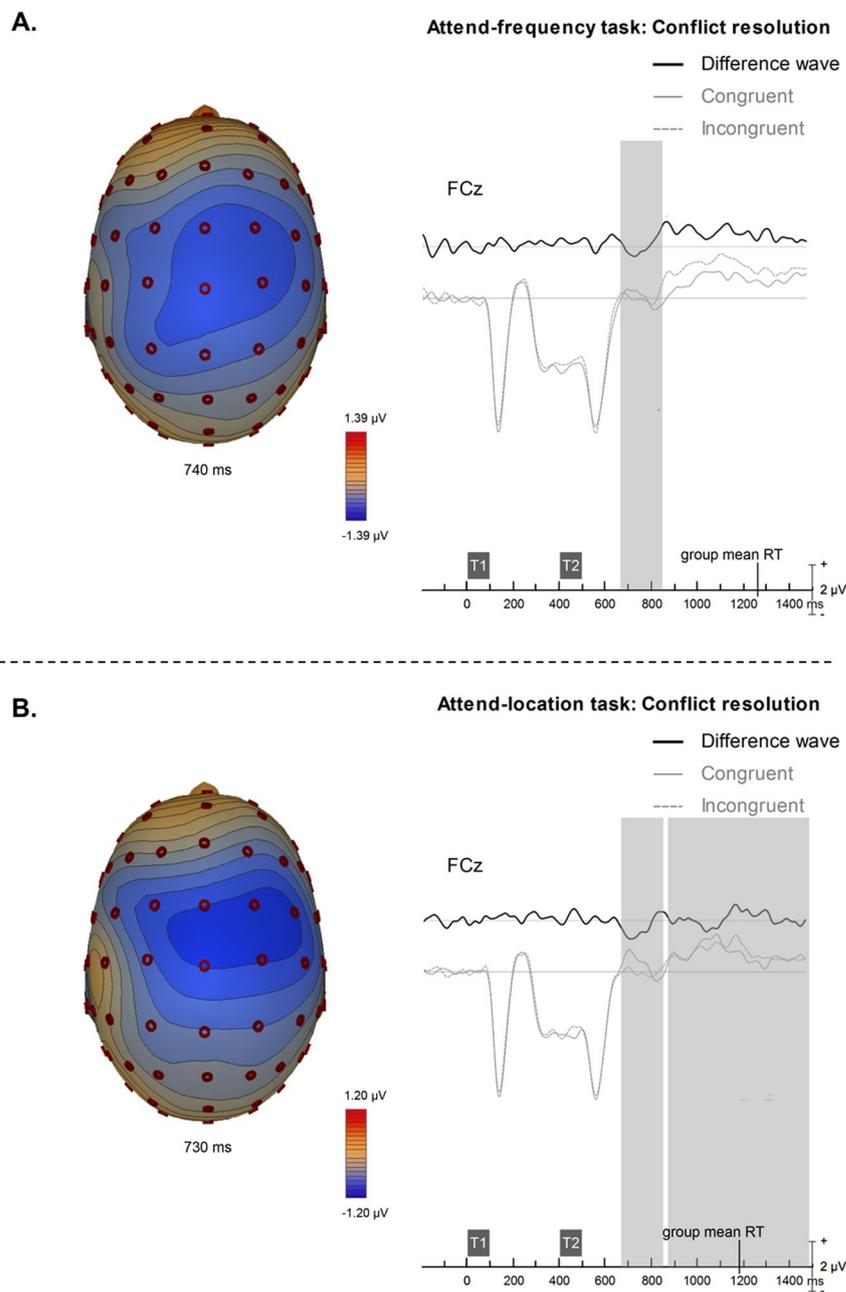


Figure 4. Conflict resolution TAIiL effect group mean event-related brain potentials recorded over the central frontal (FCz) scalp region for the (A) attend-frequency task and (B) attend-location task. T1 = first tone; T2 = second tone. Contour maps illustrate the brain activity at the peak of the difference waves. Shaded areas indicate the time period where the spatio-temporal cluster was significant, $p < 0.05$. For the attend-frequency task, the spatio-temporal cluster included the following electrodes: FCz, FC2, FC4, Cz, C2, C4, C6, CP2, CP4, CP6, Pz, P2, P4, and P6. For the attend-location, the spatio-temporal cluster included electrodes F4, FC1, FCz, FC2, FC4, FC6, C1, C2, C4, C6, CP2, CP4, CP6, and P4. Group mean RT across trials was 859 ms for the attend-frequency task and 793 ms for the attend-location task, time-locked to the onset of the second tone.

Discussion

In both the attend-frequency and attend-location tasks interference effects from the task-irrelevant tone feature were found on RT and accuracy. The EEG data showed that this main effect of *distraction* was qualified by the *conflict* resolution interaction. A change in the task-irrelevant feature significantly *distorted* the listeners regardless of whether the task-relevant feature changed or not within the trial. These behavioral effects coincided with a positive displacement that was best illustrated in the difference wave between ERP elicited by trials when the task-irrelevant tone feature changed and those obtained during trials when the task-irrelevant tone feature did not change. This ERP modulation was maximal at frontocentral sites, and peaked at about 250 ms after the onset of the second tone in TAIiL tasks. It shows similarity with the Pd and PTC component observed in visual selective

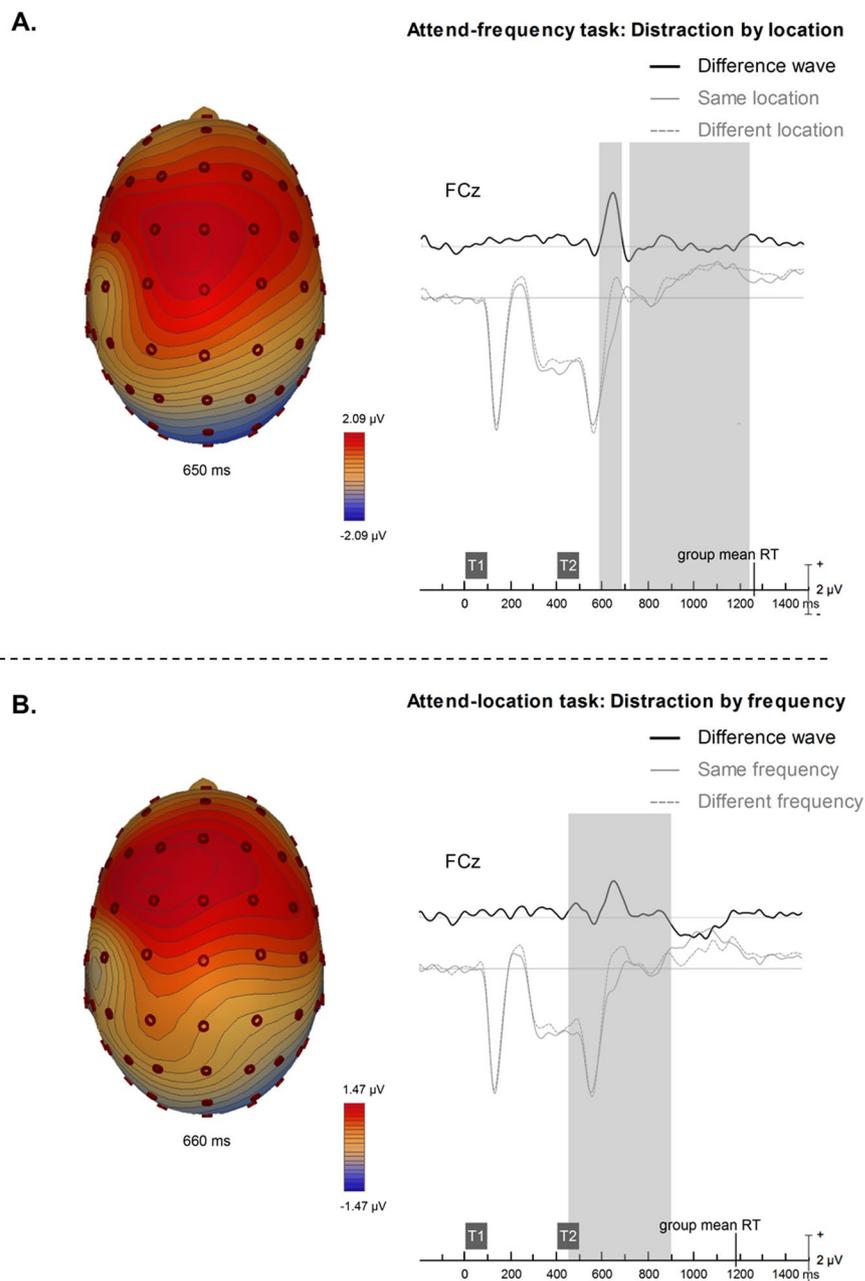


Figure 5. Distraction TAIL effect group mean event-related brain potentials recorded over the central frontal (FCz) scalp region for the (A) attend-frequency task and (B) attend-location task. T1 = first tone; T2 = second tone. Contour maps illustrate the brain activity at the peak of the difference waves. Shaded areas indicate the time period where the spatio-temporal cluster was significant, $p < 0.05$. For the attend-frequency task, the spatio-temporal cluster included the following electrodes: AF3, AFz, F5, F3, F1, Fz, F2, F4, F6, FC5, FC3, FC1, FCz, FC2, FC4, FC6, T7, C1, Cz, C2, C4, TP7, CP3, CP1, CPz, CP6, P9, P7, P4, P6, P8, P10, PO7, POz, PO4, PO9, O1, Oz, O2, and Iz. For the attend-location, the spatio-temporal cluster included the electrodes F1, Fz, F2, F4, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C1, Cz, C2, C4, C6, CP1, CP2. Group mean RT across trials was 859 ms for the attend-frequency task and 793 ms for the attend-location task, time-locked to the onset of the second tone.

attention tasks^{14–18}, but with an earlier latency. The observed ERP modulation also shows similarity with the RP component elicited by to-be-ignored sound information in auditory selective attention tasks^{6,7,19}. In the present study, the ERP modulation elicited by the distracter tone feature is thought to index the suppression, or even the rejection, of irrelevant stimulus feature when it is found not to match the attentional trace of the task-relevant stimuli.

In both visual and auditory tasks these *distraction* components have been shown to present in the area that would be enhanced if the distracter were the target¹⁸. For example, Degerman *et al.*⁶ found that during a

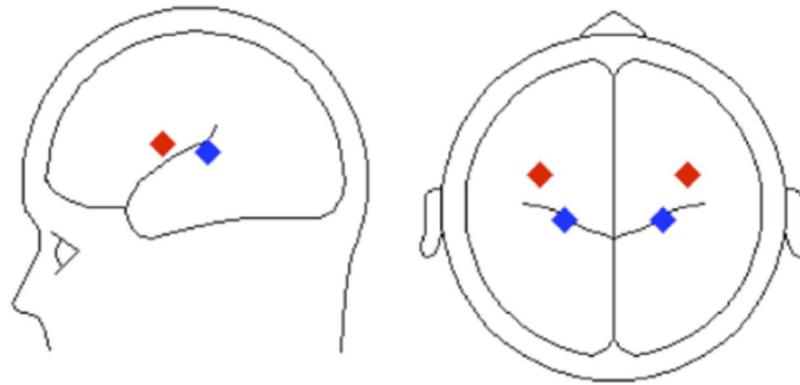


Figure 6. Dipole solution for distraction by location in the attend-frequency task (blue) (mean \pm sem: $x = 22.9 \pm 1.9$; $y = -25.3 \pm 5.4$; and $z = 13.5 \pm 3.0$) and distraction by frequency in the attend-location task (red) (mean \pm sem: $x = 34.7 \pm 1.9$, $y = 1.3 \pm 3.9$; $z = 20.1 \pm 1.9$).

continuous stream of auditory tones, when the listener was told to attend to sounds with a specific relevant feature (e.g. in the left or right ear or with high or low pitch), the RP component's source activity varied as a function of the irrelevant sound feature. This result was mirrored in the present study, with significantly different regional sources of the *distraction* component reflecting the task-irrelevant feature. That is, regional sources indexing the processing of the task-irrelevant location feature was more posterior in the attend-frequency task than the ones used to account for the task-irrelevant frequency feature effect in the attend-location task. The regional sources for the attend-frequency's task-irrelevant location feature were also more medial and inferior than the ones associated with the processing of the task-irrelevant frequency feature during the attend-location condition. These differences were unexpected and further research is needed to replicate these effects. The anterior-posterior difference in source location as a function of sound identity (pitch) and sound location is consistent with the 'what' and 'where' pathway model of auditory scene analysis^{26,30–32}, with the irrelevant location stimuli feature activating a more posterior network in the attend-frequency task than the irrelevant frequency feature in the attend-location task. In the present study, we purposely used regional sources to test the hypothesis that processing sound identity and sound location recruit different auditory areas along the superior temporal gyrus. Such an approach does not depend on identifying the exact source locations, but rather utilizes discrete source analysis as a mean to test current model of auditory processing. Further research using functional magnetic resonance imaging is needed to identify the brain regions or network of brain regions suppressing the distractor stimulus feature.

The RP *distraction* component, in this and other auditory selective attention studies, occurs within an earlier time frame than comparable visual selective attention studies^{6,11}. An earlier effect with auditory stimuli has also been shown in the Stroop task. Visual Stroop tasks typically report an onset of 350–400 ms with a peak at 450 ms post-onset of the stimuli^{20–23}. Comparably, auditory Stroop tasks show an earlier onset of 200–250 ms with a peak at around 300 ms^{11,12}.

In this study the *conflict resolution* component was found to occur between 270 and 450 ms in both TAI_L tasks. It may be that the auditory stimuli used in TAI_L causes earlier component onset than comparable studies in the visual modality as the time from stimulus onset to processing in the primary sensory cortex is markedly shorter in the auditory system when compared to the visual system³³.

However, a further negative frontocentral component was found to occur 475–1110 ms post second tone onset in the attend-location task's *conflict resolution* measure. Whilst the distribution of the *conflict resolution* component in the attend-frequency task is similar to that of previous auditory Stroop studies^{11,12}, the timing of this component in the attend-location task straddles the time frames reported in both the auditory and the visual modalities. Therefore these results provide continuing support to the behavioural results found by Stewart and Amitay²⁵ whereby TAI_L's attend-frequency *conflict resolution* measure focuses on auditory-specific non-spatial attention, whilst the attend-location *conflict resolution* measure assesses supramodal spatial based attention.

The design of TAI_L is advantageous firstly because the stimuli used in the two TAI_L tasks (attend-frequency and attend-location) are the same, just the instructions change, therefore reflecting purely cognitive processing differences rather than contrasting stimuli. Secondly, the underlying brain activation of the *distraction* and *conflict resolution* measures are calculated from the same blocks of trials allowing them to be directly compared. The results discussed in this paper suggest that through TAI_L's simple paradigm, it is possible to assess a listener's ability to deal with auditory-specific and supramodal sensory *conflict resolution* between relevant and irrelevant stimulus features. Furthermore, results indicate that the *distraction* measure does indeed result from the irrelevant, to be ignored stimuli, rather than the stimulus feature the listener is basing their decision upon.

The present study aimed to evaluate listeners' ability to suppress irrelevant sound information and to resolve conflict concurrently, as is often required in everyday listening environments. The results indicate that the TAI_L paradigm succeeds in assessing both of these skills simultaneously and display distinct neural correlates. In addition, by simply changing the task goal, via a change in instructions, the timings and distributions of the *distraction* and *conflict resolution* measures shift. The TAI_L paradigm therefore has potential use in clinical populations (e.g. children with Auditory Processing Disorder or older listeners with hearing difficulties) as a simple test to separate cognitive listening ability from sensory ability.

Method

Participants. Sixteen participants aged 18–25 (Mean = 22.25 years, SD = 2.26 years; 7 females and 9 males) were recruited through the Rotman Research Institute participant database. Inclusion criteria was normal hearing (thresholds below 25 decibel (dB) hearing level (HL) bilaterally at frequencies between 250 and 8000 Hz, inclusive, and a normal score (0–3) on the Quick Speech-in-Noise test (QuickSIN³⁴). Exclusion criteria was any self-reported history of brain damage, brain surgery, history of language-related or attention-related conditions, autism spectrum disorders or any auditory system disorders (e.g. Auditory Processing Disorder). Informed consent was signed by each participant prior to the experiment and each was paid at a rate of \$15 per hour. All methods were approved by, and performed in accordance with the relevant guidelines and regulations of, Toronto Academic Health Services Network.

Stimuli and Task. All tones were made up of sinusoids with a duration of 100 ms, gated on/off by 10 ms cos ramps, and were presented monaurally at 65 dB sound pressure level (SPL) using Etymotic ER-3A insert-earphones (Etymotic Research, Elk Grove, IL, USA). Tone frequency was randomly selected from the range 476.18–6187.50 Hz with at least 2.1 equivalent rectangular bandwidths (ERBs; ~4 semitones) between the trial's tones, therefore making it well within the listener's ability to discriminate between different frequencies³⁵. The inter-trial-interval varied randomly between 500 and 1000 ms, whilst the inter-stimulus interval (ISI) was set at 300 ms.

Listeners were asked to respond as fast and as accurate as possible after the second tone. Responses less than 200 ms and more than 2500 ms were excluded from further analysis in case of premature responding and lapses of attention or interruption of performance. As soon as the listener responded to each trial (attend-frequency task group mean: 859 ms; attend-location task group mean 793 ms) visual feedback was provided for 500 ms. If they answered correctly a smiley face was displayed. If they answered incorrectly the same face was shown with a sad expression.

TAiL stimuli were automated and presented in Matlab 2012b using PsychoPhysics Toolbox v3.0.10, a SoundMAX integrated digital HD Audio sound card. All behavioural responses were made using two buttons on the keyboard (the letters 'Q' and 'E') with the hand of the participant's choice. Listeners were tested individually in a sound-attenuated booth.

Regardless of the task type in TAiL, the stimuli and paradigm remained the same, just the instructions to the participant changed. In each trial, participants heard a tone pair where the individual tones were either the same or different in frequency and/or spatial location (ear presentation) (e.g. SfSL - same frequency and same location; DfSL - different frequency and same location; etc.). In both attend-frequency and attend-location the listener had to indicate via a button press if the task-relevant sound feature (i.e. the location of the two tones in the attend-location task) were same or different, whilst ignoring the task-irrelevant sound feature (i.e. the frequency of the two tones in the attend-location task). A full description of the task and stimuli can be found in Zhang *et al.*¹³.

From the attend-frequency and attend-location tasks measures of *baseline*, *conflict resolution* (the ability to deal with conflicting sound feature information) and *distraction* (involuntary orientation to the irrelevant sound feature) are calculated. The *baseline* measure is calculated for each task type from trials where both sound features, frequency and location, remain constant (i.e. SfSL trials). The *conflict resolution* is calculated as the difference between trials where the task-relevant and task-irrelevant features are incongruent and congruent (i.e. trials where the frequency and the location are both same/different minus trials where one feature is the same and the other different). The *distraction* measure is calculated as the difference in accuracy and response time between conditions where the task-irrelevant feature was either constant or changed between Tone 1 and Tone 2 (e.g. in the attend-location task – behavioral responses to trials where the frequency of the two sounds stayed the same was subtracted from trials in which the frequency changed).

Procedure. Both tasks (attend-frequency and attend-location) were repeated nine times. Each block included 40 trials, providing a total of 360 trials per task condition per participant. The order of the task types was alternated across the blocks, and counterbalanced across participants. For the first block of the attend-frequency and attend-location tasks there was a practice block prior. Each practice block involved five trials, accompanied by on-screen instructions. Participants had to reach 60% accuracy or more to move onto the full testing block. Instructions to define the task type and to ask the participant to be as fast and as accurate as possible were provided at the start of each block and visual feedback regarding the listener's performance was provided after each trial throughout. The total testing/recording time lasted ~45 minutes.

Behavioural analysis. RTs (ms) from correct trials, and accuracy (% correct) were used in the analysis. Repeated measure ANOVAs with the task-relevant and task-irrelevant dimensions as repeated measures were run for each TAiL task. Conflict resolution was the interaction between the relevant and irrelevant measures (i.e. comparing incongruent feature trials to congruent feature trials), and distraction was the main effect of the task-irrelevant measure (i.e. comparing different task-irrelevant feature trials to same task-irrelevant feature trials).

EEG recording and analysis. EEG was recorded from 66 scalp electrodes using a BioSemi Active Two acquisition system (BioSemi V.O.F., Amsterdam, The Netherlands). The electrode montage was according to the BioSemi electrode cap based on the 10/20 system and included a common mode sense active electrode and driven right leg passive electrode serving as ground. Ten additional electrodes placed below the hairline (both mastoid, both pre-auricular points, outer canthus of each eye, inferior orbit of each eye, two facial electrodes) to monitor

eye movements and to cover the whole scalp evenly. The latter is important because we used an average reference (i.e. the average of all scalp EEG channels as the reference for each EEG channel) for ERP analyses. Neuroelectric activity was digitized continuously at rate of 512 Hz with a bandpass of DC–100 Hz, and stored for offline analysis. All off-line analyses were performed using BESA (version 5.3; MEGIS GmbH, Gräfelfing, Germany).

The continuous EEG data were first digitally filtered with 0.1 Hz high-pass (forward, 6 dB/octave) and 20 Hz low-pass filters (zero phase, 24 dB/octave) using BESA software. The analysis epoch consisted of 200 ms of pre-stimulus activity and 1500 ms of post-stimulus activity time-locked to the onset of the first tone, from trials where the participant answered correctly. For each participant, a set of ocular movements was obtained prior to and after the experiment³⁶. From this set, averaged eye movements were calculated for both lateral and vertical eye movements as well as for eye-blinks. A principal component analysis of these averaged recordings provided a set of components that best explained the eye movements. The scalp projections of these components were then corrected from the experimental ERPs in order to minimize ocular contamination, using BESA software (version 5.3). Epochs contaminated by excessive deflections (greater than ± 120 μ V anywhere in the epoch) after correcting for ocular contaminations were excluded from the averages. For each participant, the remaining epochs were averaged according to electrode position and experimental condition (i.e., SfSL, SfDL, DfSL, DfDL). Each average was baseline-corrected with respect to the pre-stimulus interval. The group mean proportion of trials included in the ERP average for each TAiL condition was 84.27% (SD = 11.06%) in the attend-frequency task and 85.15% (SD = 8.81%) in the attend-location task. One male participant was excluded due to a high number of artifacts during recording.

BESA statistics 2.0 was used as a data-driven method to compare two conditions over all scalp regions up to 1500 ms post onset of the first tone. A two-stage analysis first computed a series of t-tests that compared the ERP amplitude between the conditions at every time point from 0 to 1500 ms after the onset of Tone 1. This preliminary step identified clusters both in time (adjacent time points) and space (adjacent electrodes) where the ERPs differed between the conditions. The channel diameter was set at 4 cm which led to around 4.03 neighbours per channel. We used a cluster alpha of 0.05 for cluster building. In the second stage of this analysis, all p-values reported for each data point were corrected for multiple comparisons using a Monte-Carlo resampling technique³⁷. The number of permutations was set at 1,000 with a 0.05 threshold.

We also performed regional source analyses using BESA software (version 5.3) on the *distraction* difference waves and compared source locations for the frequency and location tasks using student t-test. The source location was optimized for a 60 ms interval centered on the peak of the difference wave for each task. For each participant and condition, we began with a regional source in the left and right temporal lobes near Heschl's gyrus. During optimization, the location of the left and right regional source was forced to be symmetric. We used a four shell head model with head radius of 85 mm, and thickness for scalp, bone and cerebrospinal fluid of 6, 7, 1 mm, respectively. The relative conductivities were 0.33, 0.33, 0.0042, and 1 for brain, scalp, bone and cerebrospinal fluid, respectively.

References

- Alain, C., Arnott, S. A. & Dyson, B. In *The Oxford Handbook of Cognitive Neuroscience: Volume 1 Core Topics* (eds Ochsner, K. N. & Kosslin, S. M.) 215–236 (Oxford University Press, 2013).
- Michie, P. T. *et al.* An exploration of varieties of visual attention: ERP findings. *Cogn. Brain Res.* **7**, 419–450 (1999).
- Schröger, E., Marzecová, A. & Sanmiguel, I. Attention and prediction in human audition: A lesson from cognitive psychophysiology. *Eur. J. Neurosci.* **41**, 641–664 (2015).
- Ross, B., Hillyard, S. A. & Picton, T. W. Temporal dynamics of selective attention during dichotic listening. *Cereb. Cortex* **20**, 1360–71 (2010).
- Bidet-Caulet, A., Mikyska, C. & Knight, R. T. Load effects in auditory selective attention: evidence for distinct facilitation and inhibition mechanisms. *Neuroimage* **50**, 277–84 (2010).
- Degerman, A., Rinne, T., Särkkä, A.-K., Salmi, J. & Alho, K. Selective attention to sound location or pitch studied with event-related brain potentials and magnetic fields. *Eur. J. Neurosci* **27**, 3329–41 (2008).
- Michie, P. T., Solowij, N., Crawford, J. M. & Glue, L. C. The effects of between-source discriminability on attended and unattended auditory ERPs. *Psychophysiology* **30**, 205–220 (1993).
- Münte, T. F., Spring, D. K., Szycik, G. R. & Noesselt, T. Electrophysiological attention effects in a virtual cocktail-party setting. *Brain Res.* **1307**, 78–88 (2010).
- Roberts, K. L., Summerfield, A. Q. & Hall, D. A. Presentation modality influences behavioral measures of alerting, orienting, and executive control. *J. Int. Neuropsychol. Soc.* **12**, 485–492 (2006).
- Broadbent, D. E. Successive responses to simultaneous stimuli. *Quarterly Journal of Experimental Psychology* **8**, 145–152 (1956).
- Donohue, S. E. *et al.* Is conflict monitoring supramodal? Spatiotemporal dynamics of cognitive control processes in an auditory Stroop task. *Cogn. Affect. Behav. Neurosci.* **12**, 1–15 (2012).
- Buzzell, G. A., Roberts, D. M., Baldwin, C. L. & McDonald, C. G. An electrophysiological correlate of conflict processing in an auditory spatial Stroop task: the effect of individual differences in navigational style. *Int. J. Psychophysiol.* **90**, 265–71 (2013).
- Zhang, Y.-X., Barry, J. G., Moore, D. R. & Amitay, S. A New Test of Attention in Listening (TAIL) Predicts Auditory Performance. *PLoS One* **7**, e53502 (2012).
- Hickey, C., Di Lollo, V. & McDonald, J. J. Electrophysiological Indices of Target and Distractor Processing in Visual Search. *J. Cogn. Neurosci* **21**, 760–775 (2008).
- Hilimire, M. R., Mounts, J. R. W., Parks, N. A. & Corballis, P. M. Dynamics of target and distractor processing in visual search: evidence from event-related brain potentials. *Neurosci. Lett.* **495**, 196–200 (2011).
- Sawaki, R. & Luck, S. J. Capture versus suppression of attention by salient singletons: electrophysiological evidence for an automatic attend-to-me signal. *Atten. Percept. Psychophys.* **72**, 1455–1470 (2010).
- Hilimire, M. R., Mounts, J. R. W., Parks, N. A. & Corballis, P. M. Competitive interaction degrades target selection: An ERP study. *Psychophysiology* **46**, 1080–1089 (2009).
- Sawaki, R., Geng, J. J. & Luck, S. J. A common neural mechanism for preventing and terminating the allocation of attention. *J. Neurosci.* **32**, 10725–36 (2012).
- Alho, T. K., Reinikainen, K., Sams, M. & Näätänen, R. Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* **68**, 458–470 (1987).

20. Liotti, M. *et al.* An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia* **38**, 701–11 (2000).
21. Markela-Lerenc, J. *et al.* Prefrontal-cingulate activation during executive control: Which comes first? *Cogn. Brain Res.* **18**, 278–287 (2004).
22. West, R. & Alain, C. Event-related neural activity associated with the Stroop task. *Cogn. Brain Res.* **8**, 157–64 (1999).
23. West, R. Neural correlates of cognitive control and conflict detection in the Stroop and digit-location tasks. *Neuropsychologia* **41**, 1122–1135 (2003).
24. Larson, M. J., Kaufman, D. A. S. & Perlstein, W. M. Neural time course of conflict adaptation effects on the Stroop task. *Neuropsychologia* **47**, 663–670 (2009).
25. Stewart, H. J. & Amitay, S. Modality-specificity of Selective Attention Networks. *Front. Psychol.* **6** (2015).
26. Arnott, S. R., Binns, M. A., Grady, C. L. & Alain, C. Assessing the auditory dual-pathway model in humans. *Neuroimage* **22**, 401–408 (2004).
27. Alho, K., Rinne, T., Herron, T. J. & Woods, D. L. Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies. *Hear. Res.* **307**, 29–41 (2014).
28. Driver, J. & Spence, C. Cross-modal links in spatial attention. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* **353**, 1319–1331 (1998).
29. Turatto, M., Benso, F., Galfano, G. & Umiltà, C. Nonspatial attentional shifts between audition and vision. *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 628–639 (2002).
30. Rauschecker, J. P. & Tian, B. Mechanisms and streams for processing of ‘what’ and ‘where’ in auditory cortex. *Proc. Natl. Acad. Sci. USA* **97**, 11800–6 (2000).
31. Mishkin, M., Ungerleider, L. G. & Macko, K. A. Object vision and spatial vision: Two cortical pathways. *Trends Neurosci.* **6**, 414–417 (1983).
32. Goodale, M. A. & Milner, A. D. Separate visual pathways for perception and action. [Review] [61 refs]. *Trends Neurosci.* **15**, 20–5 (1992).
33. Hillyard, S. A. Electrical and magnetic brain recordings: contributions to cognitive neuroscience. *Curr. Opin. Neurobiol.* **3**, 217–224 (1993).
34. Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J. & Banerjee, S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* **116**, 2395–2405 (2004).
35. Jensen, J. K. & Neff, D. L. Development of Basic Auditory Discrimination in Preschool Children. *Psychol. Sci.* **4**, 104–107 (1993).
36. Picton, T. W. *et al.* The correction of ocular artifacts: A topographic perspective. *Clin. Neurophysiol.* **111**, 53–65 (2000).
37. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).

Acknowledgements

This research was supported by grants from the Canadian Institutes of Health Research (MOP 106619), the Natural Sciences and Engineering Research Council of Canada (NSERC) to C.A., an Erasmus Mundus Exchange Network in Auditory Cognitive Neuroscience award to H.J.S. and Medical Research Council, UK grant (Grant U135097130) to H.J.S. and S.A. Reprint requests should be sent to Claude Alain, Rotman Research Institute, Baycrest Centre, 3560 Bathurst Street, Toronto, Ontario, Canada, M6A 2E1 or via email: calain@research.baycrest.org.

Author Contributions

H.J.S., S.A. and C.A. designed this study. Data collection was conducted by H.J.S. H.J.S. and C.A. analyzed the data and wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-00811-7](https://doi.org/10.1038/s41598-017-00811-7)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017