

Effects of visual attentional load on low-level auditory scene analysis

BENJAMIN J. DYSON

University of Sussex, Brighton, England

CLAUDE ALAIN

*Rotman Research Institute, Toronto, Ontario, Canada
and University of Toronto, Toronto, Ontario, Canada*

and

YU HE

Rotman Research Institute, Toronto, Ontario, Canada

The sharing of processing resources between the senses was investigated by examining the effects of visual task load on auditory event-related brain potentials (ERPs). In Experiment 1, participants completed both a zero-back and a one-back visual task while a tone pattern or a harmonic series was presented. N1 and P2 waves were modulated by visual task difficulty, but neither mismatch negativity (MMN) elicited by deviant stimuli from the tone pattern nor object-related negativity (ORN) elicited by mistuning from the harmonic series was affected. In Experiment 2, participants responded to identity (*what*) or location (*where*) in vision, while ignoring sounds alternating in either pitch (*what*) or location (*where*). Auditory ERP modulations were consistent with task difficulty, rather than with task specificity. In Experiment 3, we investigated auditory ERP generation under conditions of no visual task. The results are discussed with respect to a distinction between process-general (N1 and P2) and process-specific (MMN and ORN) auditory ERPs.

A central tenet within current models of attention is the appeal to separate pools of processing resources for different senses (Duncan, Martens, & Ward, 1997). The existence of independent resources reserved for specific senses is consistent with the observation that stimulus processing in one sense (e.g., audition) is little affected by task demands in a second sense (e.g., vision; Alho, Woods, & Algazi, 1994; Alho, Woods, Algazi, & Näätänen, 1992). However, one problem with the idea of independent processing resources across senses is that it can rest upon the confirmation of a null hypothesis. For example, in the absence of observed changes in auditory stimulus processing as a result of visual task manipulation, it is always possible that the “harder” visual task was simply not demanding enough to have a significant effect upon processing resources in the other domain (Valtonen, May, Mäkinen, & Tiitinen, 2003). Moreover, the observation of resource independence between vision and audition in previous studies may be attributable to the use of discrete, rather than continuous, stimulus pres-

entation, thereby allowing for attentional switching between the senses during a block of trials (Otten, Alain, & Picton, 2000).

In a further attempt to elucidate the extent to which similar resources are utilized for stimuli presented in different senses, in the present experimental series, we examined the effects of visual attentional load on the processing of auditory events via the recording of human event-related potentials (ERPs). Early neuroelectric studies into the effects of cross-sensory attention, using somatosensory (i.e., shock) and visual stimuli (Spong & Lindsley, 1968) or shock and auditory stimuli (Satterfield, 1965) have shown an attenuation of neural responses for sensory input outside the current focus of attention but were often limited in terms of stimulus presentation (see Näätänen, 1975, for an early review). Nevertheless, ERPs have continued to be a powerful tool in examining the effects of both intramodal (e.g., Hillyard, Hink, Schwent, & Picton, 1973) and intermodal (e.g., Woods, Alho, & Algazi, 1992) attention on auditory stimulus processing, since ERP measurement does not rely on overt responses by the participant and, consequently, can be recorded for stimuli presented both inside and outside the focus of attention. Therefore, by presenting auditory stimuli outside the focus of attention and comparing the ERPs generated for these stimuli as a function of visual task difficulty, it is possible to assess the impact of visual processing on auditory scene analysis. Specifically, in Experiment 1, we focused on two electrophysiological

This research was funded by grants from the Canadian Institute for Health Research and the Natural Sciences and Engineering Research Council of Canada and by the Premier Research Excellence Award from the Ontario Government. We thank two anonymous reviewers for helpful comments regarding an earlier version of the manuscript. Correspondence should be addressed to B. J. Dyson, Department of Psychology, University of Sussex, Falmer, Brighton BN1 9QH, England (e-mail: b.j.dyson@sussex.ac.uk).

events that have been associated with bottom-up aspects of auditory scene analysis: the mismatch negativity (MMN) wave and the object-related negativity (ORN) wave.

MMN (see Picton, Alain, Otten, Ritter, & Achim, 2000, for a review) is a neural event generated when infrequently occurring *deviant* sounds are presented among frequently occurring *standard* sounds. Derived by subtracting the average ERP waveform of standard sounds from that of deviant sounds, MMN is typically characterized by negativity, which occurs between 100 and 200 msec after deviant onset (Näätänen & Winkler, 1999) and is thought to index the automatic detection of a mismatch between the incoming stimulus and what was expected on the basis of the organization of the previously presented stimuli (Alain, Woods, & Ogawa, 1994; Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995). Although MMN has traditionally been considered to be preattentive in nature and, therefore, independent of attentional effects (Alho, Sams, Paavilainen, Reinikainen, & Näätänen, 1989), experimental situations have been developed in which the magnitude of an MMN response can be modulated by attention (Alain & Izenberg, 2003; Alain & Woods, 1997; Woldorff, Hackley, & Hillyard, 1991). However, changes in MMN as a result of attention have, to date, largely been limited to intramodal, rather than intermodal, conditions. For example, Müller, Achenback, Oades, Bender, & Schall (2002) compared MMN amplitude under conditions of both visual and auditory discrimination. Auditory presentation consisted of a standard stimulus occurring on 82% of the trials and three kinds of deviant stimuli (duration, frequency, and “novelty”) presented for 6% of the trials each. At the same time, participants were visually presented with a series of red and green circles on a computer monitor. In the auditory discrimination case, the participants had to respond to the frequency deviant, whereas for the visual discrimination case, responses were made only when the red circle appeared. Müller et al. found that MMN modulation for the auditory stimuli was contingent upon the completion of the auditory discrimination task, and only with respect to low-contrast (in this case, duration) auditory deviants.

A second ERP component also related to the low-level processing of auditory information is the ORN. This component is characterized by negativity occurring approximately 180 msec after stimulus onset (Alain, Arnott, & Picton, 2001) and is evoked by the presentation of complex sounds comprising a number of tonal components. A critical distinction in ORN studies is made between complex sounds whose partials are harmonically related to one another (i.e., tuned stimuli) and complex sounds that contain partials harmonically unrelated to the rest of the sound (i.e., mistuned stimuli). A common finding has been that although both tuned and mistuned stimuli produce negativity within the defined range, the amplitude of the ORN tends to be greater in mistuned cases. Therefore, when the average waveform generated

by tuned stimuli is subtracted from that produced by mistuned stimuli, ORN is typically revealed. Given that the observation of ORN is related to an increase in the perceptual report of two simultaneous objects within an acoustic scene (Alain, Arnott, & Picton, 2001), it has been argued that ORN may index the number of acoustic objects present during concurrent sound segregation (Alain, Schuler, & McDonald, 2002). In a development in the literature that strongly echoes the study of MMN, sensitivity to attentional modulation is also now being considered for ORN. To date, the effect of attention on ORN has been examined only intramodally. Specifically, ORN generation does not seem to be influenced by contextual factors, such as the probability of the mistuned harmonic within a sequence of stimuli (Alain, Arnott, & Picton, 2001), or, more generally, by listener attention (Alain & Izenberg, 2003). In contrast to MMN, therefore, ORN appears to be unaffected by modulations in intramodal attention. However, it remains an open question as to whether neural indices of concurrent sound segregation such as ORN are sensitive to effects of intermodal attention.

In addition to the ERP components defined in terms of the difference between stimuli (i.e., $\text{MMN} = \text{deviant} - \text{standard}$; $\text{ORN} = \text{mistuned} - \text{tuned}$), a number of other ERPs are generated around the same time as MMN and ORN that reflect general properties of auditory stimulus processing. Indeed, as Woods et al. (1992) have reflected, assessing the extent to which MMN is modulated by attention is often complicated by the presence of additional ERP components that overlap in latency. Specifically, N1 (negative wave occurring approximately 100 msec after stimulus onset) and P2 (positive wave occurring approximately 200 msec after stimulus onset) are obligatory ERPs elicited by auditory stimuli and are thought to be related to the general mechanisms of stimulus detection (Näätänen & Picton, 1987). It is interesting to note that in contrast to MMN and ORN, these components may be susceptible to both intramodal (McArthur, Bishop, & Proudfoot, 2003) and intermodal (Valtonen et al., 2003; Woods et al., 1992) attentional modulation.

In the present study, we assessed the extent to which ERP components related to auditory stimulus processing in general (N1 and P2) and neural markers specific to either sequential (MMN) or concurrent (ORN) sound processing are affected by the modulation of visual task load. Therefore, the contribution of Experiment 1 to the literature is twofold. Specifically, this is the first time that (1) ORN generation has been assessed in an intermodal attentional case and (2) the effect of visual attentional load on both sequential and concurrent sound segregation has been compared under identical conditions. It was hypothesized that if visual and auditory stimulus processing rely on completely independent resources, the modulation of visual task load should fail to have an impact upon the neural metrics of auditory scene analysis described above. Alternatively, if sequential and/or concurrent sound processing is contingent upon the

overall attentional demands of the situation, the MMN and/or the ORN should vary with task instruction in the visual domain.

EXPERIMENT 1

Effects of Visual Task Difficulty

Method

Participants

Fifteen adults provided written informed consent to participate in the study. The data from 1 participant were excluded from further analysis because of extensive ocular contamination in the EEG signal. The final sample of 14 participants consisted of 9 women and 5 men, ranging from 20 to 36 years of age, with a mean age of 23.50 ($SD = 4.36$). All the participants had pure-tone thresholds within normal limits for frequencies ranging from 250 to 8000 Hz (both ears), and all reported normal or corrected-to-normal vision. All received an honorarium for their participation.

Stimuli and Task

Auditory presentation. The participants were presented with two different types of auditory sequences designed to investigate the effects of visual attention on the processing of concurrent and sequential sounds (i.e., spectral and temporal acoustic patterns, respectively). In all cases, the auditory stimuli were calibrated at 85 dB SPL and were 100 msec in duration, including 5-msec rise and fall times. Tones were digitally generated using a 16-bit Tucker Davis System converter at a sampling rate of 50 kHz, were passed through an anti-aliasing filter at 10 kHz, and were played binaurally through Sennheiser HD 265 headphones. Figure 1 provides a schematic representation of the stimuli presented during spectral and temporal pattern blocks.

For each spectral block, 1,320 stimuli were presented in a random order with a variable interstimulus interval (ISI) of 400–600 msec in

10-msec steps (rectangular distribution). Two different types of complex sound were used, with each stimulus having a fundamental frequency of 200 Hz and comprising 12 pure-tone components, all of equal intensity. For the first kind of stimulus (i.e., tuned), all the pure-tone components were harmonically related to the fundamental frequency (yielding 200-, 400-, 600-, 800-, 1000-, 1200-, 1400-, 1600-, 1800-, 2000-, 2200-, and 2400-Hz partials). For the second kind of stimulus (i.e., mistuned), the third harmonic was shifted either upward or downward by 16% of its original value (696 or 504 Hz, instead of 600 Hz). Tuned stimuli were presented over 660 (50%) trials, whereas upward- and downward-mistuned stimuli were presented over 330 (25%) trials each.

For each temporal block, 240 groups of tones were played at a fixed ISI of 400 msec. The first group of tones (i.e., the standard group) was presented over 192 trials (80%) and consisted of a 1189-Hz tone followed by an 841-Hz tone. The second and third groups of tones (i.e., deviant groups) were presented over 24 trials (10%) each and consisted of the groupings 1189–1189–841 Hz and 841–841–1189 Hz, respectively. Therefore, both deviant groups broke the alternating pitch pattern created by the standard group (cf. Alain et al., 1994). The use of repetition as a source of deviance is ideal, since the subsequently generated MMN is less contaminated by the contribution of nonrefractory auditory neurons. The sequencing of presentation was also designed with the proviso that deviant groups were always separated by at least two but no more than seven standard groups.

Visual presentation. In addition to the presentation of auditory stimuli, the participants were required to respond to continuously presented digits displayed on a computer monitor at a distance of 60 cm. The participants completed both an easy and a hard version of the task, although stimulus presentation was identical in both cases. Blue and red digits (1, 2, 3, 4, 6, 7, 8, and 9), approximately $2^\circ \times 2^\circ$ in visual angle, were presented at fixation for 100 msec, with stimulus onset asynchrony (SOA) varying between 1,500 and 2,000 msec in 100-msec steps. In the easy version of the task, the

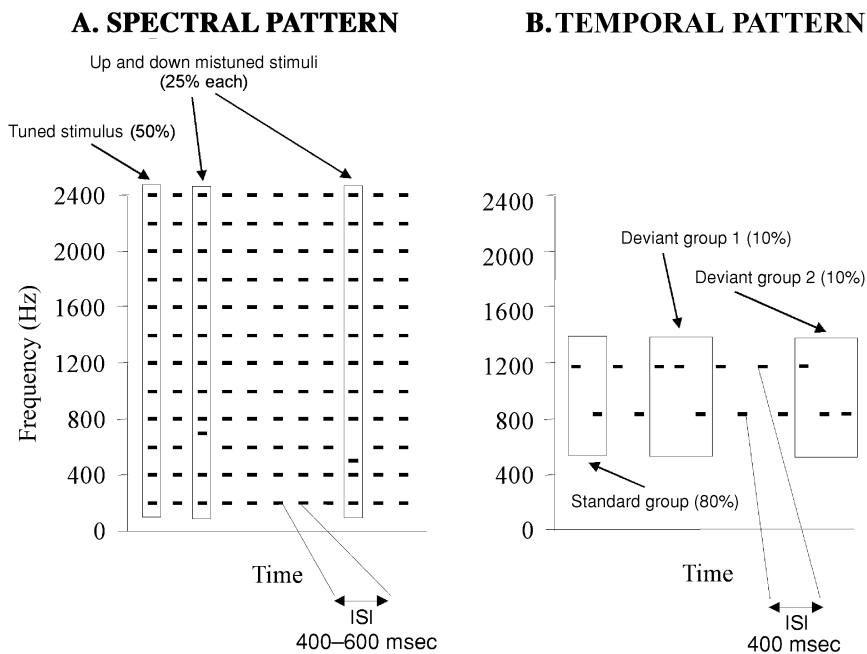


Figure 1. Schematic representation of the auditory stimulus sets used in Experiment 1. In both cases, frequency is represented on the y-axis, and time is represented on the x-axis, with all the stimuli 100 msec in length: (A) Tuned and mistuned stimuli used in the spectral pattern condition. (B) Standard and deviant groups used in the temporal pattern condition.

color of the digit was irrelevant, and the participants simply indicated whether the current digit was lower or higher than five by pressing one of two keys. In the hard version of the task, digit color was relevant. If the digit was blue, the participants were required to categorize the current digit as either above or below five. If the digit was red, the participants were required to categorize the *previous* digit as either above or below five. No feedback was given during performance, and the participants responded to an average of 224 digits per block.

The participants completed a total of four conditions, derived by the orthogonal combination of auditory stimuli (spectral or temporal) and visual task (easy or hard). Each condition consisted of 3 identical blocks of trials, with each block lasting approximately 4.5 min. All the conditions took place in a sound-attenuating room. The participants were encouraged to take a short break between conditions. After the completion of all 12 blocks, the participants were thanked for their time and were debriefed as to the nature of the investigation.

Recording of Electrophysiological Data

Electrical brain activity was digitized continuously (band-pass 0.05–50 Hz; 250-Hz sampling rate) from an array of 64 electrodes, using NeuroScan SynAmps, and were stored for offline analysis. Eye movements were recorded with electrodes placed at the outer canthi and at the superior and inferior orbits. During data collection, all electrodes were referenced to the midline central electrode (i.e., Cz), whereas for data analysis, they were rereferenced to an average reference.

Only ERPs to the auditory stimuli were recorded, with the analysis epoch representing 100 msec of prestimulus activity and 400 msec of poststimulus activity. Trials contaminated by excessive peak-to-peak deflection ($\pm 100 \mu\text{V}$) at the channels not adjacent to the eyes were automatically rejected before averaging. ERPs were then averaged separately for each site and stimulus type (tuned, mistuned, standard, and deviant) in both the easy and the hard visual conditions. In the case of mistuned and deviant stimuli, ERPs were collapsed across the two different examples of each stimulus type. ERPs were digitally low-pass filtered to attenuate frequencies above 30 Hz. For each individual, ocular artifacts (e.g., blinks, saccades, and lateral movements) were corrected by means of ocular source components, using the Brain Electrical Source Analysis software (Picton, van Roon, et al., 2000).

Analysis of Electrophysiological Data

Task and stimulus modulation of N1 and P2 components for temporal and spectral blocks were investigated with separate three-way repeated measures analyses of variance (ANOVAs). For both types of acoustic presentation, the factor of visual task (easy or hard) and the factor of electrode (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, or C2)¹ were used, whereas the remaining two-level factor of auditory stimulus differed according to the nature of the auditory stimuli. In the case of spectral patterning, tuned and mistuned stimuli were compared. In the case of temporal patterning, the comparison was between standard and deviant stimuli. Peak latency and amplitude for N1 were defined as the maximum negativity occurring between 100 and 150 msec after stimulus onset. Similarly, peak latency and amplitude for P2 were defined as the maximum positivity between 150 and 200 msec after stimulus onset. Mean amplitude for both N1 and P2 was defined according to a 20-msec time window comprising 10 msec either side of the peak latency value. This value for mean amplitude was based on the group mean of peak latency over all levels of visual task, auditory presentation, and electrode, unless analyses revealed significant differences in peak latency across these factors. In such cases, variation in peak latency for individual levels of visual task, auditory presentation, and/or electrode were taken into account, and separate mean amplitude time windows were calculated.

The effect of visual attentional load on ORN and MMN was examined in separate two-way repeated measures ANOVAs, with the factor of visual task having two levels (easy or hard) and the factor of electrode having nine levels (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, or C2). ORN was calculated as the difference between tuned and mistuned stimuli and was defined as the maximum negativity between 100 and 200 msec after stimulus onset. MMN was calculated as the difference between deviant and standard stimuli and was defined as the maximum negativity between 125 and 225 msec after stimulus onset. In a procedure similar to the examination of N1 and P2, if there were no significant differences in peak latency between visual task and electrode, mean amplitude was defined over a 40-msec time window, 20 msec either side of the group mean peak latency. If significant latency differences across factor levels were obtained, individual 40-msec time windows were calculated accordingly. In all cases, the original degrees of freedom were reported, although significance levels were based on reduced degrees of freedom, as determined by the Greenhouse–Geisser epsilon.

Results

Behavioral Data

Table 1 provides a summary of the behavioral responses made to the visual stimuli in Experiment 1. Both reaction time (RT) and error rate data were analyzed according to separate three-way repeated measures ANOVAs, consisting of the factors of auditory presentation (spectral or temporal), visual task (easy or hard), and digit color (blue or red).² The RT data revealed significant main effects for visual task [$F(1,13) = 148.56, p < .001$] and digit color [$F(1,13) = 44.44, p < .001$], in addition to a significant two-way interaction between visual task and digit color [$F(1,13) = 58.28, p < .001$]. This interaction was further explored using Tukey's HSD test, revealing all pairwise comparisons to be significant ($p < .05$), apart from the comparison between blue and red digits in the easy visual task ($p > .05$). The arcsin-transformed error rate³ revealed an identical pattern of results, with significant main effects for visual task [$F(1,13) = 22.90, p < .001$] and digit color [$F(1,13) = 7.72, p = .031$] and a significant two-way visual task \times digit color interaction [$F(1,13) = 11.64, p = .007$]. A Tukey's HSD test again showed all pairwise comparisons to be significant

Table 1
Summary Statistics for Behavioral Responses in Experiment 1

Condition	RT	SE	%Error
Spectral pattern			
Easy visual task			
Blue digits	557	16	1.65
Red digits	556	17	1.72
Hard visual task			
Blue digits	725	26	5.82
Red digits	862	34	9.65
Temporal pattern			
Easy visual task			
Blue digits	557	18	1.82
Red digits	558	19	1.57
Hard visual task			
Blue digits	720	19	5.81
Red digits	854	31	8.14

Note—RT, reaction time (in milliseconds); SE, standard error.

($p < .05$), apart from the comparison between blue and red digits in the easy visual task ($p > .05$).

The data are clear in showing that the participants were both slower and more error prone (1) in the hard visual task, relative to the easy visual task, and (2) in responding to red digits, relative to blue digits, in the hard visual task only. The former observation reveals a successful manipulation of visual task, whereas the latter observation confirms the locus of this manipulation, given that red digits in the hard visual task were the only cases in which the participants were required to respond to the previous (as opposed to the current) stimulus. Moreover, the null effect of auditory presentation in the behavioral data supports the idea that any intermodal influence observed in Experiment 1 would be essentially asymmetric (i.e., vision influencing audition), since the type of auditory stimulus sequence had no systematic impact upon performance during the visual tasks.

ERP Data

Spectral pattern condition. Figure 2 shows the group mean ERPs for tuned and mistuned stimuli at frontocentral sites, with both tuned and mistuned stimuli generating N1 and P2 deflections. Critical for the experimental hypothesis, the N1 wave peak and mean amplitude decreased with increasing visual task load [$F(1,13) = 9.04$ and 6.04 , $p < .05$ in both cases]. The N1 for the mistuned stimuli was also larger ($-0.20 \mu\text{V}$) and was delayed (134 msec) relative to the N1 elicited by the tuned stimuli ($0.08 \mu\text{V}$ and 130 msec, respectively), as revealed by

significant main effects of auditory stimulus for mean amplitude and peak latency [$F(1,13) = 12.34$ and 9.76 , $p < .01$ in both cases]. These effects were partly due to the superimposition of the ORN on the N1 (see below).

With respect to P2, visual task load showed only an effect on peak latency, with P2 occurring earlier as task difficulty increased [173 msec in the hard case, and 178 msec in the easy case; $F(1,13) = 9.40$, $p = .009$]. P2 peak and mean amplitudes were affected by the auditory stimulus, however, being smaller for the mistuned (mean, $1.10 \mu\text{V}$) than for the tuned (mean, $1.34 \mu\text{V}$) stimuli [$F(1,13) = 11.61$ and 17.60 , $p < .01$ in both cases].⁴ As for the N1, this effect can be attributed to the ORN being superimposed on the P2 wave. In addition, Figure 2 indicates task differences in a sustained negativity for both the tuned and the mistuned stimuli. Mean amplitude calculated between 300 and 400 msec after stimulus onset revealed significantly greater negativity in the easy task ($-0.77 \mu\text{V}$), relative to the hard task ($-0.52 \mu\text{V}$) [$F(1,13) = 9.21$, $p = .010$].

Figure 3 shows the ORN wave for both the easy and the hard visual tasks across the nine frontocentral electrode sites. Although ORN was produced during both visual conditions, the ANOVA revealed no significant main effect of visual task for peak latency [$F(1,13) = 2.39$, $p = .146$] or peak and mean amplitude [$F(1,13) = 3.38$, $p = .089$, and $F(1,13) = 2.18$, $p = .164$, respectively].

Temporal pattern condition. Figure 4 shows the group mean ERPs for standard and deviant stimuli at nine frontocentral sites. Due to the superimposition of

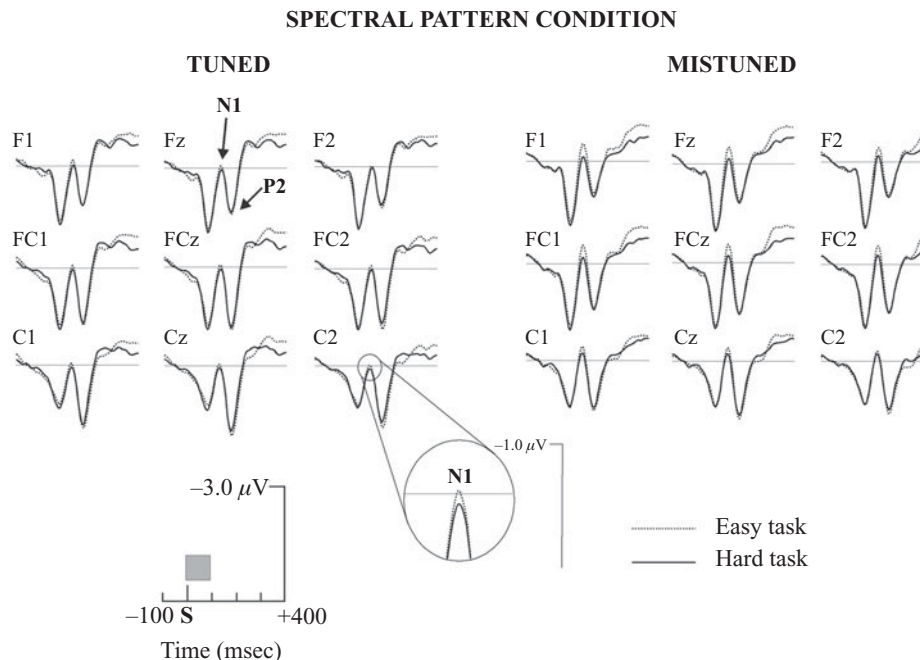


Figure 2. Group mean ($n = 14$) event-related brain potentials from nine frontocentral electrode sites (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2) during spectral pattern presentation in Experiment 1. The waveforms vary as a function of stimulus type (tuned and mistuned) and visual task (easy and hard). The gray rectangle indicates the duration of the stimulus.

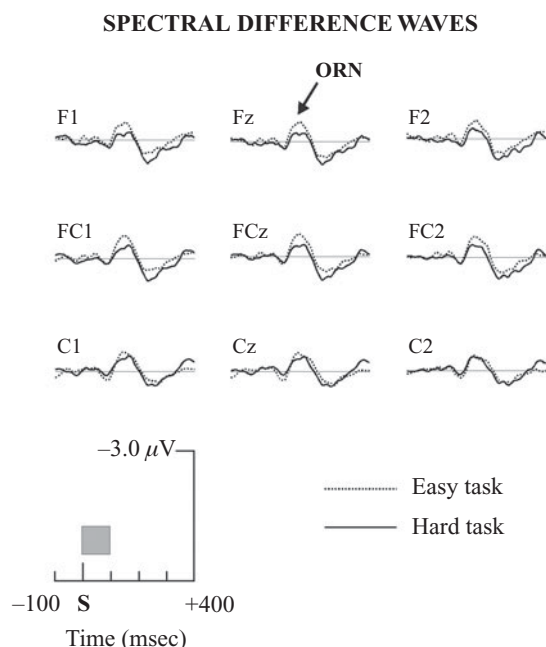


Figure 3. Group mean ($n = 14$) difference waves between the tuned and the mistuned stimuli during spectral pattern presentation over nine frontocentral electrode sites in Experiment 1. The gray rectangle indicates the duration of the stimulus. ORN, object-related negativity.

MMN upon the waveform for the deviant stimuli, N1 and P2 deflections were quantified for the standard stimuli only. Peak latency failed to reveal a significant main effect of task in either case, positioning the average peak latency of the N1 at 127 msec and the P2 at 173 msec. N1 peak amplitude was significantly larger in the easy task ($-0.36 \mu\text{V}$), relative to the hard task ($-0.13 \mu\text{V}$) [$F(1,13) = 6.37, p = .025$], whereas P2 peak amplitude showed the reverse effect of larger P2 amplitude during the hard task ($0.83 \mu\text{V}$), relative to the easy task [$0.71 \mu\text{V}$; $F(1,13) = 5.13, p = .041$]. Once again, an examination of sustained negativity between the time period 300–400 msec after stimulus onset revealed a statistically greater mean amplitude in the easy visual task ($-0.38 \mu\text{V}$), relative to the hard visual task ($-0.18 \mu\text{V}$) [$F(1,13) = 6.70, p = .022$].

A closer examination of the ERP waveform generated for the deviant stimuli in Figure 4 revealed a negative–positive deflection, later than that typically observed for N1 and P2. Figure 4 helped to quantify the time range of these effects, with the negative component defined as the maximum negativity occurring between 175 and 225 msec after stimulus onset and the positive component defined as the maximum positivity observed between 225 and 275 msec after stimulus onset. Both were submitted to a two-way repeated measures ANOVA with the factors of visual task (two levels) and electrode site (nine levels), but a statistically significant main effect of visual task

failed to be revealed using either peak or mean amplitude for both the negative [$F(1,13) = 2.74$ and $2.21, p > .05$, respectively] and the positive [$F(1,13) = 1.57$ and $2.45, p > .05$, respectively] components.

Figure 5 shows the production of MMN for both the easy and the hard visual tasks across nine frontocentral sites (the figure also shows the nonsignificant positivity following MMN, as described and analyzed above). However, the ANOVA revealed no significant main effect of visual task for MMN peak latency or peak and mean amplitude ($F < 1$) in all cases.

Further Analyses

Although the data from Experiment 1 point to a number of conclusions regarding the interaction between auditory and visual attentional systems, two issues were left outstanding.⁵ First, in contrast to the lack of modulation of MMN by visual task difficulty, the main effect of ORN amplitude was close to significance ($p = .089$ for peak amplitude). Although Experiment 1 showed statistically significant modulation of other ERP components as a consequence of visual attentional load, the ORN result may have failed to reach significance as a result of our sample size not being large enough. Second, the ERP data recorded in Experiment 1 undoubtedly contained responses to both visual and auditory stimuli, although waveforms were averaged according to auditory presentation only. However, differences in slow-wave visual ERPs may have been generated as a result of the varying task loads, which may have led to concomitant differences in the averaged auditory ERP waveforms. Although we consider this explanation unlikely, given that (1) visual stimuli were identical in both easy and hard cases and (2) auditory and visual presentation were asynchronous, we thought it was interesting to empirically investigate this issue nonetheless.

In order to address these two issues, 4 additional participants were run, consisting of 2 males and 2 females, with ages from 21 to 35 years. All the individuals had normal auditory thresholds (250 to 8000 Hz), and all were paid for their contribution. The participants completed a total of four conditions, each consisting of three blocks of trials, as before. The first two conditions consisted of counterbalanced easy and hard visual tasks, respectively, with spectral pattern auditory presentation as described in Experiment 1. These data could then be added to the previous sample, thereby allowing us to assess the effect of visual attentional load on ORN with a sample of 18. The other two conditions consisted of counterbalanced easy and hard visual presentations, with no auditory presentation. Despite the absence of acoustic stimulation, both behavioral and ERP measurements were recorded during these conditions, with ERP recording time locked to the onset of these “phantom” auditory stimuli, which the participants could not hear. By examining this baseline condition, it would be possible to assess the second issue and whether easy and hard visual

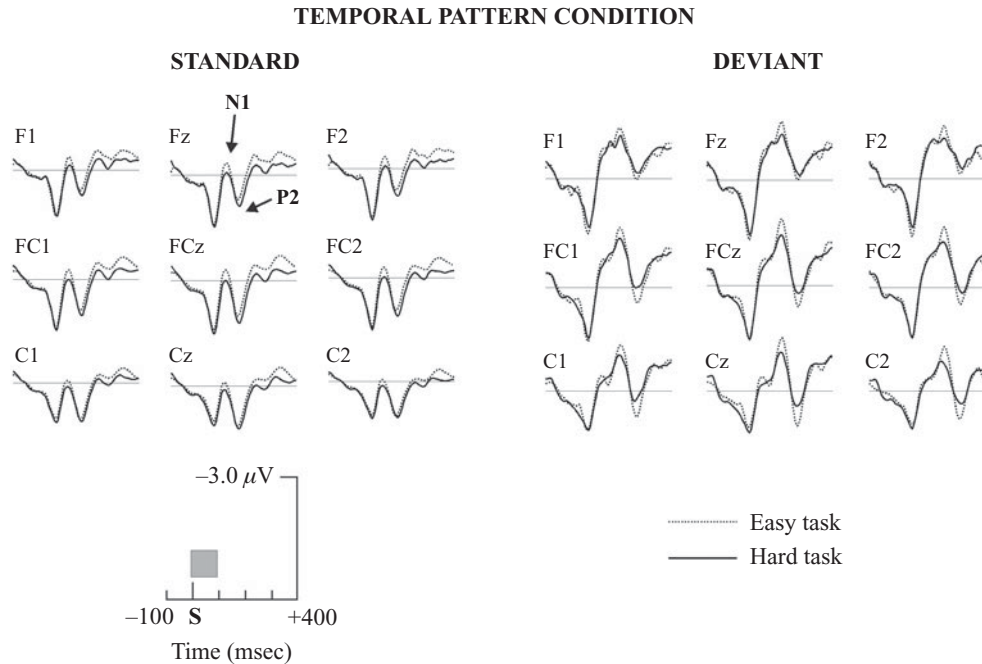


Figure 4. Group mean ($n = 14$) event-related brain potentials from nine frontocentral electrode sites during temporal pattern presentation in Experiment 1. The waveforms vary as a function of stimulus type (standard and deviant) and visual task (easy and hard). The gray rectangle indicates the duration of the stimulus.

tasks created differential slow-wave visual ERP responses, which may have systematically contaminated the recording of the auditory ERP.

Behavioral data. In order to verify that similar levels of visual task difficulty were created for the two conditions, relative to Experiment 1, the behavioral data from the additional 4 participants were visually inspected. RT means were collapsed over auditory stimulus presentation (present or absent), yielding 539 and 551 msec for the blue and the red digits, respectively, in the easy visual task and 678 and 809 msec for the blue and the red digits, respectively, in the hard visual task. With reference to the spectral pattern condition data presented in Table 1, the RT data from these 4 new participants echo those from the previous sample of 14. Specifically, responses to blue and red digits were equivalent during the easy task. Both blue and red digit RTs were inflated during the harder task, while red digit responding was slower than blue digit responding, given the additional processing indexed by the presentation of a red digit (i.e., respond to the previous stimulus). Therefore, we are satisfied as to the representative nature of this smaller subsample with respect to the attentional demands created in the two versions of the visual task.

Visual ERP contamination? The left-hand panel in Figure 6 shows the group average auditory ERP generated during no auditory stimulus presentation over both the easy and the hard visual tasks. Save for noise, the re-

sultant waveforms were essentially flat, indicating that there was no systematic contribution of visual ERP to the recording of auditory ERP. The hint of increased positivity for the hard visual condition around 225 msec after stimulus onset was not considered to raise problems, since this was both outside the temporal region of interest from ORN⁶ and in the opposite direction to what we were observing in the group average ORN data.

ORN modulation ($n = 18$). The right-hand panel in Figure 6 shows the group average ORN generated during the easy and the hard visual conditions, combining the 14 participants taken from Experiment 1 with the additional 4 individuals. Peak latency together with peak and mean amplitude of ORN as a function of visual task load were analyzed in a manner identical to that in Experiment 1. With this larger sample size, however, no statistically significant main effect of visual task load was revealed for peak latency [$F(1,17) = 2.17, p = .159$], peak amplitude [$F(1,17) = 2.55, p = .128$], or mean amplitude [$F(1,17) = 1.68, p = .212$]. In order to assess whether the absence of ORN modulation was due to the selection of an unrepresentative set of electrodes, peak global field power across all 65 scalp recordings was compared. However, this measurement also failed to capture a significant difference between the easy and the hard visual tasks [$t(17) = 0.16, p = .875$].

ORN failed to be modulated by visual task load in this larger sample, supporting the notion that interactions

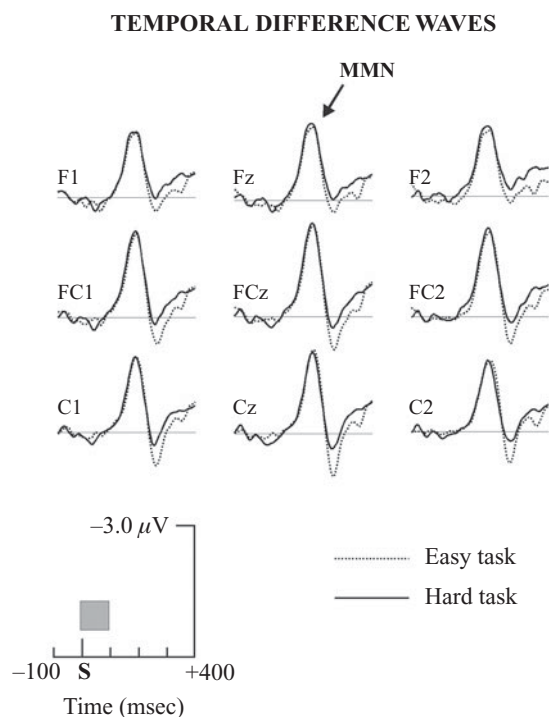


Figure 5. Group mean ($n = 14$) difference waves between the standard and the deviant stimuli during temporal pattern presentation over nine frontocentral electrode sites in Experiment 1. The gray rectangle indicates the duration of the stimulus. MMN, mismatch negativity.

between the visual and the auditory senses are more reliably revealed by auditory process-general neural components (i.e., N1 and P2) than by auditory process-specific neural components (i.e., ORN and MMN). However, it is clear from these additional analyses that ORN is a variable neural component across participants. As a final test as to whether ORN could be modulated by visual attention in the present sample, correlations between the magnitude of the RT cost and the magnitude of the difference in ORN peak amplitude between the easy and the hard visual conditions, averaged across the nine frontocentral electrodes, were calculated for each of the 18 participants who completed the ORN conditions. The correlation for ORN also failed to reach statistical significance ($r = -.255, p = .306$), further supporting the claim that auditory process-specific components appear to be resilient to attentional demands in a second modality.

Discussion

The behavioral data in Experiment 1 revealed a successful manipulation of visual task, in that the hard version yielded both slower and more error prone performance than did the easy version, suggesting the use of greater attentional resources in the former case. The increase in visual task load was paralleled by a reduction in N1, an increase in P2, and a decrease in sustained neg-

ativity occurring between 300 and 400 msec after stimulus onset. Specific neural indices of sequential (MMN) and concurrent (ORN) auditory processing, however, failed to be significantly modulated by differential attentional demands in the visual domain.

In terms of addressing the issue of independent or shared resources between the senses, a central result was the consistent observation of N1 reduction during both concurrent and sequential sound processing as visual task difficulty increased. In contrast to intramodal studies, which have reported the robustness of the N1 response in the presence of irrelevant acoustic stimulation (McArthur et al., 2003), the observation of decreased N1 amplitude for auditory stimuli during a more attentionally demanding visual task is consistent with the idea of shared processing resources between vision and audition (i.e., Kahneman, 1973). These data are consistent not only with previous research that shows N1 amplitude modulation for auditory stimuli when attention is directed toward vision has been reported (Valtonen et al., 2003; Woods et al., 1992), but also with studies that have demonstrated the early modulation of visual ERPs when auditory stimuli cue the spatial location of a visual target (McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003; McDonald & Ward, 2000).

However, the direction of the N1 effect is critical for the interpretation of the data. Specifically, although the present data are consistent with the *decrease* in N1 observed by Woods et al. (1992), Valtonen et al. (2003) reported an *increase* in the magnetic counterpart of N1 (N1m) under conditions of high visual attentional load. Whereas Valtonen and colleagues appealed to increased arousal in the hard visual case as an explanation of increased N1m, a similar case cannot be made for Experiment 1, given that N1 modulation was in the reverse direction. However, differences in arousal may not be the only contributing factor. Specifically, accentuated N1m in Valtonen et al.'s study appeared to be limited to a rehearsal, as opposed to an encoding, phase of the visual task, in which participants were required to repeat a string of digits in sequential order. Therefore, differences in the use of certain aspects of working memory (e.g., the articulatory loop; Baddeley & Hitch, 1974) may also determine the direction of N1 modulation.

In contrast to the effect of visual task load on auditory N1 and P2 response, additional ERP components occurring around the same time (i.e., MMN and ORN) failed to demonstrate statistically significant effects for visual task. Although neither the differences in MMN nor those in ORN reached statistical significance as a result of visual task load, it is important to consider how the configuration of individual ERP components, such as N1 and P2, have an impact upon the nature of overlapping components, such as MMN and ORN (cf. Woods et al., 1992). Specifically, it will be noted that only a significant modulation of N1 was observed for both tuned and mistuned stimuli in the case of ORN. This is in contrast to MMN, in which both N1 and P2 modulation for stan-

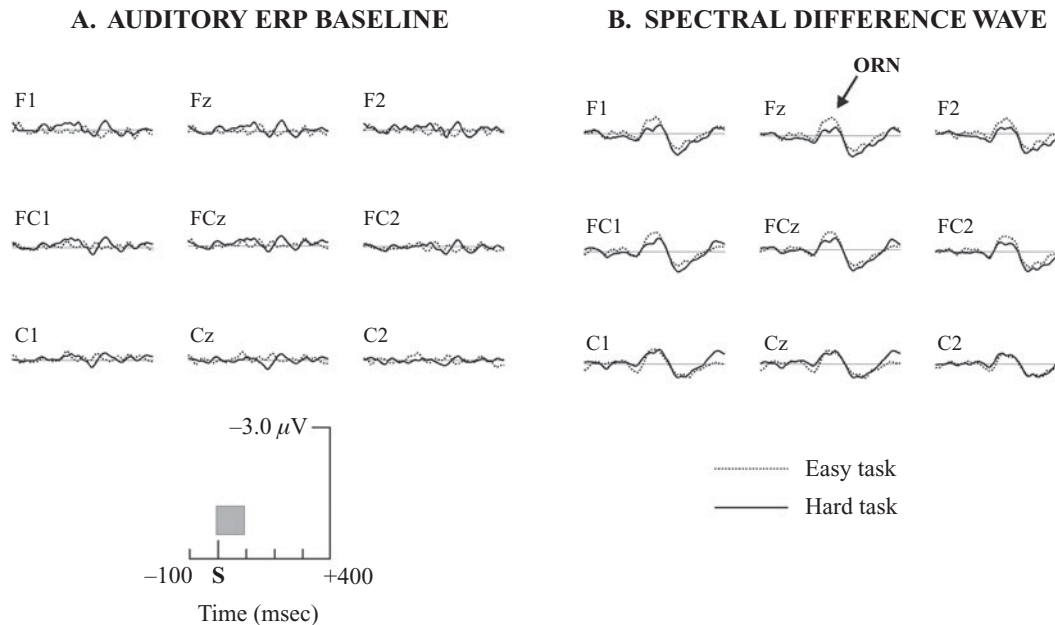


Figure 6. (A) Group mean ($n = 4$) event-related brain potentials (ERPs) from nine frontocentral electrode sites recorded from “phantom” auditory stimuli in the spectral pattern condition unheard by the participants during both easy and hard visual tasks. (B) Group mean ($n = 18$) difference waves between the tuned and the mistuned stimuli during spectral pattern presentation over nine frontocentral electrode sites. The gray rectangle indicates the duration of the stimulus. ORN, object-related negativity.

dard stimuli was noted. These underlying differences between N1 and P2 modulation for ORN and MMN are compounded by the fact that, in the latter example, these modulations were expressed in opposite directions. With reference to Figure 4, N1 amplitude was attenuated during the hard task, whereas P2 amplitude was enhanced. Therefore, when it came to calculating MMN difference, it is possible that these two effects of visual task load essentially canceled one another out.⁷ Since there was no concomitant P2 modulation during spectral sound processing, this is perhaps one reason why ORN modulation appeared to be different from MMN at first sight.

The failure to observe MMN change as a result of visual task manipulation is consistent with previous research in which it has been claimed that MMN is sensitive only to intramodal attentional constraints (Alho et al., 1994; Alho et al., 1992; Müller et al., 2002). Moreover, it extends previous work in demonstrating MMN resistance to visual attentional load, using an *x*-back task (Winkler et al., 2003), with MMN in the present experiment operationalized as a change in an abstract rule (i.e., the sequential order of stimuli), as opposed to simple alterations in basic acoustic features, such as frequency (Otten et al., 2000). Although the resilience of ORN amplitude to visual attentional load is a novel finding, it is consistent with other work that has shown ORN insensitivity under conditions of intramodal attentional manipulation (Alain & Izenberg, 2003). In this respect, Experiment 1 shows that the influence of visual load on audition

appears to be limited to general characteristics of auditory stimulus processing (as indexed by N1; Näätänen & Picton, 1987), which operate independently from subsequent (and assumedly, more specific) processes associated with the presentation of sequential (i.e., MMN) or concurrent (i.e., ORN) sounds.

EXPERIMENT 2

Effects of Visual Task Content

By designing Experiment 1 in such a way that the visual and the auditory stimuli were not related in any meaningful way to one another, it was possible to assess the extent to which fundamental processing constraints existed when simultaneously presented auditory and visual stimuli were dealt with. However, additional behavioral work in which the processing of auditory information has been examined as a function of visual load (e.g., Tellinghuisen & Nowak, 2003) has employed experimental paradigms in which, although the auditory stimuli were task irrelevant, certain aspects of the acoustic signal were either compatible or incompatible with task-relevant information in the visual domain, thereby increasing the likelihood of similar forms of coding for the visual and the auditory stimuli.

Therefore, an alternative explanation for the absence of auditory MMN modulation as a result of visual task load differences in Experiment 1 may be the absence of a shared level of description between the two senses.

Consequently, Experiment 2 was designed to test the hypothesis that auditory ERP modulation as a function of visual task load may also depend on the extent to which similar types of coding are shared between the senses—the assumption being that shared coding would increase the use of similar functional operations between vision and audition, thereby increasing the likelihood of observing constraints among putative intersensory processing resources. One framework for thinking about common modes of processing for auditory and visual stimuli is the distinction between identity (i.e., *what*) and localization (i.e., *where*) analyses. Both the visual (Ungerleider & Mishkin, 1982) and the auditory (Anourova et al., 2001; Clarke, Adriani, & Bellmann, 1998) systems have previously been described with respect to these two separate functional streams of processing, with *what* information being processed along a ventral pathway and *where* information running to completion relative to a dorsal pathway (Alain, Arnott, Hevenor, Graham, & Grady, 2001).⁸

In Experiment 2, the participants were required to make judgments regarding either the direction (i.e., *what*) or the position (i.e., *where*) of visually presented arrows (after Farrell, 1979). The visual experimental design was based on a standard Garner interference paradigm (Pomerantz, 1983), in which the participants processed single stimulus attributes in both the presence (i.e., filtering conditions) and the absence (i.e., baseline conditions) of variation in the second attribute across separate blocks. If similar conceptualizations of visual and auditory stimuli led to the use of sensory-independent but task-specific operations, then processing one kind of information in vision (i.e., visual *what*) should limit the extent to which similar information could be detected in audition (i.e., auditory *what*) but should not affect the registration of information from a different attribute (i.e., auditory *where*). Therefore, during the completion of the various visual blocks, the participants were exposed to two acoustic patterns similar in nature to the temporal pattern condition employed in Experiment 1. Again, these two patterns established a standard pattern of alternation and introduced deviance by repetition. Critically, auditory stimuli were defined relative to frequency (i.e., *what*) alternation in one instance and location (i.e., *where*) alternation in the second, so that sensitivity to *what* and *where* processing in audition, as indexed by ERPs, could be examined as a function of *what* and *where* processing in vision. Therefore, the primary goal of Experiment 2 was not to modulate auditory ERP as a function of visual task difficulty but, rather, to modulate auditory ERP as a function of visual task specificity by allowing visual and auditory stimuli to be processed using the same or different putative pathways. Specifically, if auditory ERPs were attenuated as a result of common modes of processing across the senses, these effects should be clearer in the visual baseline conditions as a result of fixing the task-irrelevant attribute, thereby allowing a single putative pathway (i.e., *what* or *where*)

to be activated during each condition. In contrast, the presence of irrelevant variation in the visual filtering conditions might lead to less clear modulation of auditory ERPs, given the activation of both putative pathways (i.e., *what* and *where*) in each condition.

Method

Participants

Informed consent for 12 participants was obtained before the experiment. Ages ranged from 20 to 38 years, with a mean age of 27.83 years ($SD = 5.06$), and the sample consisted of 8 females and 11 right-handers. All the participants had hearing thresholds below 35 dB SPL between 250 and 8000 Hz (both ears) and reported normal or corrected-to-normal vision. All volunteers received an honorarium for their participation.

Stimuli and Apparatus

Visual stimuli. A gray fixation cross ($1^\circ \times 1^\circ$) was used to orient the participants to center screen. The cross was also used to provide feedback, changing from gray to green or red, depending on whether the participant made a correct or an incorrect response, respectively. Four white arrows ($2^\circ \times 2^\circ$) were presented in Experiment 2, each having a different orientation (pointing up, down, left, or right), which could be presented 2° away from fixation in up, down, left, or right positions. All the visual stimuli were presented over a computer monitor positioned approximately 57 cm from the participant.

Auditory stimuli. As in Experiment 1, two different types of auditory pattern were played concurrently with the completion of the visual tasks. A modified version of the temporal pattern used in Experiment 1 was employed as the frequency (i.e., *what*) condition in Experiment 2. Frequency separation between the two stimuli was now 118 Hz (1059 and 941 Hz for the high and the low tones, respectively), and the ISI was set at a constant 150 msec, allowing for the presentation of 340 groups of stimuli per block. Stimulus rate was identical for the location (i.e., *where*) condition and involved the presentation of alternating leftward (-30°) and rightward ($+30^\circ$) locale 1000-Hz tones. Location differences were created via the use of head-related transfer functions. Both *what* and *where* auditory patterns were presented binaurally, and the location condition was presented 6 dB louder than the frequency condition, in order to equate for subjective loudness. In both cases and as in Experiment 1, pattern deviance represented 20% of group presentation, with the *what* condition breaking the alternating frequency pattern by introducing pitch repetition and the *where* condition breaking the alternating location pattern by introducing locale repetition. All other aspects of auditory stimulus presentation were identical to those in Experiment 1.

Design

Visual tasks. The participants were required to complete four visual conditions (two baseline and two filtering conditions). Each condition involved four subtasks, with each subtask consisting of a practice block of 16 trials (4 presentations of the 4 stimuli in the baseline conditions, 1 presentation of the 16 stimuli in the filtering conditions), followed by an experimental block of 128 trials (32 presentations of the 4 stimuli in the baseline conditions, 8 presentations of the 16 stimuli in the filtering conditions). Condition order was counterbalanced between individuals, and subtask order and trial order within each block were randomized.

In each of the four conditions, the participants were required to respond to the specific value of a single stimulus dimension. The four conditions were defined according to the dimension being judged (position or direction) and the nature of the task (baseline or filtering). In the baseline conditions, the stimulus dimension critical to responding was allowed to vary, whereas the irrelevant di-

mension was held at a different constant value for each of the four subtasks. In the filtering conditions, both the response-relevant and the response-irrelevant dimensions varied orthogonally during each subtask. Therefore, in the position baseline condition, the participants made judgments regarding the location of an arrow (up, down, left, or right of fixation), and at each subtask, the identity of the arrow was fixed (pointing up, down, left, or right). Similarly, in the direction baseline condition, the participants made judgments regarding the orientation of an arrow (pointing up, down, left, or right), and at each subtask, the location of the arrow was fixed (up, down, left, or right of fixation). In contrast, for both the location-filtering and the identity-filtering conditions, both stimulus dimensions varied over each subtask. That is, the participants responded to a variation in one stimulus dimension (e.g., position) while attempting to ignore irrelevant variation in the second stimulus dimension (e.g., direction).

Auditory presentation. Concurrently with the completion of the visual conditions, the participants were also presented with pure tones through headphones during experimental blocks only. Given that each of the four main conditions involved four subtasks, both the frequency (i.e., *what*) and the location (i.e., *where*) conditions were presented during two subtasks each, with the allocation of auditory stimuli presentation to subtask randomized between individuals. Recording and analysis of the electrophysiological response generated by the auditory stimuli were identical to those in Experiment 1.

Procedure

Before the start of the experiment, the participants were informed that they were required to respond to various singular properties of visual stimuli over a number of trials. They were also told that auditory stimuli would be presented at the same time that they were completing these conditions, but since no overt response would ever be required to these stimuli, they could essentially ignore any sounds presented during the experiment. At each trial, a central gray fixation cross was presented on a computer monitor for 250 msec, after which a single arrow was presented. The arrow was defined relative to position and direction (in both cases, up, down, left, or right). With 5 as an orienting position on a numeric keypad, the participants were instructed to make a fast and accurate ballistic movement, using the index finger of their preferred hand, to the key that was compatible with the task-relevant dimension. For example, if the participants were completing a position condition and the trial contained an arrow presented left of fixation, the correct response would be 4 on the keypad (regardless of the identity of the arrow). Similarly, if the participants were completing a direction condition and the trial contained a downward-pointing arrow, the correct response would be 2 on the keypad (regardless of the location of the arrow). After a response had been made, the arrow disappeared, and feedback was provided center screen in the form of an appropriately colored fixation cross for 500 msec, after which the cross returned to gray for a further 300 msec. Prior to the initiation of each condition, the participants were informed as to which visual stimulus dimension (position or direction) was critical to current task performance.

Results

Behavioral Data

Table 2 provides a summary of the behavioral responses made to the visual stimuli in Experiment 2. Both the RT and the arcsin-transformed error rate data were analyzed according to separate three-way repeated measures ANOVAs, consisting of the factors of auditory pattern (frequency or location), condition (baseline or filtering), and visual dimension (position or direction).

The RT data revealed significant main effects for both condition [$F(1,11) = 47.13, p < .001$] and visual dimension [$F(1,11) = 90.27, p < .001$], in addition to a statistically significant two-way interaction between condition and visual dimension [$F(1,11) = 5.21, p = .043$]. A Tukey's HSD test revealed all pairwise comparisons to be significant from one another ($p < .05$), apart from the difference between the position baseline and the position filter conditions. Arcsin-transformed error rates failed to show any statistically significant main effects or interactions. Therefore, the RT data show that the participants were (1) significantly faster in responding to visual position (*where*) information, relative to visual direction (*what*) information, and (2) slowed in the filtering condition, relative to the baseline conditions, but only when responding to *what* information while attempting to suppress *where* information. Once again, the fact that auditory pattern failed to produce a significant main effect or interactions with the other factors suggests that the type of auditory signal delivered had no systematic influence on visual stimulus processing.

Electrophysiological Data

Frequency (*what*) pattern condition. N1 and P2 components of auditory ERP in the frequency condition were quantified for standard stimuli in a way identical to that in Experiment 1 (see Figure 7). Data related to peak latency, peak amplitude, and mean amplitude were entered into three-way repeated measures ANOVAs with the factors of condition (baseline or filtering), visual dimension (position or direction), and electrode (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, or C2). N1 peaked around 136 msec after stimulus onset in all the conditions and revealed no statistically significant main effects or interactions for any of three measures. In contrast, maximal P2 amplitude was observed around 179 msec following stimulus presentation for all the conditions and revealed a statistically significant two-way interaction between condition and visual dimension with respect to peak amplitude [$F(1,11) = 7.25, p = .021$]. A Tukey's

Table 2
Summary Statistics for Behavioral Responses in Experiment 2

Condition	RT	SE	%Error
Frequency pattern			
Visual baseline task			
Position	461	18	2.33
Direction	532	22	2.83
Visual filtering task			
Position	486	26	2.58
Direction	574	22	3.00
Location pattern			
Visual baseline task			
Position	463	19	1.83
Direction	523	21	1.83
Visual filtering task			
Position	477	20	2.33
Direction	582	22	3.58

Note—RT, reaction time (in milliseconds); SE, standard error.

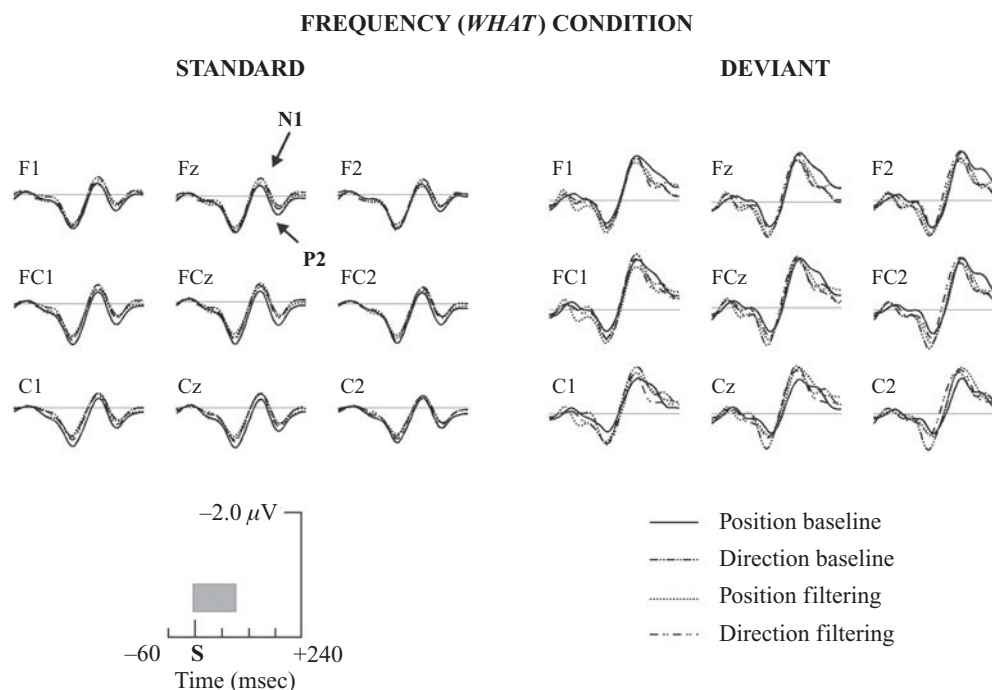


Figure 7. Group mean ($n = 12$) event-related brain potentials from nine frontocentral electrode sites in the auditory frequency (i.e., *what*) conditions in Experiment 2. Waveforms are shown separately for standard and deviant stimuli, and each individual plot represents the auditory-evoked response during one of the four visual conditions (position baseline, direction baseline, position filtering, and direction filtering). The gray rectangle indicates the duration of the stimulus.

HSD test revealed P2 amplitude to be significantly greater ($p < .05$) in the position (i.e., *where*) baseline condition than in the direction (i.e., *what*) baseline condition.

Difference waves between standard and deviant stimuli in the frequency condition were calculated in order to assess the extent to which frequency MMN was modulated by visual task demands (see Figure 8). Frequency pattern MMN was observed for all four conditions, with peak amplitude occurring around 168 msec after stimulus onset. Despite Figure 8's suggesting greater MMN amplitude at frontal sites for the position baseline condition, the results of the various ANOVAs failed to reveal any statistically significant main effects or interactions as a consequence of visual task load for peak amplitude, peak latency, or mean amplitude.

Location (*where*) pattern condition. Figure 9 shows the group average auditory ERPs produced from standard and deviant stimuli in the location condition. Analysis of the N1 (approximately 139 msec) and P2 (approximately 185 msec) components generated by the standard stimuli failed to reveal any statistically significant main effects or interactions.

Location difference waves were calculated in a way identical to that for frequency difference waves (see Figure 10). However, the differences between locational standards and deviants revealed a qualitatively different

pattern of results than did the comparison between stimulus types in the frequency condition. Specifically, locational difference waves were generally noisy and did not exhibit clear increased negativity around the time observed for frequency MMN. Indeed, group differences between locational standards and deviants appeared to be marginal or characterized by general enhanced negativity beginning after stimulus onset and continuing throughout the epoch. Therefore, locational MMN could not be quantified in any reliable way.

Discussion

In Experiment 2, we investigated the claim that if auditory ERPs were modulated by task content as well as by task difficulty, significant interactions between visual task (i.e., position and direction) and the type of auditory ERP examined (i.e., frequency and location) should have been observed, according to shared *what* and *where* codings of the stimuli. However, the only statistically significant finding to be revealed in the auditory ERP data was greater P2 peak amplitude for frequency standards during the position (i.e., *where*) baseline condition, relative to the direction (i.e., *what*) baseline condition. The data from Experiment 2 are ambiguous with respect to a task specificity hypothesis, for two reasons. First, the direction of this effect was not reversed with respect

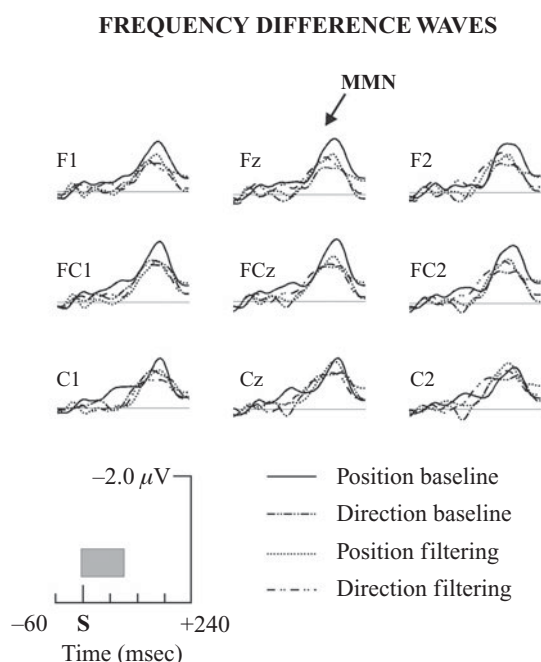


Figure 8. Group mean ($n = 12$) difference waves between the standard and the deviant stimuli during auditory frequency (i.e., *what*) presentation over nine frontocentral electrode sites in Experiment 2. Each line represents a single visual condition (position baseline, direction baseline, position filtering, and direction filtering). The gray rectangle indicates the duration of the stimulus. MMN, mismatch negativity.

to location standards, as would be anticipated with a strict reading of the task specificity account. Specifically, processing visual positional information should have also inhibited auditory locational analysis (due to the division of the putative *where* resource between the senses), relative to visual directional processing, in the same respect that visual positional information was observed to facilitate auditory frequency analysis (due to the use of independent *what* and *where* resources for audition and vision, respectively) in Experiment 2. Second, the observed pattern of data is also consistent with a generic task difficulty account. That is, auditory-evoked responses were enhanced when the participants took part in an easy (i.e., position baseline, 462 msec), relative to a harder (i.e., direction baseline, 528 msec), visual condition, thereby allowing greater attentional resources for audition. It will also be noted that the direction of this P2 effect was in opposition to that observed in Experiment 1. However, no concomitant effect on N1 was observed in Experiment 2, and therefore, the specific direction of the ERP modulation may have been the result of a number of factors, including the interaction between various neural components and the degree to which certain memory systems were implicated (Valtonen et al., 2003).

Therefore, it seems that the more conservative conclusion to be drawn from Experiment 2 is further sup-

port for process-general auditory ERPs (i.e., P2) being sensitive to the overall level of visual task difficulty (Valtonen et al., 2003; Woods et al., 1992). However, task difficulty manipulation in the present experiment was a secondary feature of the design, and therefore, the effect was not as large as that observed in Experiment 1.

EXPERIMENT 3

Auditory ERPs Under Baseline Conditions

Perhaps the most striking feature of the auditory ERP data in Experiment 2 is the apparent difference in morphology between standards and deviants produced by frequency and location patterns. Whereas frequency pattern seemed to produce a standard MMN, locational pattern did not (cf. Figures 7 and 9). Instead, location deviants were characterized by increased negativity relative to location standards from stimulus onset, which continued throughout the epoch. Possible reasons for the absence of a traditional location MMN were examined in Experiment 3. One possibility was that frequency MMN was more robust than locational MMN (perhaps due to differences in the salience between frequency and location values). Consequently, basing both frequency and location auditory ERPs on two blocks in each visual task, as in Experiment 2, may have been sufficient to produce MMN for frequency pattern but insufficient to show locational pattern MMN. As a result, we compared the generation of electrical differences between frequency and location standards and deviants over two and four blocks of presentation in Experiment 3. Averages taken from the first two blocks would provide a suitable comparison with the averages yielded for each of the visual conditions in Experiment 2, whereas a comparison between the two- and the four-block conditions would provide some indication as to the robustness of these differences.

A second possibility was that locational MMN failed to be produced due to a misconceptualization of the visual tasks in Experiment 2. That is, every response in all of the visual conditions could be coded according to location (i.e., up, down, left, and right).⁹ Therefore, the very *absence* of locational MMN across all four visual conditions may, in fact, be evidence for the task specificity hypothesis, in that the visual processing system was so concerned with *where* coding that only *what* coding in the auditory domain survived these attentional demands. This begs the question; if participants were to ignore the auditory stimuli under a control condition in which no visual task (and hence, no explicit location-based coding) was required, what morphology would frequency and location auditory ERPs take under this baseline condition? This idea of conducting a *no visual task* baseline condition was also useful for exploring a final hypothesis, in that evidence for auditory MMN modulation as a result of task difficulty may be contingent upon the comparison between the *active* visual conditions reported in Experiment 2 and a *passive* baseline

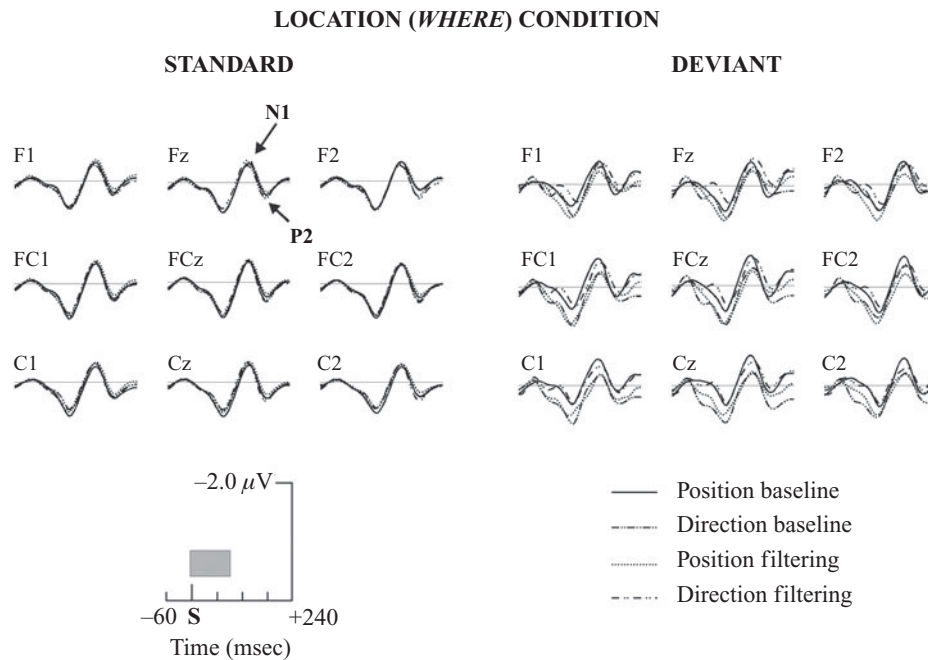


Figure 9. Group mean ($n = 12$) event-related brain potentials from nine frontocentral electrode sites during auditory location (i.e., *where*) presentation in Experiment 2. Waveforms are shown separately for standard and deviant stimuli, and each individual plot represents the auditory-evoked response during one of the four visual conditions (position baseline, direction baseline, position filtering, and direction filtering). The gray rectangle indicates the duration of the stimulus.

condition in Experiment 3. In order to increase the validity of the comparison between these two experiments as far as possible, the same participants were used for a second time.

Method

Participants

Eleven of the 12 participants in Experiment 2 gave their informed consent to take part in Experiment 3. All volunteers received an additional honorarium for their participation.

Design and Procedure

All aspects of auditory presentation and data analysis were as per Experiment 2. The participants were played short blocks of a frequency or a location pattern over headphones while watching a movie with subtitles. A total of eight blocks (four frequency and four location) were recorded. ERP data were initially averaged according to the first two blocks of each type of pattern to yield an appropriate passive condition baseline which to compare the active visual conditions in Experiment 2. Auditory-evoked waveforms for both conditions were then averaged over all four blocks, in order to assess the robustness of frequency and location MMN. Block presentation was randomized across participants.

Results

Figure 10 shows the group average waveforms elicited by frequency and location standards and deviants, over both two and four blocks of trials. The results from Experiment 3 replicated the general morphological differences between standards and deviants as a function of

the acoustic pattern presented in Experiment 2. Specifically, frequency deviants produced additional negativity, relative to frequency standards, which temporally overlapped with the N1 and P2 components. In contrast, location deviants elicited greater negativity, relative to location standards, following stimulus onset, which continued throughout the epoch. Therefore, the ERPs produced by frequency and location patterns were analyzed differently. Specifically, N1 and P2 were examined for frequency standards only, whereas P1, N1, and P2 deflections were quantified for both location standards and deviants. Therefore, N1 and P2 for frequency standards were examined in two-way repeated measures ANOVAs with the factors of block (two or four) and electrode (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, or C2), whereas ERPs for location pattern were assessed in a three-way repeated measures ANOVA for stimulus (standard or deviant) \times block (two or four) \times electrode (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, or C2).

Frequency and Location Standard and Deviant Waves

An analysis of frequency standards failed to reveal any theoretically relevant significant main effects or interactions with respect to peak amplitude, peak latency, or mean amplitude, with N1 and P2 peaks occurring at 135 and 174 msec, respectively, after stimulus onset. As for location ERPs, peak latency also failed to reveal any

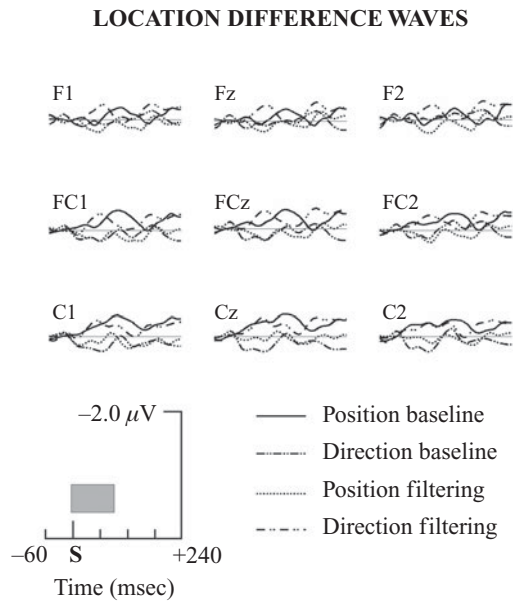


Figure 10. Group mean ($n = 12$) difference waves between the standard and the deviant stimuli during auditory location (i.e., where) presentation over nine frontocentral electrode sites in Experiment 2. Each line represents a single visual condition (position baseline, direction baseline, position filtering, and direction filtering). The gray rectangle indicates the duration of the stimulus.

significant main effects or interactions, yielding average latency of the P1 at 74 msec, the N1 at 137 msec, and the P2 at 180 msec. The negative shift in location deviants relative to location standards observed in Figure 11 was reflected in significant stimulus main effects for P1 mean amplitude [$F(1,10) = 8.77, p = .014$], N1 peak amplitude [$F(1,10) = 8.97, p = .013$], and peak and mean P2 amplitude [$F(1,10) = 8.05, p = .018$, and $F(1,10) = 10.76, p = .008$, respectively]. In addition, a significant main effect of block was revealed for N1 peak amplitude [$F(1,10) = 8.99, p = .013$] and P2 mean amplitude [$F(1,10) = 9.07, p = .013$]. Figure 10 shows that N1 was reduced when this ERP component was averaged over four blocks ($-0.35 \mu\text{V}$), relative to two blocks ($-0.42 \mu\text{V}$), whereas P2 showed an increase in positivity as the number of blocks that were averaged increased (0.17 and $0.25 \mu\text{V}$, for two and four blocks, respectively). All other main effects and interactions failed to reach statistical significance.

Frequency and Location Difference Waves

MMN for both frequency and location pattern was assessed in Experiment 3 with a three-way repeated measures ANOVA, relative to the factors of auditory pattern (frequency or location), block (two or four), and the standard nine-level frontocentral factor of electrode. Peak latency failed to significantly modulate according to the above factors, with peak MMN being observed, on average, 165 msec after stimulus onset. Peak amplitude re-

vealed a statistically significant main effect of block [$F(1,10) = 5.42, p = .042$], suggesting a reduction in the magnitude of MMN when averaging is done on the basis of four ($0.71 \mu\text{V}$) rather than two ($0.82 \mu\text{V}$) blocks, although this effect failed to be replicated with mean amplitude. More important, both peak and mean amplitude revealed a significant main effect of auditory pattern [$F(1,10) = 6.41, p = .030$, and $F(1,10) = 7.78, p = .019$, respectively]. As is graphically represented in Figure 12, mean MMN produced in the frequency pattern condition ($-0.75 \mu\text{V}$) was significantly greater than any maximal negativity generated around the same time window in the location pattern condition ($-0.24 \mu\text{V}$). All other main effects and interactions were nonsignificant.

Intrasubject Comparisons of Frequency MMN

For the 11 participants who completed both investigations, mean amplitude frequency MMN was compared between data from the two-block condition in Experiment 3 and the location baseline visual condition in Experiment 2. This comparison between the control condition in Experiment 3 and the easiest visual task in Experiment 2 provided a stringent test of the hypothesis that visual task load, in its broadest conception, would modulate change detection for acoustic frequency processing. A related t test revealed the respective amplitudes from the control condition and the easiest visual task to be nonsignificantly different from one another [$t(10) = 0.30, p = .771$]. A more relaxed test of this *task versus no task* hypothesis, comparing frequency MMN in Experiment 3 with the average frequency MMN produced over all four visual conditions in Experiment 2, also failed to reveal a statistically significant effect [$t(10) = -0.83, p = .426$].

Discussion

The ERP data in Experiment 3 reveal a number of findings. First, similar frequency MMN amplitudes were generated under conditions of passive listening in the absence of a visual task in Experiment 3, relative to when the participants were actively completing various kinds of visual processing in Experiment 2. Therefore, a simple dichotomy between the presence and the absence of a visual task was also insufficient to produce MMN modulation. Relative to Experiment 1, however, this is perhaps none too surprising, since modulation of task difficulty was a secondary design feature of Experiment 2. Second, ERPs to locational pattern seemed to be more sensitive to averaging over different numbers of blocks than were frequency pattern ERPs. Therefore, one possibility is that in Experiment 2, in which the participants were exposed to eight blocks of frequency and location patterns, locational registration and the concomitant ERP differences between the standard and the deviant stimuli were simply produced against a noisier background. This should not have had a systematic effect upon the data, given the counterbalancing employed in Experiment 2, but the relative robustness of frequency

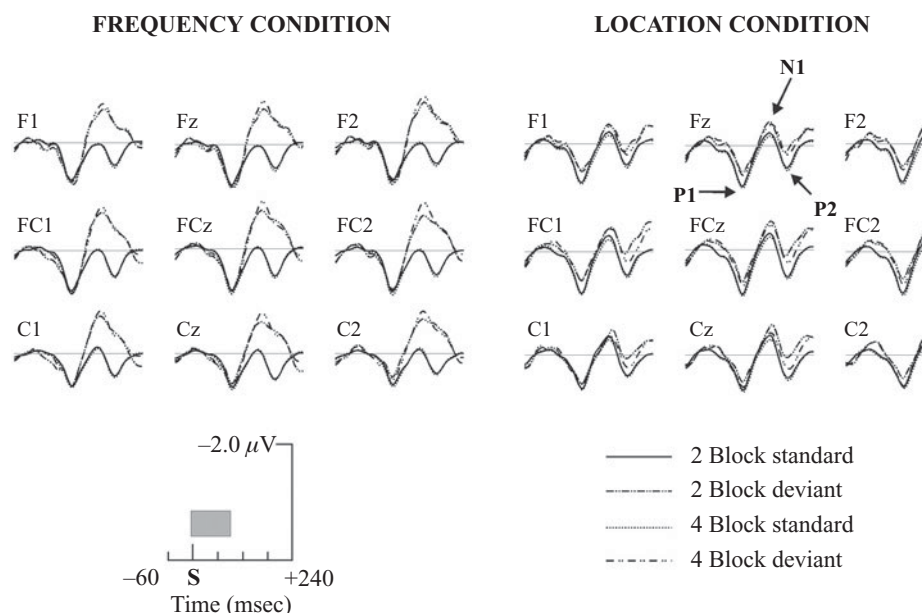


Figure 11. Group mean ($n = 11$) event-related brain potentials from nine frontocentral electrode sites during frequency and location pattern presentation in Experiment 3. The individual plots represent standard and deviant stimuli averaged on the basis of either two or four blocks of presentation. The gray rectangle indicates the duration of the stimulus.

and locational detection should be pursued in further work.

Third, signature MMN for repetition-defined location deviance failed to be recorded for a second time. Therefore, the absence of locational MMN in Experiment 2 cannot have been due to a misconceptualization of the visual tasks, in that all four visual conditions were predominately locational in nature, thereby limiting the detection of locational deviance in a secondary (i.e., auditory) sense. Moreover, the data from Experiment 3 replicated the finding in Experiment 2 that representation of locational pattern deviants by the auditory system was qualitatively different from the waveform elicited by frequency pattern deviants. Specifically, locational pattern deviance was reflected in an overall increase in negativity after stimulus onset that continued throughout the epoch, relative to frequency pattern deviance, which was characterized by increased negativity around N1 and P2 (i.e., standard MMN). Therefore, the isolation of *locational MMN* appears to raise problems when deviance is defined according to pattern and, more specifically, locale repetition, since the increased deviant negativity around 150 msec, which traditionally signals MMN, does not substantially differ from the increased negativity observed earlier in the epoch.

One possible account as to why the auditory system shows differential sensitivity to locational and frequency pattern deviation may be derived from a consideration of current theoretical views regarding auditory object formation. Kubovy (1981; Kubovy & Van Valkenburg, 2001) appeals to the notion of indispensable attributes

for certain senses, specifically claiming that for audition, frequency acts as the auditory *edge* that defines and segregates an auditory object from other energy within the field (although see the recent communiqués between Neuhoff, 2003, and Van Valkenburg & Kubovy, 2003). The purported significance of frequency information, relative to locational information, in acoustic processing has also been emphasized by Woods, Alain, Diaz, Rhodes, and Ogawa (2001). Here, a model of auditory feature binding was put forward in a manner analogous to visual feature integration theory (Treisman & Gelade, 1980). Specifically, Woods et al. (2001) stated that frequency in audition has a status analogous to location in vision, in that the constituent features of an object remain unbound until attention is focused on this critical property. This is similar to Kubovy's indispensable attribute account, in that frequency invariance may be much more critical for tracking the same acoustic object over time than is locational invariance (see also Dyson & Quinlan, 2004). Therefore, if location changes were not as critical as frequency changes in terms of auditory object formation and maintenance, it follows that change detection (as indexed by MMN) should be less sensitive in the former case.

This is, of course, not to deny the results of studies that have demonstrated locational MMN under standard oddball conditions (Deouell & Bentin, 1998; Schröger, 1995; Schröger & Wolff, 1997). However, as far as we are aware, this the first attempt to produce locational MMN defined by locale repetition, and the absence of a traditional MMN effect in both Experiments 2 and 3 sug-

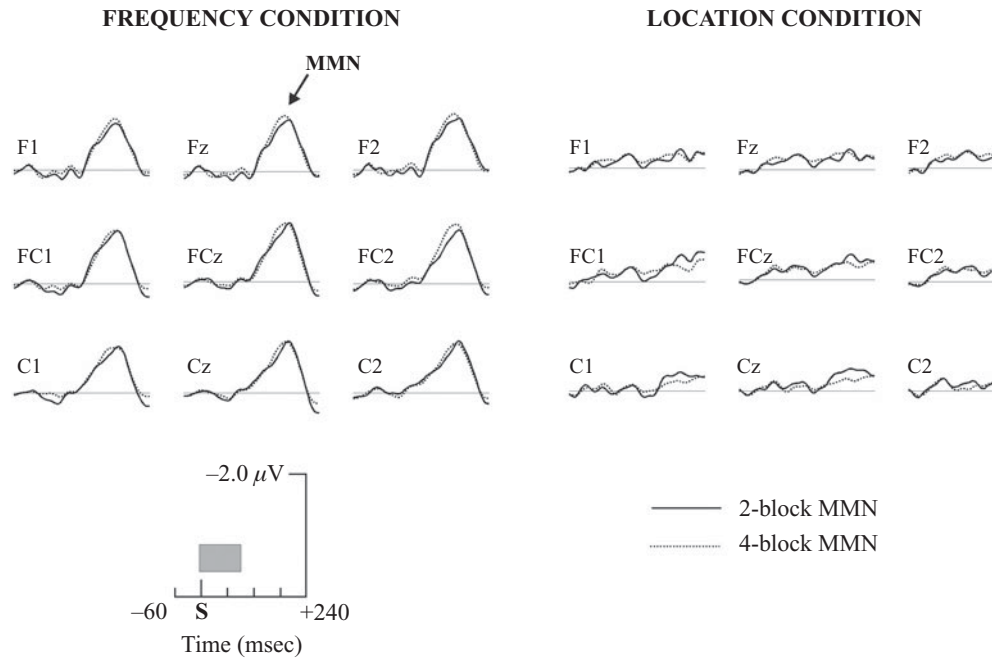


Figure 12. Group mean ($n = 11$) difference waves between the standard and the deviant stimuli during frequency and location pattern presentation over nine frontocentral electrode sites in Experiment 3. The individual plots represent difference waves averaged on the basis of either two or four blocks of presentation. The gray rectangle indicates the duration of the stimulus. MMN, mismatch negativity.

gests critically different mechanisms with respect to the way in which the auditory system detects and tracks variation along these two acoustic dimensions.

GENERAL DISCUSSION

The present experimental series was an attempt to further elucidate the nature of interaction between visual and auditory senses—specifically, with respect to the way in which putative processing resources were distributed during the active completion of visual tasks and the passive registration of changes in acoustic properties. In Experiment 1, we initially examined the effect of general visual task difficulty on auditory ERPs produced for both temporal (i.e., MMN) and spectral (i.e., ORN) patterns. The behavioral data revealed a successful manipulation of task difficulty, which was reflected in auditory ERPs as a concomitant reduction in N1 and an increase in P2. A distinction was made between process-general (i.e., N1 and P2) and process-specific (i.e., MMN and ORN) auditory-evoked components, with evidence for shared processing resources between the senses occurring in the former group. In Experiment 2, we set out to test a secondary hypothesis that modulation of auditory ERPs by visual task may also be dependent on similar processing codes between the two senses. To this end, both visual and auditory stimuli were defined relative to variation in generic *what* and *where* dimensions, and the interaction between the two domains was assessed. Once again, P2 amplitude was shown to vary according to vi-

sual task demands. Although this effect was consistent with a task specificity hypothesis, the data were also supportive of a general task difficulty account, as put forward in Experiment 1. Experiment 3 was conducted in order to resolve a number of outstanding issues in Experiment 2. Specifically, Experiment 3 confirmed that the absence of traditional location MMN in Experiment 2 was replicable and that the auditory system may deal with pattern deviation as defined by frequency and locale repetition in very different ways.

The present data stand in contrast to previous research showing interactions between visual and auditory stimulus processing that are driven by task overlap (e.g., Tellinghuisen & Nowak, 2003) or contingent upon attention to both senses (e.g., Talsma & Kok, 2001). Specifically, the present experimental series demonstrated that manipulating the degree of general attentional load in vision influenced purportedly exogenous neural responses to auditory stimuli presented outside the field of attention. This suggests a fundamental constraint in the ability to allocate separate processing resources to different senses. However, this conclusion must be taken with a degree of caution, because although extreme positions regarding the interaction of processing resources between vision and audition are in evidence (see Tellinghuisen & Nowak, 2003, for a discussion of these views), the present research shows that these positions are perhaps untenable, since this interaction depends on a number of factors. The article will conclude with two key factors that may help delineate experimental cases in

which sensory independence and interdependence may be observed.

As was revealed across Experiments 1 and 2, an examination of neural components, such as N1, that are thought to relate to mandatory processing mechanisms in audition (Näätänen & Picton, 1987) showed modulation as a result of visual task load, suggesting the use of shared processing resources between the senses. In contrast, an examination of additional neural components specific to certain kinds of sound segregation (i.e., ORN in Experiment 1 and MMN in Experiments 1 and 2) failed to show modulation as a result of visual task, thereby supporting the independent processing of vision and audition. Therefore, the data may provide evidence for either the interdependence (i.e., Kahneman, 1973) or the independence (i.e., Duncan et al., 1997) of resources across the senses, according to whether researchers focus on the impact of visual task upon process-general (i.e., N1 and P2) or process-specific (i.e., MMN and ORN) auditory ERP components. It should be noted, however, that this distinction is further compounded by differences in signal-to-noise ratio between these two types of components. It is a simple consequence of the calculation of MMN that deviant waveforms are based on a smaller number of trials than are standard stimuli. Therefore (and putting aside issues of intrasubject reliability for the moment), analysis of ERP components elicited for both standard and deviant stimuli that do not overlap the MMN latency range (i.e., P1 and possibly N1) are much more robust than deviant-specific deflections and the MMN itself. For example, in the temporal conditions in Experiment 1, deviant groups represented 20% of the trials, whereas standard groups made up the remaining 80%. Consequently, the fidelity of the signal was improved by a factor of four during standard stimuli analysis.¹⁰ One issue for future research, therefore, is how to dissociate a failure to modulate MMN with a poor signal-to-noise ratio (Cacace & McFarland, 2003; Picton, Alain, et al., 2000). This highlights the importance of Experiment 3 in the present series and the empirical investigation of ERP component reliability under specific experimental contexts.

A secondary, although no less important, factor is the conceptualization of visual task employed in the investigation. Although Experiment 2 revealed little evidence for the task specificity hypothesis as defined by auditory and visual *what* and *where* systems (Alain, Arnott, Hevenor, et al., 2001; Ungerleider & Mishkin, 1982), this is not to deny that broader aspects of the visual task do not influence the presence of an effect. For example, employing an extreme manipulation of task difficulty as in Experiment 1 appears to be relevant to the observation of resource constraints, although how one achieves external validation of task difficulty in the absence of modulation in a secondary sense is a problematic issue (Valtonen et al., 2003). However, robust differences in task difficulty may not be sufficient in and of themselves to produce critical changes in the processing resources available for other tasks in other senses. For example,

Otten et al. (2000) emphasized the importance of using continuous, as opposed to discrete, preliminary tasks, since the latter case may allow for attentional switching between the two senses. Most recently, Yucel, Petty, McCarthy, and Belger (in press) have demonstrated frequency MMN modulation at frontocentral sites as a result of increasing the difficulty of a continuous visual-tracking task. Therefore, while MMN modulation driven by intermodal differences is a novel finding, and not one that we were able to reproduce in the present experimental series, detailed examination of the visual paradigm in Yucel et al.'s study would be a useful starting point for defining the conditions necessary for producing such a unique effect. In addition to the possibility that task differences give rise to sensory interaction, other issues such as stimulus salience need also to be considered, given the principle of inverse effectiveness as demonstrated by Meredith and Stein (1986). It is clear that only by considering the interaction between these and other experimental parameters will the conditions under which cross-sensory interaction occurs become known.

In conclusion, simple statements regarding the interdependence or independence of processing resources across the senses or regarding straightforward dissociations between those auditory ERP components that are sensitive to the influence of attention and those that are not fail to acknowledge the advanced state of the current literature. Rather, an increased awareness of the sensitivity of specific ERP components to intermodal conditions and a further specification and detailing of global experimental conditions will ultimately assist in setting the level of description most appropriate for sensory-independent frameworks of cognitive processing and resolve the degree to which initially separate sensory systems may, at a higher stage of operation, become isomorphic with one another.

REFERENCES

- ALAIN, C., ARNOTT, S. R., HEVENOR, S., GRAHAM, S., & GRADY, C. L. (2001). "What" and "where" in the human auditory cortex. *Proceedings of the National Academy of Sciences*, **98**, 12301-12306.
- ALAIN, C., ARNOTT, S. R., & PICTON, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 1072-1089.
- ALAIN, C., & IZENBERG, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of Cognitive Neuroscience*, **15**, 1063-1073.
- ALAIN, C., SCHULER, B. M., & McDONALD, K. L. (2002). Neural activity associated with distinguishing concurrent auditory objects. *Journal of the Acoustical Society of America*, **111**, 990-995.
- ALAIN, C., & WOODS, D. L. (1997). Attention modulates auditory pattern memory as indexed by event-related brain potentials. *Psychophysiology*, **34**, 534-546.
- ALAIN, C., WOODS, D. L., & OGAWA, K. H. (1994). Brain indices of automatic pattern processing. *NeuroReport*, **6**, 140-144.
- ALHO, K., SAMS, M., PAAVILAINEN, P., REINIKAINEN, K., & NÄÄTÄNEN, R. (1989). Event-related brain potentials reflecting processing of relevant and irrelevant stimuli during selective attention. *Psychophysiology*, **26**, 514-528.
- ALHO, K., WOODS, D. L., & ALGAZI, A. (1994). Processing of auditory

- stimuli during auditory and visual attention as revealed by event-related potentials. *Psychophysiology*, **31**, 469-479.
- ALHO, K., WOODS, D. L., ALGAZI, A., & NÄÄTÄNEN, R. (1992). Intermodal selective attention: II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography & Clinical Neurophysiology*, **82**, 356-368.
- ANOUROVA, I., NIKOULINE, V. V., ILMONIEMI, R. J., HOTTA, J., ARONEN, H. J., & CARLSON, S. (2001). Evidence for dissociation of spatial and nonspatial auditory information processing. *NeuroImage*, **14**, 1268-1277.
- BADDELEY, A., & HITCH, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-90). New York: Academic Press.
- BELIN, P., & ZATORRE, R. J. (2000). "What," "where" and "how" in auditory cortex [Letter to the editor]. *Nature Neuroscience*, **3**, 965-966.
- CACACE, A. T., & MCFARLAND, D. J. (2003). Quantifying signal-to-noise ratio of mismatch negativity in humans. *Neuroscience Letters*, **341**, 251-255.
- CLARKE, S., ADRIANI, M., & BELLMANN, A. (1998). Distinct short-term memory systems for sound content and sound localization. *NeuroReport*, **9**, 3433-3437.
- DEOUELL, L. Y., & BENTIN, S. (1998). Variable cerebral responses to equally distinct deviance in four auditory dimensions: A mismatch negativity study. *Psychophysiology*, **35**, 745-754.
- DUNCAN, J., MARTENS, S., & WARD, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, **387**, 808-810.
- DYSON, B. J., & QUINLAN, P. T. (2004). Stimulus processing constraints in audition. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 1117-1131.
- FARRELL, W. S., JR. (1979). Coding left and right. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 42-51.
- GOODALE, M. A., & MILNER, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20-25.
- HILLYARD, S. A., HINK, R., SCHWENT, V., & PICTON, T. (1973). Electrical signs of selective attention in the human brain. *Science*, **182**, 177-180.
- KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- KUBOVY, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 55-98). Hillsdale, NJ: Erlbaum.
- KUBOVY, M., & VAN VALKENBURG, D. (2001). Auditory and visual objects. *Cognition*, **80**, 97-126.
- MCAARTHUR, G. M., BISHOP, D. V. M., & PROUDFOOT, M. (2003). Do video sounds interfere with auditory event-related potentials? *Behavior Research Methods, Instruments, & Computers*, **35**, 329-333.
- MCDONALD, J. J., TEDER-SÄLEJÄRVI, W. A., DI RUSSO, F., & HILLYARD, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, **15**, 10-19.
- MCDONALD, J. J., & WARD, L. M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science*, **11**, 167-171.
- MEREDITH, M. A., & STEIN, B. E. (1986). Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, **56**, 640-662.
- MÜLLER, B. W., ACHENBACK, C., OADES, R. D., BENDER, S., & SCHALL, U. (2002). Modulation of mismatch negativity by stimulus deviance and modality of attention. *NeuroReport*, **13**, 1317-1320.
- NÄÄTÄNEN, R. (1975). Selective attention and evoked potentials in humans: A critical review. *Biological Psychology*, **2**, 237-307.
- NÄÄTÄNEN, R., & PICTON, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, **24**, 375-425.
- NÄÄTÄNEN, R., & WINKLER, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Review*, **125**, 826-859.
- NEUHOFF, J. G. (2003). Pitch variation is unnecessary (and sometimes insufficient) for the formation of auditory objects. *Cognition*, **87**, 219-224.
- OTTEN, L. J., ALAIN, C., & PICTON, T. W. (2000). Effects of visual attentional load on auditory processing. *NeuroReport*, **11**, 875-880.
- PICTON, T. W., ALAIN, C., OTTEN, L., RITTER, W., & ACHIM, A. (2000). Mismatch negativity: Different water in the same river. *Audiology & Neuro-Otology*, **5**, 111-139.
- PICTON, T. W., VAN ROON, P., ARMILIO, M. L., BERG, P., ILLE, N., & SCHERG, M. (2000). The correction of ocular artifacts: A topographic perspective. *Clinical Neurophysiology*, **111**, 53-65.
- POMERANTZ, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General*, **112**, 515-540.
- RITTER, W., DEACON, D., GOMES, H., JAVITT, D. C., & VAUGHAN, H. G., JR. (1995). The mismatch negativity of event-related potentials as a probe of transient auditory memory: A review. *Ear & Hearing*, **16**, 52-67.
- SATTERFIELD, J. H. (1965). Evoked cortical response enhancement and attention in man: A study of responses to auditory and shock stimuli. *Electroencephalography & Clinical Neurophysiology*, **19**, 470-475.
- SCHRÖGER, E. (1995). Processing of auditory deviants with changes in one versus two stimulus dimensions. *Psychophysiology*, **32**, 55-65.
- SCHRÖGER, E., & WOLFF, C. (1997). Fast preattentive processing of location: A functional basis for selective listening in humans. *Neuroscience Letters*, **232**, 5-8.
- SPONG, P., & LINDSLEY, D. B. (1968). Cortical evoked responses and attentiveness in man: Differential effects of selective attentiveness and general alertness level. *Electroencephalography & Clinical Neurophysiology*, **24**, 396-397.
- TALSMAN, D., & KOK, A. (2001). Nonspatial intermodal selective attention is mediated by sensory brain areas: Evidence from event-related potentials. *Psychophysiology*, **38**, 736-751.
- TELLINGHUISEN, D. J., & NOWAK, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, **65**, 817-828.
- TREISMAN, A., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- UNGERLEIDER, L. G., & MISHKIN, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.
- VALTONEN, J., MAY, P., MÄKINEN, V., & TIITINEN, H. (2003). Visual short-term memory load affects sensory processing of irrelevant sounds in human auditory cortex. *Cognitive Brain Research*, **17**, 358-367.
- VAN VALKENBURG, D., & KUBOVY, M. (2003). In defense of the theory of indispensable attributes. *Cognition*, **87**, 225-233.
- WINER, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- WINKLER, I., SUSSMAN, E., TERVANIEMI, M., HORVÁTH, J., RITTER, W., & NÄÄTÄNEN, R. (2003). Preattentive auditory context effects. *Cognitive, Affective, & Behavioral Neuroscience*, **3**, 57-77.
- WOLDORFF, M. G., HACKLEY, S. A., & HILLYARD, S. A. (1991). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, **28**, 30-42.
- WOODS, D. L., ALAIN, C., DIAZ, R., RHODES, D., & OGAWA, K. H. (2001). Location and frequency cues in auditory selective attention. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 65-74.
- WOODS, D. L., ALHO, K., & ALGAZI, A. (1992). Intermodal selective attention: I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography & Clinical Neurophysiology*, **82**, 356-368.
- YUCEL, H. G., PETTY, C., MCCARTHY, G., & BELGER, A. (in press). Graded visual attention modulates brain responses evoked by task-irrelevant auditory pitch changes. *Journal of Cognitive Neuroscience*.

NOTES

1. These nine electrode sites were selected on the basis of previous research (Alain & Izenberg, 2003) indicating strongest MMN activation at frontocentral sites and strongest ORN activation at central sites.
2. In three individual cases, RTs and error rates were based on only two blocks of behavioral data (as opposed to three), due to recording errors.
3. Winer (1962) has suggested that when data are expressed as percentages, the arcsin transformation should be used.
4. Although tangential to the thrust of the article, it is also important to point out that the ERP modulation observed as a result of the spectral properties of the auditory stimuli—specifically, the increase in N1

mean amplitude and the decrease in P2 mean amplitude for mistuned, relative to tuned, stimuli (see Figure 2)—is consistent with previous examinations of the neural correlates of inharmonicity (Alain, Arnott, & Picton, 2001).

5. We thank an anonymous reviewer for these suggestions.

6. This was also supported by the nonsignificance of a related t test upon mean amplitude observed between 100 and 200 msec after stimulus onset (the particular region of interest relating to ORN) between the easy and the hard visual tasks [$t(3) = 0.05$, $p = .962$ for auditory ERP baseline].

7. Readers should be mindful that the opposite direction of N1 and P2 effects in the temporal pattern condition might also be reflective of a general sustained positive shift in the easy condition, relative to the hard condition (see Figure 4).

8. This simple distinction between ventral *what* and dorsal *where* processing in both vision and audition, however, is likely to be an oversimplification of a much more complex situation. Alternative functionality for the dorsal route has been proposed by Goodale and Milner (1992) in vision and by Belin and Zatorre (2000) in audition.

9. We thank Terry Picton and Roxane Itier for this alternative interpretation.

10. A similar argument regarding differences in the signal-to-noise ratio could also be made regarding the observation of near-significant task load modulation of ORN, relative to MMN, in Experiment 1.

(Manuscript received June 4, 2004;
revision accepted for publication June 6, 2005.)