

# Selective Attention and Multisensory Integration: Multiple Phases of Effects on the Evoked Brain Activity

Durk Talsma and Marty G. Woldorff

## Abstract

■ We used event-related potentials (ERPs) to evaluate the role of attention in the integration of visual and auditory features of multisensory objects. This was done by contrasting the ERPs to multisensory stimuli (AV) to the sum of the ERPs to the corresponding auditory-only (A) and visual-only (V) stimuli [i.e., AV vs. (A + V)]. V, A, and VA stimuli were presented in random order to the left and right hemispaces. Subjects attended to a designated side to detect infrequent target stimuli in either modality there. The focus of this report is on the ERPs to the standard (i.e., nontarget) stimuli. We used rapid variable stimulus onset asynchronies (350–650 msec) to mitigate anticipatory activity and included “no-stim” trials to estimate and remove ERP overlap from residual anticipatory processes and from adjacent stimuli in the sequence. Spatial attention effects on the processing of

the unisensory stimuli consisted of a modulation of visual P1 and N1 components (at 90–130 msec and 160–200 msec, respectively) and of the auditory N1 and processing negativity (100–200 msec). Attended versus unattended multisensory ERPs elicited a combination of these effects. Multisensory integration effects consisted of an initial frontal positivity around 100 msec that was larger for attended stimuli. This was followed by three phases of centro-medially distributed effects of integration and/or attention beginning at around 160 msec, and peaking at 190 (scalp positivity), 250 (negativity), and 300–500 msec (positivity) after stimulus onset. These integration effects were larger in amplitude for attended than for unattended stimuli, providing neural evidence that attention can modulate multisensory-integration processes at multiple stages. ■

## INTRODUCTION

In everyday life a considerable number of cognitive processes depend on the integration of information from multiple senses. Despite this importance, the scientific study of the physiological processes underlying multisensory integration in humans has been heretofore rather limited (Foxy & Schroeder, 2002). Until relatively recently, studies of multisensory integration processes focused mostly on either animal physiology (e.g., Stein & Meredith, 1993; Wallace, Meredith, & Stein, 1992) or human behavior (e.g., Stein, London, Wilkinson, & Price, 1996; Saldaña & Rosenblum, 1993; McGurk & MacDonald, 1976). Such behavioral studies have shown, for example, that the simultaneous, or near-simultaneous, presentation of an auditory stimulus can influence the perceived temporal characteristics of a visual stimulus (e.g., Shams, Kamitani, & Shimojo, 2000). Similarly, visual stimuli are known to affect the perceived location of an auditory stimulus (e.g., Bertelson, 1999; Platt & Warren, 1972; Pick, Warren, & Hay, 1969), the illusion known as the “ventriloquism”

effect. Furthermore, McDonald, Teder-Sälejärvi, and Hillyard (2000) and McDonald and Ward (1999, 2000) have shown that irrelevant sounds at a given location can enhance the perception of visual stimuli at that same location.

In general, one can conclude that the simultaneous, or near-simultaneous, presentation of stimuli in two sensory modalities reduces stimulus ambiguity (e.g., Calvert, Campbell, & Brammer, 2000), or enhances stimulus detection (Stein et al., 1996). It is still unclear, however, whether this integration process occurs preattentively for all perceived multisensory objects and does not require attention, or whether attended objects are integrated differently from those that are not attended. A large number of event-related potential (ERP) studies have shown that the human brain is highly effective in relatively attenuating the further processing of irrelevant stimuli. These ERP studies have shown that this filtering, as a result of selectively attending, can take place at least as early as 80 msec after stimulus onset for visual stimuli (see Wijers, Mulder, Gunter, & Smid, 1996, for a review), or as early as 20 msec after auditory stimulation (Woldorff, Gallen, et al., 1993; Woldorff & Hillyard, 1991). Therefore, one

might expect that multisensory integration does take place for stimuli that are task relevant (i.e., attended), but may not do so, or may do less so for stimuli that are task irrelevant (or unattended).

Although the number of studies that have investigated interactions between attention and integration of multisensory input—using any methodology—is still rather small, Bertelson, Vroomen, De Gelder, and Driver (2000) and Bertelson (1999) have argued on the basis of behavioral studies that such interactions with attention do not occur for the multisensory integration processes underlying the ventriloquism effect. In addition, Driver (1996) has shown that the localization of speech sounds can be influenced by lip reading and has argued that integration of audiovisual stimuli occurs at a preattentive processing stage. Thus, although the abovementioned studies comprise only behavioral results, they argue against the hypothesis that attention can enhance or otherwise modulate multisensory integration, but that integration is a preattentive process instead. This finding therefore also suggests that integration of audiovisual inputs takes place at a very early stage in the processing stream.

The possibility that such integration effects occur at early preattentive stages would fit with some results from a couple of recent multisensory ERP studies. More specifically, Giard and Peronnet (1999), applying an approach from some earlier animal single-unit studies (see Stein & Meredith, 1993), studied multisensory integration using ERPs. Their method was based on the notion that the early parts of the ERP waveforms reflect activity mainly in the sensory processing pathways. Therefore, under the assumption that these electrical activities summate at the scalp, one can sum the ERP responses elicited by auditory-only (A) and visual-only (V) objects together and compare the sum of the combined (A + V) waveforms to ERPs elicited by true multisensory objects (AV) that are composed of the same auditory and visual components as the unisensory objects. The effects of multisensory integration are expected to be found as differences between the AV and A + V waveforms, which can simply be expressed mathematically as  $[AV - (A + V)]$ . Based on this approach, Giard and Peronnet (1999) reported that such integrative processes can take place as early as 40 msec in the sensory processing chain and, based on their scalp distributions, suggested that they operate in both sensory-specific and nonspecific cortical structures. Along similar lines, Fort, Delpuech, Pernier, and Giard (2002b) and Molholm et al. (2002) also reported an early enhancement effect and suggested that the onset of the early effect comprises a modulation of the early C1 component, which is thought to represent initial activity in the primary visual cortex.

As pointed out by Teder-Sälejärvi, McDonald, Di Russo, and Hillyard (2002), however, a potential prob-

lem with these studies derived from the possible occurrence of prestimulus anticipatory waves in conjunction with the analysis approach employed. In particular, the relatively long ISIs might have led to the production of anticipatory ERP waves, known as the contingent negative variation (CNV), due to stimulus expectation (Walter, Cooper, Aldridge, McCallum, & Winter, 1964) that could overlap onto the beginning of the actual stimulus responses. Due to the nature of the contrast of multisensory versus sum of unisensory responses, this could lead to some problems. More specifically, if stimulus expectancy generates a CNV wave, it would be present for the unisensory visual, unisensory auditory, and multisensory audiovisual objects alike. Under such circumstances, the sum of the unisensory visual and unisensory auditory (A + V) waveforms would contain a contribution from two CNV waves, whereas ERP waveforms to the multisensory (AV) objects would contain a contribution from only one CNV. Therefore, the  $[AV - (A + V)]$  comparison could contain CNV artifacts that can be falsely identified as an early integration effect. Indeed, pursuing this possible artifactual source for an early effect, Teder-Sälejärvi et al. (2002) also found what appeared to be a similarly early effect of the  $[AV - (A + V)]$  integration. However, they were able to demonstrate that applying a high-pass filter to their data resulted in the elimination of this effect. The high-pass filter eliminated slow-wave activity (such as what would come from CNVs) but left intact higher-frequency components (such as would be present in the C1 wave or other early sensory cortical activity).

The main goal of the present study was to address the question of whether attention affects multisensory integration. To do so, we presented a rapid sequence of auditory-only, visual-only, and multisensory stimuli unilaterally to either the left or right hemispaces, at an average rate of two stimuli per second. During each 5-minute run, subjects focused on both the visual and auditory stimuli (and consequently also on the multisensory stimuli) that were presented in a designated hemispaces, while ignoring all stimuli in the opposite hemispaces. Thus, each hemispaces was attended in half of the runs and unattended in the other half. The subject's task was to detect infrequent target stimuli (in either or both modalities) within the attended hemispaces and to report these by making a speeded manual response. Target stimuli were characterized by a brief dip in intensity in the middle of the stimulus duration. In the case of AV objects, this intensity decrease was always present for both the visual and auditory features of the stimulus. A general finding in the behavioral literature has been that response times to stimuli containing redundant features are shorter than response times to either unimodal stimulus feature alone (Miller, 1982, 1986). Combined with the finding that multisensory

stimuli are more easily detectable (e.g., Stein et al., 1996), we expected that responses to the multisensory targets would not only be faster, but also more accurate, than the responses to the unisensory stimuli alone.

In addition, a secondary goal of the present study was to demonstrate that apparent effects of multisensory integration, stemming from artifactual sources, such as from anticipatory ERP waves or other overlapping ERP components, can be removed by applying a novel subtraction technique. To mitigate the generation of anticipatory ERP waveforms, which might have distorted the early ERPs (Teder-Sälejärvi et al., 2002), and to facilitate the selective focusing of attention, we used an average rate of stimulus presentation (of about 2 stimuli per second) that was much higher than the rates used in the studies of Molholm et al. (2002), Teder-Sälejärvi et al. (2002), and Giard and Peronnet (1999). At these fast stimulus presentation rates, however, one needs to deal with the methodological problem of overlap between the ERP responses to adjacent trials in the sequence (Woldorff, 1993), which could lead to similar problems as the ones caused by overlapping CNV activity. To address this problem, we applied an approach developed in the functional MRI literature (Buckner et al., 1998; Burock, Buckner, Woldorff, Rosen, & Dale, 1998) of including randomly intermixed “no-stim” trials with the other trial types as a means to estimate the overlapping ERP responses from adjacent trials (see Methods for details). After subtracting off the overlap using this approach, the ERP waveforms to auditory-only and visual-only stimuli were added together and compared with the ERP waveforms elicited by multisensory audiovisual objects, separately for the attended and unattended conditions.

Because one of the functions of both attention and selective attention is to enhance perception, our expectation was that multisensory integration and attention would interact. It has been shown previously that spatial attention is not limited to one sensory modality only, but spreads between visual and auditory modalities (Eimer, van Velzen, & Driver, 2002; Talsma & Kok, 2002; Eimer, Cockburn, Smedley, & Driver, 2001; Eimer & Driver, 2001; Talsma & Kok, 2001; Eimer & Schröger, 1998; Spence & Driver, 1996, 1997). In addition, McDonald and Ward (2000) and McDonald et al. (2000) have shown behaviorally that involuntary orienting to irrelevant sounds facilitates motor responses to visual stimuli presented at nearby locations, suggesting that spatial orientation to stimuli in one modality facilitates integrating auditory and visual stimuli. Therefore, we expected that, to further aid in perceptual enhancement, effects of multisensory integration would be larger for stimuli that were behaviorally (and therefore perceptually) relevant, that is, for those stimuli that were attended.

## RESULTS

### Behavioral Data

Table 1 presents the mean reaction times and hit rates for each target type. A main effect of the factor stimulus type confirmed that response times to visual-only, auditory-only, and multisensory stimuli differed significantly from each other [ $F(2,30) = 14.9; p < .001$ ]. Consistent with previous behavioral studies, subsequent planned comparisons showed that subjects responded faster to multisensory stimulus targets than to visual ones [ $F(1,15) = 26.9; p < .001$ ]. In addition, responses to multisensory stimuli were also significantly faster than responses to auditory stimuli [ $F(1,15) = 10.5; p < .01$ ]. Finally, responses to visual stimuli were somewhat slower than responses to auditory stimuli [ $F(1,15) = 6.32; p < .05$ ].

Hit rates showed similar effects as response times (Table 1). Hit rates were lowest for visual stimuli, intermediate for auditory stimuli, and highest for multisensory stimuli. These effects were statistically expressed as an effect of the within-subjects factor stimulus type [ $F(2,30) = 25.3; p < .001$ ]. Planned comparisons showed that accuracy differed significantly between visual and auditory stimuli [ $F(1,15) = 17.0; p < .001$ ], and visual and multisensory stimuli [ $F(1,15) = 67.2; p < .001$ ], but only marginally between auditory and multisensory stimuli [ $F(1,15) = 3.2; p < .1$ ]. Neither significant main effects of location (left vs. right), nor significant interactions between location and stimulus type were found.

### Event-related Potentials: Overlap Correction using No-Stims

In order to remove overlapping ERP activity from adjacent trials and from anticipatory CNV waves, ERPs to no-stim trials were subtracted from stimulus ERPs. To show

**Table 1.** Mean Response Times and Hit Rates

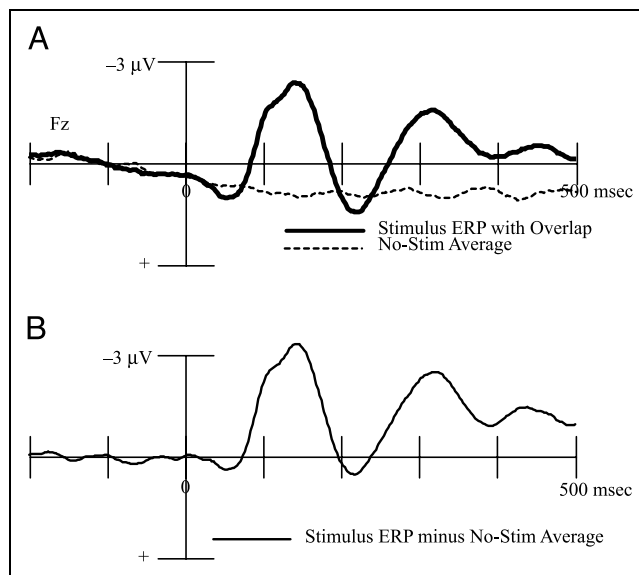
	<i>Attend Left</i>	<i>Attend Right</i>
<i>Response Times</i>		
Visual (V)	725 (155)	677 (127)
Auditory (A)	653 (147)	618 (130)
Multisensory (AV)	568 (121)	573 (109)
<i>Hit Rates</i>		
Visual (V)	72.2 (15.1)	72.0 (13.3)
Auditory (A)	82.4 (16.3)	87.2 (10.2)
Multisensory (AV)	87.3 (10.6)	89.6 (8.51)

Response times are in milliseconds (msec). Hit rates are in percentage of correctly responded targets. Standard deviation values are given in parentheses.

graphically the effectiveness of this approach, visual-only, auditory-only, and multisensory ERPs were collapsed for one attention condition (attend left), and no-stim ERPs were subtracted from this waveform (Figure 1). This figure clearly shows that nonzero baseline activity (resulting from overlapping responses) was equally present in both the ERPs evoked by stimulus trials as well as ERPs time-locked to the no-stim trials. Subtracting the no-stim ERPs from the stimulus ERPs therefore eliminated this slow overlapping baseline activity from the stimulus ERPs.

To further quantify the effectiveness of using the no-stims to subtract out previous response overlap, we tested whether multisensory and combined unisensory waveforms differed significantly from 0  $\mu\text{V}$  just before stimulus onset, at which time it is theoretically impossible to record evoked activity triggered by the current stimulus. This test was done after baselining the ERPs to the time window of  $-200$  to  $0$  msec, relative to stimulus onset. Thus, if no overlapping ERP activity was present from preceding stimulus events, one would expect that the mean voltage of all ERPs would not significantly deviate from 0  $\mu\text{V}$ , at any time point in this 200-msec baseline interval. We estimated whether this

was the case, and in addition also estimated whether the combined unisensory (A + V) and multisensory (AV) waveforms differed from each other at this time point, before and after subtracting off the no-stims. Mean voltages were calculated at six centro-medial electrodes (Cza, C1a, C2a, C1p, C2p, and Pzs) in the interval between  $-20$  and  $0$  msec relative to stimulus onset. These mean values were then submitted to an analysis of variance (ANOVA) that contained the following within-subjects factors: attention (two levels: attended or unattended), integration [also two levels: multisensory (AV) or summed unisensory (A + V) ERPs], channel (six levels: corresponding to each of the input channels), and a factor mean (which tested whether the grand mean differed significantly from zero). Notice that this test is similar in design to the one described below to assess multisensory integration effects. When no-stim trials were *not* subtracted from AV and A + V trials, this test indicated that ERP waveforms differed significantly from baseline, as indicated by a main effect of the factor mean [ $F(1,15) = 40.75$ ;  $p < .0001$ ]. In addition, we found a main effect of the factor integration [ $F(1,15) = 8.61$ ;  $p < .01$ ], which presumably derived from the significant previous-response overlap activity being included only once in the multisensory response but twice in the sum of the unisensory responses. After subtracting out the no-stim trials, however, neither the factors of mean nor of integration was significant anymore [ $F(1,15) < 1$ ;  $p > .5$  for both factors]. These results show that the no-stim trial approach effectively removes distortion caused by overlapping activity from previous-trial ERP waves and/or from any ongoing CNV/expectation waves.

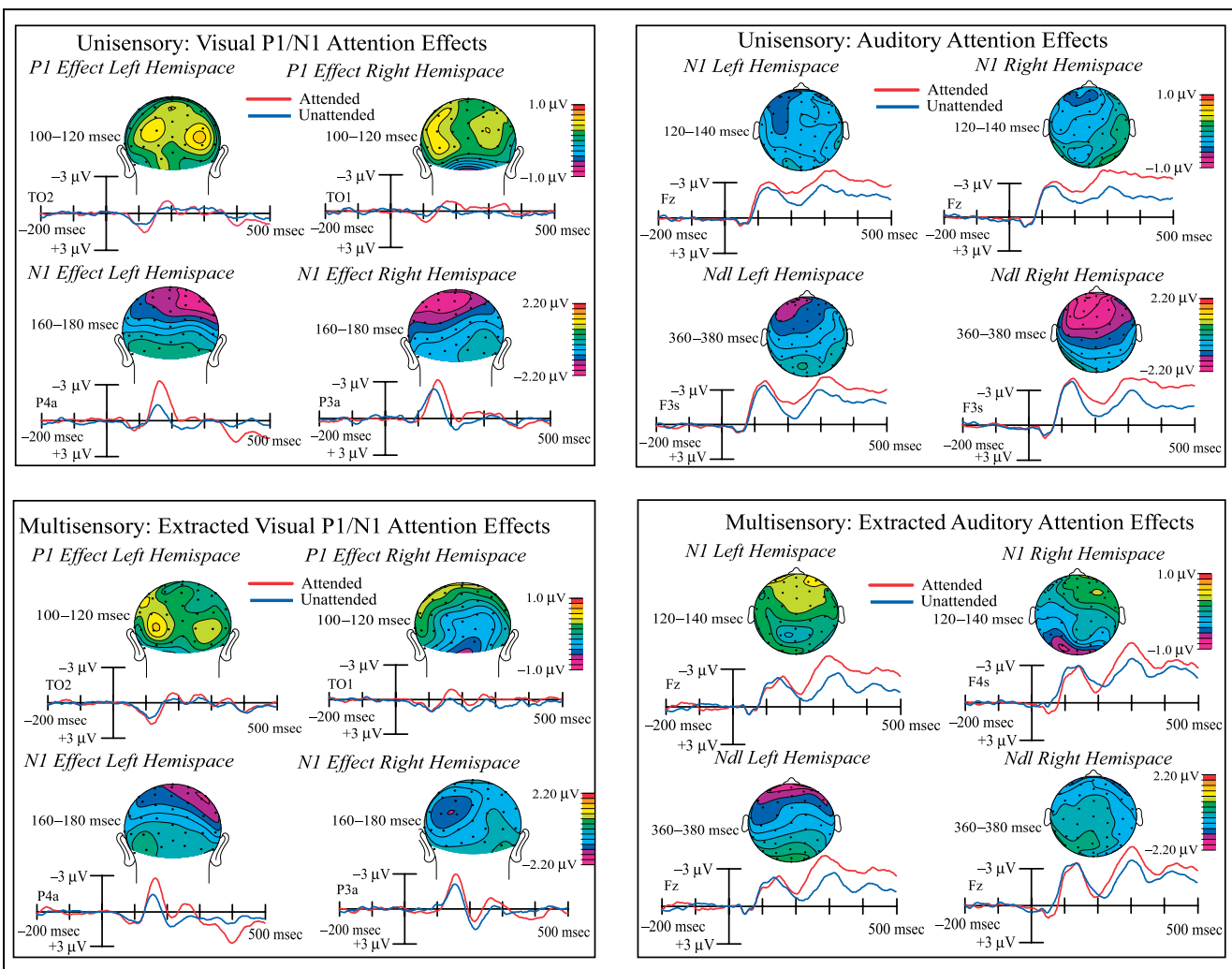


**Figure 1.** Illustration of the effectiveness of subtracting “no-stim” ERPs from ERPs elicited by real stimuli for removing the distortion from overlapping activity from previous trials or from CNV baseline shifts. To gain a high signal-to-noise ratio in this example, ERPs were collapsed across all trial types in the attend-left condition. Panel (A) shows the ERP waveforms for the real stimulus trials, with the time-locked averages for the no-stim trials superimposed. Note the low-frequency overlap activity present in both traces. Panel (B) shows the stimulus-evoked ERP waveforms after the responses to the no-stim trial averages were subtracted. Notice that the overlap-derived slanted baseline that was equally present for both the stimulus trials and the no-stim trials in panel (A), and that was resulting in distortion in the ERP epoch, has been eliminated in panel (B).

### Event-related Potentials: Unisensory Visual Stimuli

Figure 2 (top left) shows the early contralateral P1 and N1 waveforms. As expected, attended standards elicited enhanced P1 and N1 components, as compared to unattended standards. P1 amplitude was determined as the mean amplitude between 90 and 150 msec after stimulus onset at two occipital-temporal electrode sites (TO1 and TO2). These amplitudes were subjected to an ANOVA that contained the following within-subjects factors: attended location (two levels: attending left or right), stimulus location (two levels: left or right hemisphere), and hemisphere (also two levels: left or right). Attended stimuli elicited significantly larger P1 components, as was expressed by a significant interaction between attended location and stimulus location [ $F(1,15) = 7.38$ ;  $p < .05$ ]. No clear contralateral organization of the P1 component and P1 attention effect could be observed, however.

N1 amplitudes were determined by finding the mean amplitude between 160 and 200 msec after stimulus



**Figure 2.** Attention effects for unisensory and multisensory stimuli. Top left: The early attention effects on visual stimuli consisted of amplitude enhancements of posterior P1 and N1 components. Notice that N1 shows a clear contralateral enhancement, whereas the P1 is more bilaterally distributed. The scalp topographies plotted in this figure represent the difference between the responses in the attended and unattended conditions. Top right: Waveforms and scalp topographies of the early (top) and late (bottom) auditory processing negativities. Both early and late effects were characterized by frontal scalp distributions. Although the figure suggests a left hemispheric lateralization, no statistical evidence could be observed for such a lateralization. Topographies plotted in this figure represent the difference between the attended and unattended conditions. Bottom left: Attention effects on multisensory stimuli caused by visual processing and possible integration effects, obtained by subtracting auditory ERPs from multisensory ERPs (i.e., AV - A). Shown here are amplitude enhancements of posterior P1 and N1 visual ERP components. As with the visual-only stimuli, the N1 shows a clear contralateral enhancement, whereas P1 is more bilaterally distributed. Notice also that the P1 effect appears to be present mainly for stimuli presented to the left hemisphere. The scalp topographies represent the difference between the attended and unattended conditions. Bottom right: Attention effects on multisensory stimuli caused by auditory processing and possible integration effects, obtained by subtracting visual ERPs from multisensory ERPs (i.e., AV - V). Note that in contrast to auditory-only stimuli, multisensory attention effects attributed to auditory processing yielded only a somewhat later processing negativity, but not an early N1 effect. Scalp topographies represent the difference between the attended and unattended conditions.

onset on two parietal electrodes (P3a and P4a) and subjecting these amplitude measures to the same analysis as was performed on the P1 amplitudes. N1 amplitude was larger contralateral to the stimulus location [ $F(1,15) = 6.65$ ;  $p < .05$ ] and was also larger contralateral to the direction of attention [ $F(1,15) = 7.19$ ;  $p < .05$ ]. Finally, a statistically significant interaction between attention and stimulus location indicated that the contralateral attention N1 was larger for

attended stimuli than for unattended [ $F(1,15) = 23.83$ ;  $p < .001$ ].

### Event-related Potentials: Auditory Stimuli

Attended auditory stimuli elicited an early N1 enhancement, which was followed by a later frontocentrally distributed processing negativity (Ndl) that occurred from about 200 msec after stimulus onset

(see Figure 2, top right). The N1 effect was quantified by computing the mean amplitude in the interval between 90 and 150 msec after stimulus onset on two fronto-central electrodes (F3s and F4s) in each condition and for each subject. The resulting mean amplitudes were subjected to an ANOVA, which contained the following within-subjects factors: attention (two levels: attending left or right), stimulus location (two levels: left or right hemisphere), and hemisphere (also two levels: left or right). Attended stimuli yielded significantly larger N1 amplitudes, as shown by an interaction between attention and stimulus location [ $F(1,15) = 6.94; p < .02$ ]. In addition, a main effect of hemisphere showed that the N1 amplitude was slightly larger over the right hemisphere than over the left [ $F(1,15) = 12.66; p < .01$ ]. This lateralization was larger when the auditory stimuli were presented in the right hemisphere than when presented in the left hemisphere, as shown by an interaction between stimulus location and hemisphere [ $F(1,15) = 17.58; p < .001$ ].

The late Nd was quantified by computing the mean amplitudes of each condition between 200 and 400 msec after stimulus onset on the same fronto-central channels (F3s and F4s). These mean amplitudes were subjected to a similar within-subjects ANOVA as described for the N1 effect. The main finding was that attended stimuli elicited larger late processing negativities, as shown by a significant interaction between attention and stimulus location [ $F(1,15) = 29.43; p < .0001$ ]. Although Figure 2 suggests that this attention effect was somewhat lateralized to the left hemisphere, no statistical evidence for this observation could be found.

### Event-related Potentials: Multisensory Stimuli

In general, ERPs to multisensory stimuli contained a combination of visual and auditory components and effects. That is, multisensory ERPs consisted of contralateral occipital P1 and N1 components (see Figure 2, bottom left) plus a fronto-central N1 component (Figure 2, bottom right). Attention effects were reflected in an occipital N1 enhancement that was combined with a fronto-central late processing negativity. To prove that the multisensory ERPs were indeed mainly a combination of visual and auditory ERP responses, the ERPs to unisensory auditory stimuli were subtracted from the ERPs to the multisensory stimuli, separately for the attended and unattended conditions. The resulting difference wave would be expected to mostly reflect visual stimulus processing, along with some multisensory interaction effects. As can be seen in Figure 2 (bottom left), after subtracting out the unisensory auditory ERPs from the multisensory ERPs, the resulting residual activity was indeed similar to the ERPs

observed to visual-only stimuli. Analyzing this activity, a possible P1 amplitude effect, expressed in either an interaction between attended location and stimulus location or a three-way interaction between attended location, stimulus location, and hemisphere, failed to reach significance [ $F(1,15) < 1; p > .1$ ].

After subtracting off the auditory ERPs, the contralateral N1 attention effect on the multisensory responses was significant, however, and expressed in an interaction between attention location and stimulus location [ $F(1,15) = 18.19; p < .001$ ]. Although the contralaterality of the N1 component itself fell just short of significance [ $F(1,15) = 2.81; p > .1$ ], the contralaterality of the N1 attention effect was expressed as a marginally significant three-way interaction between the factors attention, stimulus location, and hemisphere [ $F(1,15) = 3.94; p < .066$ ].

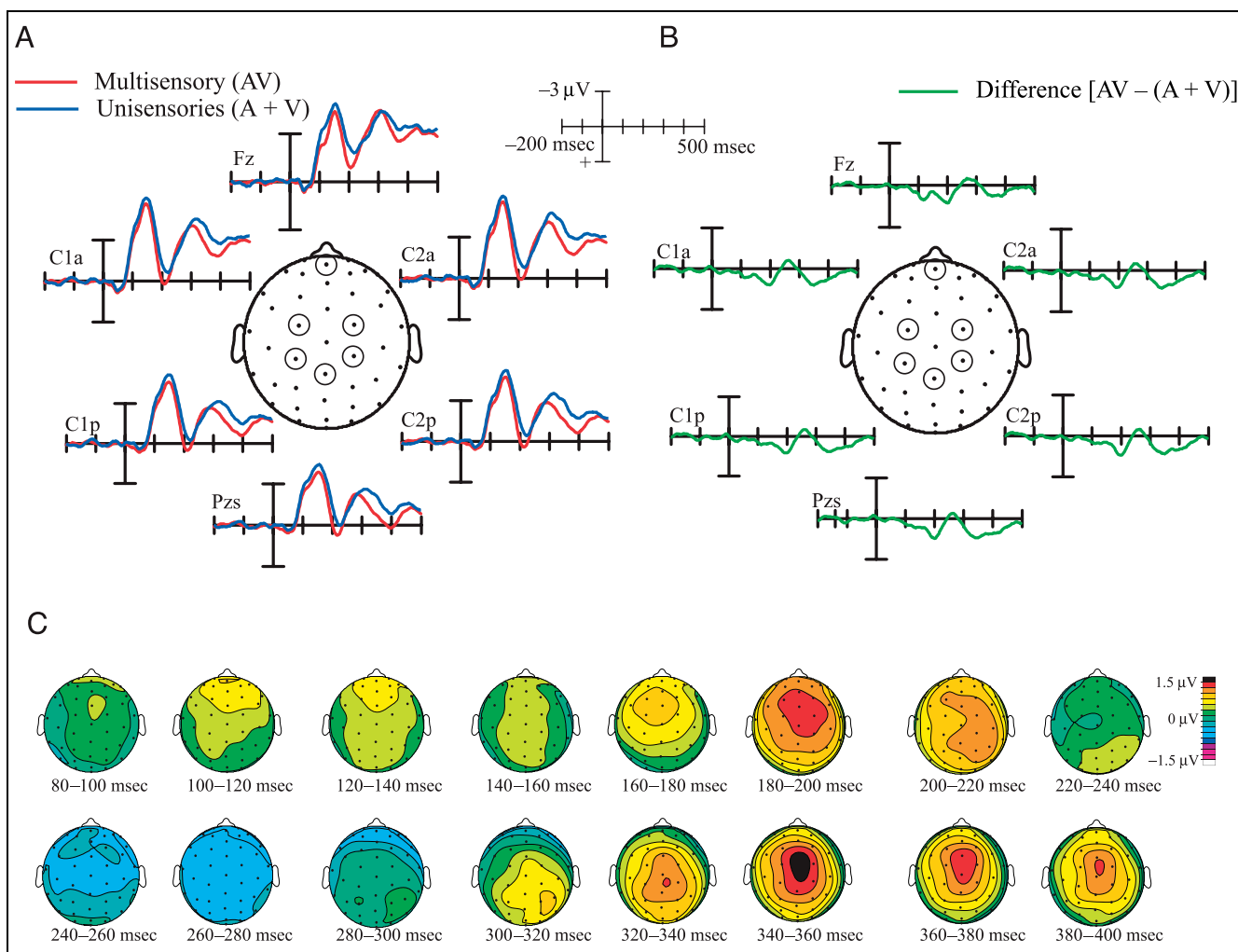
Analogously, contributions of auditory processing on attention effects of multisensory stimuli were estimated by subtracting visual-only ERPs from multisensory stimuli. Again, this was done separately for the corresponding attended and unattended locations. Interestingly, no clear direct N1 effect or early processing negativity was observed, as evidenced by the absence of an interaction between factors attention and stimulus location ( $F < 1$ ). The later attention-related frontal negativity did become significant, however, as expressed in an interaction between attention and stimulus location [ $F(1,15) = 15.27; p < .001$ ].

### Multisensory Integration

Multisensory integration processes were studied using a similar approach as the one taken by Giard and Peronet (1999), derived originally from approaches used in the animal single-unit literature (Stein & Meredith, 1993). That is, ERPs from the unisensory auditory (A) and visual (V) stimuli were summed and compared with the ERPs elicited by multisensory (AV) stimuli, according to the  $[AV - (A + V)]$  equation (after subtracting off the time-locked averages to the no-stims to correct for any summed overlap differences). Because the multisensory integration effects did not show strong lateralization effects, ERP responses to left and right hemisphere stimuli were collapsed to further improve the signal-to-noise ratio of the ERPs.

#### *Integration Effects for Attended Stimuli*

The first multisensory integration effect started at around 100 msec after stimulation and consisted of an enhanced frontal positivity for the multisensory stimulus response, relative to the sum of the unisensory responses. This effect was followed by three more phases of effects that were marked by a mostly centro-medial scalp distribution (see Figure 3). These later



**Figure 3.** Effect of multisensory integration for *attended* stimuli. (A) Overlay of the multisensory (AV) and the summed unisensory (A + V) waveforms. (B) Difference waves between the multisensory (AV) and the summed auditory + visual (A + V) ERP responses, reflecting the multisensory integration effects for attended stimuli. (C) Scalp topographies from 80 to 400 msec after stimulus onset for the integration effect difference waves plotted in panel (B).

centro-medial integration effects included an initial positivity that started at approximately 160 msec and peaked at approximately 190 msec after stimulus onset. This was followed by a centro-medial negativity that peaked at approximately 250 msec and then a late spatially broad positivity between about 350 and 450 msec after stimulus onset.

#### Integration Effects for Unattended Stimuli

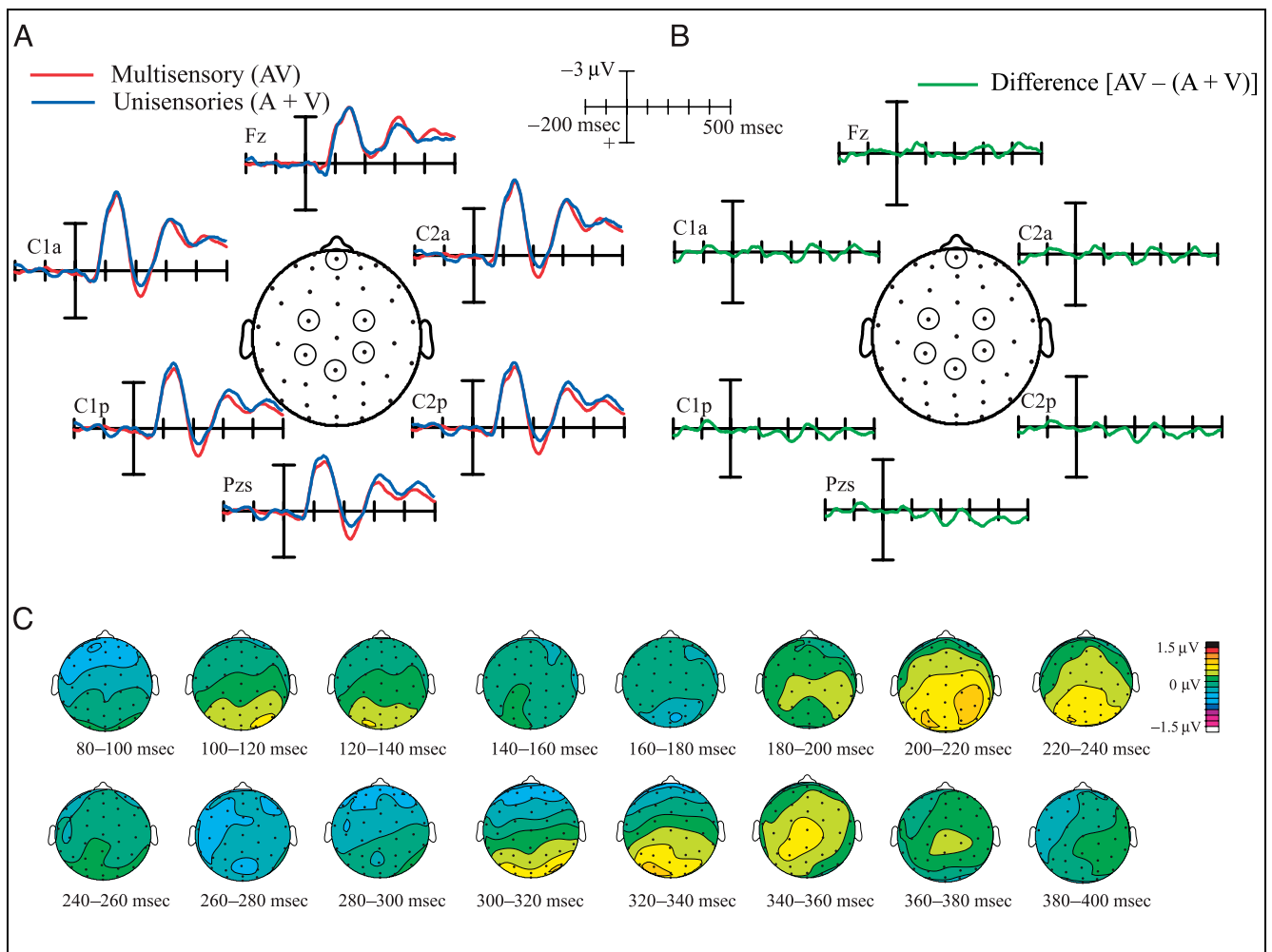
Figure 4 shows the ERPs to AV and (A + V) waveforms (A), and the difference between these (B), for the unattended stimuli. As can be seen from panel (B), the early 100-msec integration effect did not appear to be present for the unattended stimuli, and the later three phases of centro-medial effects all appeared to be substantially smaller than they were for attended stimuli (compared with Figure 3). In addition, the to-

pographies of the integration effect to unattended stimuli appear to be less centrally focused than those for attended stimuli.

### Statistical Tests and Interaction between Attention and Integration

#### Frontal Effects

The early frontal integration effect (see Figure 3) was analyzed using a similar ANOVA on six frontal electrodes (FP1m, FP2m, F3a, F4a, F3s, and F4s), using mean amplitudes between 100 and 140 msec. This analysis yielded a significant effect of the factor integration [ $F(1,15) = 5.34; p < .05$ ] as well as a significant interaction between attention and integration [ $F(1,15) = 4.35; p < .05$ ]. This finding is of particular interest, because it shows that attention effects on multisensory integration processes can occur early in time. No signif-



**Figure 4.** Effects of multisensory integration for *unattended* stimuli. (A) Overlay of the multisensory (AV) and summed unisensory (A + V) waveforms. (B) Difference waves between the multisensory (AV) and the summed auditory + visual (A + V) ERP responses, reflecting the multisensory integration effects for unattended stimuli. (C) Scalp topographies from 80 to 400 msec after stimulus onset for the integration-effect difference waves plotted in panel (B).

icant main effect of attention was observed in this test [ $F(1,15) = 1.60; p > .2$ ].

### Centro-Medial Effects

The temporally extensive effects of integration and attention over centro-medial scalp were statistically tested on the centro-medial electrodes Cza, C1a, C2a, C1p, C2p, and Pzs. Mean voltages were calculated at these electrodes at successive 20-msec intervals between stimulus onset and 500 msec after stimulus onset (see Table 2 for an overview of the centro-medial statistical results). At each latency window, these mean values were then submitted to an ANOVA that contained the following within-subjects factors: attention (two levels: attended or unattended), integration [also two levels: multisensory (AV) or summed unisensory (A + V) ERPs], and channel (six levels: corresponding to each

of the electrode sites). To correct for the increased possibility of type I errors that might have arisen from the application of multiple tests in this analysis, only results that were statistically significant in two or more consecutive time windows will be reported here.

At these centro-medial electrodes, significant main effects of the factor integration were found in three latency ranges: (1) 100–140 msec, which corresponded with an enhanced positivity elicited by multisensory stimuli (which was characterized by a more frontal maximum, but nevertheless became significant at these centro-medial electrodes; see Figure 3); (2) 180–220 msec, which corresponded to an enhanced centro-medial positivity elicited by multisensory stimuli; and (3) 320–380 msec after stimulus onset, which showed a similar topography as the effect found in the 180–220 msec time window. In the latency range of the centro-medial negativity (~250 msec), the main effect of integration was not significant.



**Table 2.** Statistical Results (Centro-Medial Sites)

Time Window (msec)	Int		Att		Int by Att	
	F	p	F	p	F	p
000–020	0.39	.5438	0.40	.5379	0.04	.8524
020–040	0.22	.6445	1.02	.3276	1.18	.2935
040–060	0.03	.8545	0.04	.8526	0.29	.5976
060–080	0.16	.6992	0.17	.6873	5.94	.0277*
080–100	1.42	.2516	0.11	.7476	4.12	.0605
100–120	12.07	.0034*	0.40	.5353	1.12	.3075
120–140	13.28	.0024*	2.57	.1298	1.03	.3261
140–160	2.26	.1535	30.86	.0001*	1.16	.2990
160–180	3.00	.1038	36.77	.0000*	5.33	.0356*
180–200	11.93	.0035*	24.27	.0002*	11.08	.0046*
200–220	14.23	.0018*	17.70	.0008*	2.01	.1771
220–240	3.38	.0860	24.46	.0002*	0.31	.5835
240–260	0.17	.6823	29.79	.0001*	1.36	.2617
260–280	2.47	.1372	19.33	.0005*	0.18	.6780
280–300	0.19	.6696	6.44	.0228*	0.18	.6777
300–320	2.29	.1509	3.09	.0993	2.93	.1073
320–340	6.21	.0249*	1.71	.2109	6.86	.0194*
340–360	16.62	.0010*	1.74	.2072	5.99	.0272*
360–380	9.26	.0082*	1.70	.2121	6.98	.0185*
380–400	2.66	.1240	1.55	.2328	6.76	.0201*
400–420	1.05	.3208	1.07	.3168	7.46	.0155*
420–440	0.79	.3889	2.87	.1110	4.43	.0526
440–460	1.38	.2590	3.79	.0705	1.68	.2140
460–480	0.70	.4144	8.36	.0112*	0.20	.6590
480–500	0.59	.4527	10.43	.0056*	0.65	.4317

Overview of statistical test results over centro-medial scalp for the multisensory integration and attention effects for the first 500 msec of the ERP. All tests were conducted with (1,15) degrees of freedom. Int = Integration; Att = Attention.

\* = Significant.

The factor attention was significant between 140 and 300 msec after stimulus onset, being most strongly significant around 180 msec (around the time when attention also interacted with integration; see below) and around 250 msec (at which time attention did not interact with integration). In this time window attended stimuli elicited more negative ERP waves than unattended.

Also at these centro-medial sites, integration and attention interacted at two different latencies (see e.g., Figure 5): (1) 160–200 msec and (2) 320–420 msec after stimulus onset. In these two latency ranges, the multi-

sensory integration effect was significantly larger for attended than for unattended stimuli.

## DISCUSSION

### Timing and Distribution of Multisensory Integration Effects

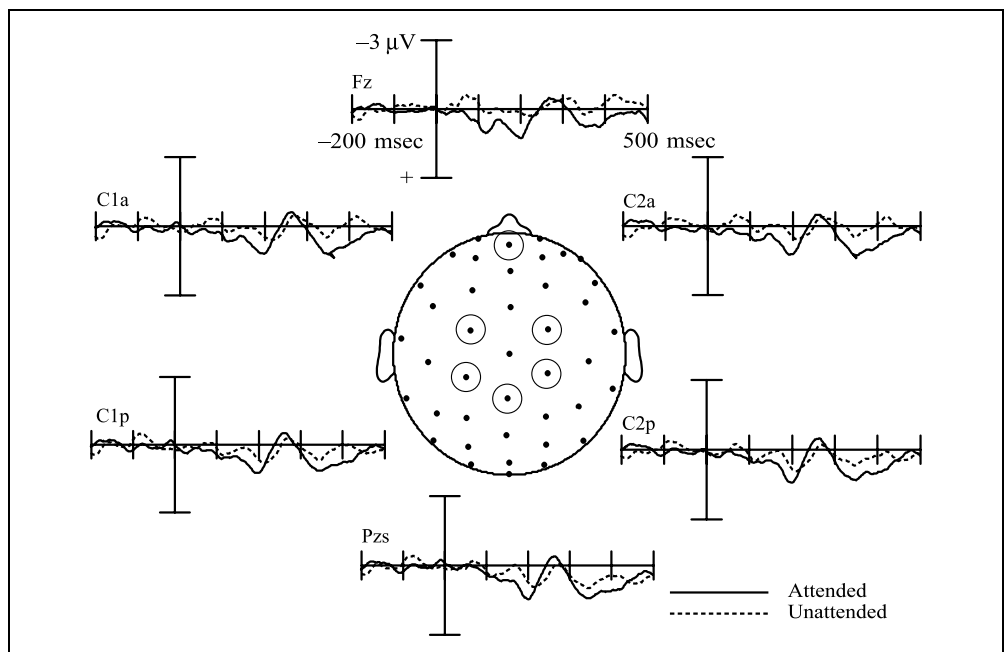
#### Timing

The present study applied the [AV – (A + V)] method used by Giard and Peronnet (1999), to investigate multisensory integration with ERPs, and previously used in many single-unit animal studies (e.g., Stein & Meredith, 1993). We extended this method by including no-stim trials in our design to correct for ERP activity resulting from overlapping ERP components, resulting from either anticipatory ERPs or from overlapping response to adjacent trials. This method proved successful in eliminating overlap resulting from adjacent stimuli, and, similar to Teder-Sälejärvi et al. (2002), we did not observe the early (i.e., 40 msec) multisensory integration effects in the ERPs reported by Giard and Peronnet after applying this correction. In the present study, the earliest multisensory integration effects were expressed in a fronto-medial effect that started around 100 msec, which was present only in the attended channel. This was followed by a centro-medial series of effects that started around 140 msec after stimulus onset, and which were enhanced by attention starting at 160 msec after stimulus onset.

Considering that after correcting for the contribution of CNV or adjacent ERP overlap the early effects were eliminated (i.e., current study, as well as Teder-Sälejärvi et al., 2002), it is possible that these very early ERP effects previously reported were actually artifactual. On the other hand, it is possible that multisensory effects occurred earlier in time even in the present study, but were producing ERP activity in striate areas that were too weak to be picked-up. The latter conclusion would be in line with the results from both Fort et al. (2002b) and Molholm et al. (2002), who successfully replicated an early integration effect, even after controlling for possible artifactual CNV contributions. However, using a somewhat different task, Fort, Delpuech, Pernier, and Giard (2002a), only found the later effects of multisensory integration, and not the early one. Thus, it is possible that multisensory integration processes might occur early in time and be reflected in ERPs under certain circumstances, but not under others.

The present study differs in some key ways from earlier studies on multisensory integration (e.g., Fort et al., 2002b; Molholm et al., 2002). These differences might, at least in part, account for the somewhat later emergence of the multisensory integration effects we report here. Molholm et al. (2002) maximized the evocation of early C1 activity by optimizing the location

**Figure 5.** Comparison of the [AV - (A + V)] difference waves for attended and unattended stimuli, showing the effects of attention on multisensory integration processes.



of the visual stimulus, in such a way that visual stimulation generated maximal striate (C1) activity. To further increase the generation of early striate activity, they also used a slow (but randomly jittered) ISI. Because we examined the influence of spatial selective attention on multisensory integrative processes, stimuli were presented in two streams at a relatively large angle from fixation. This might have led to a reduced activation of early striate activation, which would be reflected as a reduced early C1 wave. In addition, to compel subjects to selectively focus their attention on one designated side, stimuli were presented at a relatively high rate on the two lateral sites, which might have led to a further attenuation of early C1 activity.

In the present study, multisensory integration effects were characterized by an initial frontal positive scalp topography (at around 100 msec) that was followed in time by strongly focused centro-medial positivities that were present in both the 160–210 msec and 300–400 msec latency ranges. The observed scalp distributions of these effects are similar with those of the later multisensory integration effects described by previous studies (Fort et al., 2002b; Molholm et al., 2002; Teder-Sälejärvi et al., 2002; Giard & Peronnet, 1999). It is currently unclear what the neural generators of these integration effects are. The scalp topographies, associated with these effects, however, do not seem likely to arise from processes in sensory (visual and auditory) cortices.

A particularly important result in the present study is that attention interacted with multisensory integration at various points in time, including an early effect at around 100 msec. More specifically, when stimuli were attended, the multisensory integration effect was larger

than when stimuli were unattended. This finding suggests that attention can act as an early filter (i.e., by at least 100 msec after stimulus onset) that increases the binding together of auditory and visual components of a multisensory stimulus.

#### *Scalp Topographies and Brain Areas*

The earliest multisensory integration effect was marked by a frontal scalp topography (at 100 msec) that was followed by a more centrally distributed effect for the later phases of integration. There are a number of brain areas outside the sensory regions that are known to be specifically responsive to multisensory inputs, and thus, were possibly activated in the present study. For instance, single-cell recordings in animals have shown that there are polymodal representations of external space in the superior colliculus of cats (e.g., Wallace, Meredith, & Stein, 1998; Wallace & Stein, 1997) and monkeys (Wallace & Stein, 2001). These cells are not only responsive to visual and auditory stimulation alone, but are superadditive, responding much more intensely to the simultaneous presentation of visual and auditory inputs than would be expected on the basis of the sum of the visual and auditory responses alone.

Although the neuroanatomy and depth inside the brain of the superior colliculus would seem to make it unlikely to be the generators of the ERP effects here, the human superior colliculus has many reciprocal connections with the parietal cortex and other cortical regions, and it is also believed to play an important role in orienting attention toward locations in space and in coordinating spatial attention across modalities (LaBerge, 1995; Stein & Meredith, 1993; Meredith &

Stein, 1986; Goldberg & Wurtz, 1972). Recent animal studies have provided evidence for the hypothesis that superadditive responsiveness of neurons in cats' superior colliculi are mediated by interactions with cortical areas (Jiang, Wallace, Jiang, Vaughan, & Stein, 2001; Wallace et al., 1992). These results are in agreement with recent fMRI findings that have identified a number of cortical and subcortical brain areas to be involved in multisensory processing. For instance Downar, Crawley, Mikulis, and Davis (2000) suggest that the left and right temporo-parietal junctions, right middle temporal gyrus, left and right inferior gyrus, left SMA/CMA, right posterior and anterior insula are sensitive to polymodal activations. Calvert, Hansen, Iversen, and Brammers (2001) used fMRI to study audiovisual integration processes and found, in addition to the human superior colliculus, the right superior parietal lobula, right inferior parietal sulcus, insula, and several frontal areas (including the inferior frontal gyrus) as sensitive to multisensory stimulation. Future multimethodological studies may help determine whether these brain areas are involved in the multisensory integration processes reported in the present study.

### **Effects of Attention**

The results from the present study showed strong effects of attentional manipulations to visual, auditory, and multisensory objects. For visual stimuli these manipulations consisted of contralateral occipital "sensory gain" P1 and N1 enhancements that have been reported by numerous previous unisensory visual attention studies (Woldorff, Fox, et al., 1997; Wijers et al., 1996; Mangun, 1995; Mangun & Hillyard, 1988, 1990a, 1990b, 1991, 1995; Mangun, Hillyard, & Luck, 1992; Mangun, Hansen, & Hillyard, 1987; Hillyard & Münte, 1984; Eason, 1981). Attention effects on auditory stimuli consisted of an enhancement of the early fronto-central N1 component, which was followed by a late processing negativity (NdI). These results are also in line with what previously has been reported (reviewed in Woldorff, 1999; Hillyard, Mangun, Woldorff, & Luck, 1995; Näätänen, 1982, 1992). The visual P1/N1 and auditory N1 attention effects discussed above demonstrate that subjects were able to stay highly focused and were capable of ignoring stimuli presented at the unattended hemispaces. These components are generally believed to be attention-related enhancements of neural activities in the visual and auditory sensory cortices (Woldorff, Fox, et al., 1997; Heinze et al., 1994; Woldorff, Gallen, et al., 1993). Modulations of the amplitude of these components are therefore likely to reflect filtering at an early (i.e., perceptual) stage of analysis.

Attention effects on the multisensory stimuli were somewhat more complex in time course and scalp distribution, which was mainly due to their consisting of a combination of visual and auditory attention effects.

It appeared that the ERP attention effects on the multisensory stimuli emerged somewhat later in time than attention effects on the unisensory stimuli that were described above. For the visual component of the multisensory stimuli, as significant early P1 effect was not observed, with the first effect appearing at around the N1 latency. Similarly, a significant early effect on the auditory N1 was not observed for the auditory component, with the earliest significant effect being an enhancement of the late processing negativity, starting around 200 msec. A possible explanation for this relative late emergence of attention effects on multisensory stimuli may be due to the fact that multisensory stimuli are generally more easily detectable or less ambiguous than unisensory stimuli (e.g., Calvert, Campbell, et al., 2000; Stein et al., 1996). Therefore, the simultaneous presentation of an auditory and visual stimulus possibly triggers a larger initial neural response in the sensory visual and auditory cortex anyhow, regardless of whether this stimulus is attended or not, thus mitigating the earliest observed effect of attention on these multisensory stimuli.

### **Interactions of Attention and Multisensory Integration**

The most important new findings of the present study were the effects of attention on the multisensory integration processes. More specifically, we found that the activity associated with multisensory integration processes was stronger when subjects were attending to stimuli than when they were not attending (i.e., attending elsewhere). This result provides clear physiological evidence for the hypothesis that attention is capable of affecting the processes involved in the integration of visual and auditory stimulus information.

Previous studies by Bertelson and colleagues (Vroomen, Bertelson, & De Gelder, 2001a, 2001b; Bertelson et al., 2000; Bertelson, 1999) have reported, based on behavioral measures, that spatial attention does not influence the degree or direction of the ventriloquism effect, and thus, concluded that attention does not influence multisensory integration. The current study would not be consistent with this view, in that the results indicate that attention can indeed affect multisensory interaction processes at multiple phases of processing, at least under certain circumstances. In addition, other studies have indicated other ways in which attention and multisensory processing can interact. For example, Eimer and Schröger (1998), Hillyard, Simpson, Woods, Van Voorhis, and Münte (1984), and Talsma and Kok (2002) have shown that spatial attention can be supramodal in nature, specifically when the spatial positions of the visual and auditory locations are closely matched (Eimer & Schröger, 1998). This is presumably caused by a supramodal top-down spatial attentional control system that can selectively bias the

sensitivity of areas responsive to specific corresponding spatial locations in both visual and auditory cortical regions simultaneously. Under these conditions, early effects of spatial attention can be recorded on visual stimuli when audition was attended in that region of space and vice versa (Talsma & Kok, 2002). Notably, these supramodal effects of attention occurred at around the same time as when the present study found the first interactions between attention and multisensory integration. In addition, Busse and Woldorff (2003a) have recently reported that task-irrelevant auditory stimuli elicited an enhanced long-latency processing negativity when they occurred synchronously with an attended visual stimulus relative to with an unattended visual stimulus, even when these auditory stimuli were always presented from the same (unattended) location. These authors interpret these results as suggesting a bottom-up perceptual binding mechanism first linking temporally synchronous auditory and visual stimuli together, after which attention to the visual stimuli then spreads across modality and space to enhance the processing of the task-irrelevant auditory stimuli.

The abovementioned results thus suggest that either spatial or temporal proximity of auditory and visual stimuli can result in selective attention in one modality, affecting stimulus processing in another one. In the present study, audiovisual stimuli were presented both at the same location and at the same time, thus creating particularly optimal conditions for perceiving these stimuli as a coherent audiovisual object (see also Lewald & Guski, 2003). In addition, in the present study, participants were attending to both the visual and auditory modalities, thereby presumably also enhancing the conditions for such integration. Thus, top-down supramodal spatial attention that encompasses both of these modalities, along with possible contributions from bottom-up temporal binding mechanisms, could together explain the interactions between attention and multisensory integration that we observed. Future studies examining the specific roles of attending to either visual or auditory modalities of multisensory objects will aid in unraveling the relative contributions of the various mechanisms by which attention can influence multisensory integration processes.

Interestingly, the  $[AV - (A + V)]$  effects were not only larger when stimuli were attended, they also appeared to occur somewhat earlier in time than the integration effects for unattended stimuli (see Figures 3 and 4). These figures suggest that the effect at around 200 msec started somewhat earlier in the attended channel (at around 160 msec) than in the unattended one (at around 200 msec). In addition, the scalp topographies for this effect appeared to have a more central scalp distribution than the scalp topographies of the unattended integration effects, which were more posteriorly distributed. This latter result suggests that this multisensory integration effect is likely to have contributions

from additional brain areas that are activated when these stimuli are attended, but not when they are unattended. Regardless, the results show clear effects of attention on multisensory-integration brain activity.

The behavioral results also indicate that visual and auditory stimuli are effectively integrated. Because responses to auditory stimuli were faster and more accurate than responses to visual stimuli, one might argue that subjects were strategically focusing on the easier stimulus (i.e., the auditory stimulus) and trying to filter a potentially harder-to-discriminate stimulus. This explanation, however, would also predict that behavioral performance would be fastest and more accurate in the auditory condition, because filtering of the visual part of the multisensory stimulus should occur at an observable cost. In contrast to this alternative explanation, performance on these multisensory stimuli was actually the best, providing evidence that integrating the combined signals of these multisensory stimuli improved the efficiency with which these targets could be detected. One could also argue, however, that this improvement in behavioral performance results from an independent processing of auditory and visual stimulus features alone (e.g., Miller, 1982, 1986). According to this account, the facilitation in RT, accuracy, and the ERP effects when both visual and auditory targets are presented simultaneously is due to a race between redundant target signals that are independently selected by modality-specific visual and auditory selection mechanisms. Thus, the processing of the multisensory stimuli would be faster and more accurate because on trials in which noise in the auditory processing channel slows the processing of that stimulus, the processing of the visual stimulus can continue uninterrupted and finish before the auditory target is completely processed. Although the observed pattern of results is consistent with such an alternative explanation, this explanation would also predict the absence of integration effects on the  $[AV - (A + V)]$  ERP difference wave. If no integration were to take place, the amounts of noise in the auditory and visual channels would be exactly the same in the  $A + V$  ERPs as in the AV ERPs, and therefore cancel out in the difference wave. Thus, the integration effects observed in the ERP difference wave clearly show that the processing carried out when processing multisensory stimuli is more than the result of two independent mechanisms of channel selection.

### **Generic Events and the $[AV - (A + V)]$ Subtraction**

Because the  $[AV - (A + V)]$  subtraction would be sensitive to ERP responses that might be equally produced by nonsensory generic events, Giard and Peronnet (1999) limited their analysis of ERP effects to the first 200 msec after stimulus presentation, arguing that pro-

cesses such as target discrimination or decision making would be equally present for both unisensory and multisensory stimuli, which typically start at around 200 to 250 msec after stimulus onset (see, e.g., Talsma, Wijers, Klaver, & Mulder, 2001). The current study includes data from somewhat longer latencies. For a number of reasons, we argue that these effects are still unlikely to originate from processes other than multisensory interactions and integration. First, we used a task of detecting infrequent targets in a stream, with subjects required to respond to only 20% of all the stimuli at the attended side. Therefore, by focusing our analyses on the remaining 80% standard stimuli, we excluded target-related and motor-response-related ERP activity from our data. Secondly, in tasks involving detection of infrequent targets in a rapidly presented stream, target identification and processing is mainly reflected in N2b and P3 components, that are present for target stimuli, but not, or much less so, for the standard (nontarget) stimuli that were the focus of the present article (e.g., Brookhuis, Mulder, Mulder, & Gloerig, 1983); indeed we saw no evidence for any of such componentry in the nontargets of the current study. Third, the topography of the later positive phase of the [AV – (A + V)] difference wave (350–450 msec) is similar to the topography of the earlier (about 160–200 msec) positive phase, suggesting that similar brain areas are activated in both time windows. Finally, some of the longer latency effects are also observed (albeit smaller in amplitude) in the unattended channel, again suggesting that these effects are related to stimulus processing and not to generic target identification processes.

### Summary and General Conclusions

Within the context of a spatial attention task, streams of unisensory and multisensory stimuli were presented to lateral spatial locations that were either attended or unattended in different runs. By comparing the response to the multisensory stimuli to the sum of the unisensory responses, separately for when they were in the attended and unattended channels, and correcting analytically for any possible differential overlap in this contrast, we were able to investigate both the effects of multisensory integration on stimulus processing and how such effects are influenced by spatial attention. Several phases of multisensory integration effects were found. The earliest of these integration effects, which was present only for attended stimuli, consisted of an initial fronto-medial positivity at 100 msec. This effect was followed by two later phases of attentional modulation of multisensory integration processes, with those processes being larger and/or earlier in the attended channel. These consisted of a centro-medial positivity beginning at around 160 msec poststimulus and peaking at around 190 msec, followed by a positive wave peaking

between about 370 msec after stimulus onset. In addition, a negative wave peaking at about 250 msec was significantly enhanced by attention, but was not significantly affected by integration. Both the early frontal effect and the later centro-medial positive effects had scalp distributions dissimilar from those that are normally associated with visual or auditory perceptual processes. Therefore, the multisensory integration waveform is unlikely to reflect modulations in sensory processing alone, but is more likely to reflect integrative processes that originate from association areas in the brain.

A key new finding in this study is that spatial attention affected the amplitude of most of the phases of the multisensory integration effects, including the early frontal effect at 100 msec after stimulation. In addition, when stimuli were attended, the longer latency integration effects also appeared to occur somewhat earlier in time than when these stimuli were not attended. Future studies using hemodynamic imaging (e.g., functional MRI), especially in coordination with electrophysiological recordings such as used here, could help delineate the neuroanatomical areas underlying these multisensory integration processes, and their modulation by attention.

## METHODS

### Participants

Sixteen healthy volunteers participated in the experiment (age 18–27, mean 21.1; 7 men and 9 women). All participants had normal or corrected-to-normal vision and normal hearing capabilities. Participants were recruited through local advertisements at the campus of Duke University and were paid \$10 per hour. All participants gave written informed consent for their participation.

### Stimuli and Task

Streams of unisensory visual, unisensory auditory, and multisensory (auditory and visual components occurring simultaneously) were presented to the left and right hemispaces. The unisensory visual stimuli consisted of white horizontal square wave gratings ( $5.8 \times 5.8$  cm, subtending a visual angle of about  $6^\circ$ ) presented against a black background. These visual stimuli were presented unilaterally to lateral locations on either the left or right of the display at an angle of about  $15^\circ$  from a centrally presented fixation point, in the lower visual fields (about  $6^\circ$  below the horizontal meridian), each with a duration of 105 msec.

The unisensory auditory stimuli consisted of a 1600 Hz tone pip, with a total duration of 105 msec and linear rise and fall times of 10 msec, and an amplitude of 65 dB(a). These stimuli were presented through two

speakers that were placed on a table slightly lateral to and behind the monitor, such that the speakers were hidden from the subject's view and the observed location of the sound matched the location of the visual stimuli (see Eimer & Schröger, 1998). Multisensory stimuli consisted of a combination of both auditory and visual features. Presenting the visual and auditory stimuli simultaneously created the subjective impression of a single multisensory audiovisual object.

Throughout the experiment, participants were required to focus attention on one side (either left or right) and to attend to all the stimuli (auditory-only, visual-only, and multisensory) on that side. The subject's task was to detect occasional targets (20% of all stimuli) on the attended side by pressing a button on a gamepad joystick. Target stimuli were highly similar to standards, but contained a transient dip in intensity halfway through the duration of the stimulus, which caused the subjective impression that the stimulus appeared to flicker (visual target) or to stutter (auditory target). The degree of intensity reduction was determined for each subject individually during a training session, prior to the experiment (see below). Multisensory targets always contained the mid-stimulus intensity decrease in both the visual and auditory modalities.

In sum, the present study contained 12 different stimulus categories (trial types) consisting of the combination of stimulus modality (three levels: unimodal visual, unimodal auditory, or multimodal audiovisual), presentation side (two levels: left or right), and stimulus identity (two levels: targets or standards).

Preceding each session, a computer generated a new first-order counterbalanced randomized stimulus order for each subject. The stimulus onset asynchrony (SOA) varied randomly between 350 and 650 msec (mean SOA 500 msec). For each condition (attend left/attend right), 700 visual, 700 auditory, and 700 multisensory stimuli were presented across the experimental session. For each of these 700 stimuli in these categories, 350 were presented on the left and 350 on the right. Of these 350 stimuli, 280 stimuli were standards and the remaining 70 stimuli in each category were targets (20%).

Because of the high stimulation rates, the ERPs to successive stimuli could overlap in time, thereby distorting the ERP averages (Woldorff, 1993). This could include anticipatory CNV activity preceding each stimulus. This problem could be particularly important in the present experiment because in the contrast of multisensory response versus sum of unisensory responses [ $AV - (A + V)$ ], there would be double the amount of such overlap in the unisensory sum than in the multisensory response. In order to address this issue in the present study, a total of 350 "no-stim" events were included per attention condition, in addition to the other stimulus types. No-stims are points in time that are randomly inserted into the stimulus stream, which have the same randomization as the regular stimuli have, but without

the physical presence of a stimulus (Buckner et al., 1998; Burock et al., 1998). Because the occurrence of no-stims is randomized in the sequence, their time-locked averages contain, on average, the same response overlap from adjacent trials as any other trial type. Because the proportion of no-stims and the jitter rate of the SOA between trial types satisfy the conditions shown by Busse and Woldorff (2003b) to be ones for which it is unlikely for the no-stims to elicit any response themselves (such as an omitted stimulus response), one can assume that selectively averaging the no-stim events will only reflect the summated response overlap from adjacent trials. Thus, a contrast between the no-stims and the average for the other trial types will subtract out the overlap, revealing the corrected ERP waveform for these other trial types.

## Procedure

To familiarize participants with the stimulus material, they were first given a discrimination task, in which only unimodal visual or auditory stimuli were presented in a single run. In the auditory practice run, participants were presented randomly either a standard tone or a target (50% probability) and they were required to indicate whether the stimulus was a standard or a target. Based on the subject's accuracy, the difficulty of the presented target was changed by increasing or decreasing the level of the mid-stimulus intensity decrement (see Task and Stimulus section). If a subject's accuracy was below 90% correct, target difficulty was decreased by increasing the mid-duration intensity decrement (i.e., making the decrement larger and more discriminable), and when subject's accuracy was above 90% correct, target difficulty was increased by decreasing the mid-duration intensity decrement, thus making standards and targets more similar. For the visual stimuli, a similar procedure was used. The required accuracy was set to 90% correct in these unimodal practice sessions, so that subjects would still be able to achieve reasonable levels of accuracy during the main experimental session, where subjects were required to divide their attention between visual and auditory stimuli and detect targets (20% probability) in both modalities.

After the practice session was completed, the electrocaps for recording brain electrical activity were put in place on the head and participants were seated and given a description of the task, along with a number of practice blocks. Participants continued training until the experimenter was convinced that the participants understood and could perform the task. Prior to the experiment, participants were given task-specific instructions (see above) and, in addition, they were instructed to try to respond as fast and accurately as possible. To avoid movement artifacts, participants were further instructed to try to minimize blinking and making bodily movements and to fixate onto a centrally presented

fixation dot. Prior to each run, participants were instructed which side to attend to, and after the run was completed participants were given feedback about their performance. Participants were allowed to take short breaks between runs.

## Apparatus

Stimulus presentation was controlled by a personal computer running "Presentation" software (Neurobehavioral Systems, Albany, CA). EEG was recorded from 64 tin electrodes, mounted in an elastic cap (Electro-Cap International, Eaton, OH) and referenced to the right mastoid during recording. Electrode impedances were kept below 2 k $\Omega$  for the mastoids and ground, 10 k $\Omega$  for the eye electrodes, and 5 k $\Omega$  for the remaining electrodes. Horizontal eye movements were monitored by two electrodes at the outer canthi of the eyes. Vertical eye movements and eye blinks were detected by electrodes placed below the orbital ridge of both eyes, which were referenced to two electrodes directly located above the eyes. During recording, eye movements were also monitored using a closed-circuit video monitoring system. EEG was recorded using a Neuroscan (SynAmps) acquisition system (Neuroscan, Inc., El Paso, TX). The EEG channels were recorded with a band-pass filter of 0.01–100 Hz and a gain setting of 1000. Raw signals were continuously digitized with a sampling rate of 500 Hz and digitally stored for off-line analysis. Recordings took place in a sound-attenuated, dimly lit, electrically shielded chamber.

## Data Analysis

### Behavioral Data

Reaction times (RTs) for correct detections of targets, hit rates (HR), and false alarm (FA) rates were computed separately for the different conditions. These measures were subjected to an ANOVA with the following within-subject factors: stimulus type (three levels: visual-only, auditory-only, or multisensory), and location (two levels: left or right hemispace). To assess which levels of the factor stimulus type differed from each other, planned comparisons were run contrasting visual-only with auditory-only, visual-only with multisensory, and auditory-only with multisensory.

### ERP Analysis

Artifact rejection was performed off-line by discarding epochs of the EEG that were contaminated by eye movements, eye blinks, excessive muscle-related potentials, drifts, or amplifier blocking. Averages were calculated for the different stimulus types from 1000 msec before to 1200 msec after stimulus onset. The averages were digitally filtered with a noncausal, zero-phase

running average filter of 9 points, which strongly reduces signal frequencies at and above 56 Hz at our sample frequency of 500 Hz. After averaging, all channels were re-referenced to the algebraic average of the two mastoid electrodes. ERP responses were extracted by selective time-locked averaging to the different stimulus types. The various contrasts, subtractions, and analyses of the ERP averages are described in the appropriate places in the Results section.

## Acknowledgments

We thank Sebastian Bindschuh, Laura Busse, Tracy Doty, Tineke Grent-t Jong, Chad Hazlett, and Daniel Senkowski, for their assistance during various stages of this project. In addition, we also thank three anonymous reviewers for helpful comments on an earlier version of this article. This study was supported by NIH grants R01 MH 64015 and P01 NS41328 (Proj. 2) to M. G. W.

Reprint requests should be sent to Marty G. Woldorff, Duke University, Box 90999, LSRC-Bldg., Room D203, Durham, NC 27708-0999, or via e-mail: woldorff@duke.edu.

## REFERENCES

- Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Acherlesleben, T. Bachman, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events*. Amsterdam: Elsevier.
- Bertelson, P., Vroomen, J., De Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception and Psychophysics*, *62*, 321–332.
- Brookhuis, K. A., Mulder, G., Mulder, L. J. M., & Gloerig, A. B. M. (1983). The P3 complex as an index of information processing: The effects of response probability. *Biological Psychology*, *17*, 277–296.
- Buckner, R. L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., Rosen, B., & Dale, A. M. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron*, *20*, 285–296.
- Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *NeuroReport*, *9*, 3735–3739.
- Busse, L., & Woldorff, M. G. (2003a). Attention-related brain mechanisms underlying the auditory-visual perceptual grouping in the ventriloquism effect. *Cognitive Neurosciences Abstracts*, D6, 123.
- Busse, L., & Woldorff, M. G. (2003b). The ERP omitted stimulus response to "no-stim" events and its implications for fast-rate fMRI. *Neuroimage*, *18*, 856–864.
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*, 649–657.
- Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammers, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage*, *14*, 427–438.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of

- changes in the sensory environment. *Nature Neuroscience*, *3*, 277–283.
- Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381*, 66–68.
- Eason, R. G. (1981). Visual evoked potential correlates of early neural filtering during selective attention. *Bulletin of the Psychonomic Society*, *18*, 203–206.
- Eimer, M., Cockburn, D., Smedley, B., & Driver, J. (2001). Cross-modal links in endogenous spatial attention are mediated by common external locations: Evidence from event-related brain potentials. *Experimental Brain Research*, *139*, 398–411.
- Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioural Reviews*, *25*, 497–511.
- Eimer, M., & Schröger, E. (1998). Effects of intermodal attention and cross-modal attention in spatial attention. *Psychophysiology*, *35*, 313–327.
- Eimer, M., van Velzen, J., & Driver, J. (2002). Cross-modal interactions between audition, touch, and vision in endogenous spatial attention: ERP evidence on preparatory states and sensory modulations. *Journal of Cognitive Neuroscience*, *14*, 254–271.
- Fort, A., Delpuech, C., Pernier, J., & Giard, M.-H. (2002a). Early auditory–visual interactions in human cortex during nonredundant target identification. *Cognitive Brain Research*, *15*, 20–30.
- Fort, A., Delpuech, C., Pernier, J., & Giard, M.-H. (2002b). Dynamics of cortico-subcortical operations involved in audio-visual object detection in humans. *Cerebral Cortex*, *12*, 1031–1039.
- Foxe, J. J., & Schroeder, C. E. (2002). Preface. *Cognitive Brain Research*, *14*, 1.
- Giard, M.-H., & Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473–490.
- Goldberg, M. E., & Wurtz, R. H. (1972). Activity of superior colliculus in behaving monkey: II. Effects of attention in neural responses. *Journal of Neurophysiology*, *35*, 560–574.
- Heinze, H.-J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Münte, T. F., Gös, A., Scherg, M., Johannes, S., Hundeshagen, H., Gazzaniga, M. S., & Hillyard, S. A. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature*, *372*, 543–546.
- Hillyard, S. A., Mangun, G. R., Woldorff, M. G., & Luck, S. J. (1995). Neural systems mediating selective attention. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 661–681). Cambridge: MIT Press.
- Hillyard, S. A., & Münte, T. F. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Perception and Psychophysics*, *36*, 185–198.
- Hillyard, S. A., Simpson, G. V., Woods, D. L., Van Voorhis, S., & Münte, T. F. (1984). Event-related brain potentials and selective attention to different modalities. In F. Renoso-Suarez & C. Ajmone-Marsan (Eds.), *Cortical integration* (pp. 395–414). New York: Raven Press.
- Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., & Stein, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *Journal of Neurophysiology*, *85*, 506–522.
- LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge: Harvard University Press.
- Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research*, *16*, 468–478.
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*, 4–18.
- Mangun, G. R., Hansen, J. C., & Hillyard, S. A. (1987). The spatial orienting of attention: Sensory facilitation or response bias? In R. Johnson Jr., W. Rohrbaugh, & R. Parasuraman (Eds.), *Current trends in event-related potential research, electroencephalography and clinical neurophysiology* (pp. 118–124). Amsterdam: Elsevier.
- Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual attention: Behavioral and electrophysiological evidence. *Electroencephalography and clinical neurophysiology*, *70*, 417–428.
- Mangun, G. R., & Hillyard, S. A. (1990a). Allocation of visual attention to spatial locations: Tradeoff functions for event-related brain potentials and detection performance. *Perception and Psychophysics*, *47*, 532–550.
- Mangun, G. R., & Hillyard, S. A. (1990b). Electrophysiological studies of visual selective attention in humans. In A. B. Schleibel & A. Wechsler (Eds.), *The neurobiological foundations of higher cognitive function* (pp. 271–295). New York: Guilford Press.
- Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual–spatial priming. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 1057–1074.
- Mangun, G. R., & Hillyard, S. A. (1995). Mechanisms and models of selective attention. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 40–85). Oxford: Oxford University Press.
- Mangun, G. R., Hillyard, S. A., & Luck, S. J. (1992). Electrooculographic substrates of visual selective attention. In D. A. Meyer & S. Kornblum (Eds.), *Attention and performance: XIV. Synergies in experimental psychology, artificial intelligence and cognitive neuroscience* (pp. 219–243). Cambridge: Birkhäuser.
- McDonald, J. J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*, 906–908.
- McDonald, J. J., & Ward, L. M. (1999). Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1234–1252.
- McDonald, J. J., & Ward, L. M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science*, *11*, 167–171.
- McGurk, H., & MacDonald, J. W. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.
- 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.
- Miller, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279.
- Miller, J. O. (1986). Time course of coactivation in bimodal divided attention. *Perception and Psychophysics*, *40*, 331–343.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans:



- A high-density electrical mapping study. *Cognitive Brain Research*, *14*, 115–128.
- Näätänen, R. (1982). Processing negativity: An evoked potential reflection of selective attention. *Psychological Bulletin*, *92*, 605–640.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception and Psychophysiology*, *6*, 203–205.
- Platt, B. B., & Warren, D. H. (1972). Auditory localization: The importance of eye movements and a textured visual environment. *Perception and Psychophysiology*, *12*, 245–248.
- Saldaña, H. M., & Rosenblum, L. D. (1993). Visual influence on auditory pluck and bow judgments. *Perception and Psychophysiology*, *54*, 406–416.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788.
- Spence, C. J., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1005–1030.
- Spence, C. J., & Driver, J. (1997). Audiovisual links in exogenous covert spatial attention. *Perception and Psychophysiology*, *59*, 1–22.
- Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience*, *8*, 497–506.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge: MIT Press.
- Talsma, D., & Kok, A. (2001). Non-spatial intermodal selective attention is mediated by sensory brain areas: Evidence from event related potentials. *Psychophysiology*, *38*, 736–751.
- Talsma, D., & Kok, A. (2002). Intermodal spatial attention differs between vision and audition: An event-related potential analysis. *Psychophysiology*, *39*, 689–706.
- Talsma, D., Wijers, A. A., Klaver, P., & Mulder, G. (2001). Working memory processes show different degrees of lateralization: Evidence from event related potentials. *Psychophysiology*, *38*, 425–438.
- Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, *14*, 106–114.
- Vroomen, J., Bertelson, P., & De Gelder, B. (2001a). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychologica*, *108*, 21–33.
- Vroomen, J., Bertelson, P., & De Gelder, B. (2001b). The ventriloquist does not depend on the direction of automatic visual attention. *Perception and Psychophysiology*, *63*, 651–659.
- Wallace, M. T., Meredith, M. A., & Stein, B. E. (1992). Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research*, *91*, 484–488.
- Wallace, M. T., Meredith, M. A., & Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology*, *80*, 1006–1010.
- Wallace, M. T., & Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *Journal of Neuroscience*, *17*, 2429–2444.
- Wallace, M. T., & Stein, B. E. (2001). Sensory and multisensory responses in the newborn monkey superior colliculus. *Journal of Neuroscience*, *21*, 8886–8894.
- Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, C. V. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384.
- Wijers, A. A., Mulder, G., Gunter, T. C., & Smid, H. G. O. M. (1996). Brain potential analysis of selective attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of perception and action* (Vol. 3, pp. 333–387). London: Academic Press.
- Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, *30*, 98–119.
- Woldorff, M. G. (1999). Auditory attention. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 50–52). Cambridge: MIT Press.
- Woldorff, M. G., Fox, P. T., Matzke, M., Lancaster, J. L., Veeraswamy, S., Zamarripa, F., Seabolt, M., Glass, T., Gao, J. H., Martin, C. C., & Jerabek, P. (1997). Retinotopic organization of early visual spatial attention effects as revealed by PET and ERPs. *Human Brain Mapping*, *5*, 280–286.
- Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences, U.S.A.*, *90*, 8722–8726.
- Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*, 170–191.