Intermodal spatial attention differs between vision and audition: An event-related potential analysis

DURK TALSMA AND ALBERT KOK

Psychonomics Department, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Subjects were required to attend to a combination of stimulus modality (vision or audition) and location (left or right). Intermodal attention was measured by comparing event-related potentials (ERPs) to visual and auditory stimuli when the modality was relevant or irrelevant, while intramodal (spatial) attention was measured by comparing ERPs to visual and auditory stimuli presented at relevant and irrelevant spatial locations. Intramodal spatial attention was expressed differently in visual and auditory ERPs. When vision was relevant, spatial attention showed a contralateral enhancement of posterior N1 and P2 components and enhancement of parietal P3. When audition was relevant, spatial attention showed a biphasic fronto-central negativity, starting after around 100 ms. The same effects were also present in ERPs to stimuli that were presented in the irrelevant modality. Thus, spatial attention was not completely modality specific. Intermodal attention effects were also expressed differently in vision and audition. Taken together, the obtained ERP patterns of the present study show that stimulus attributes such as modality and location are processed differently in vision and audition.

Descriptors: ERPs, Intermodal spatial attention, Cross-modal links

Research has indicated that focusing attention to a given location in space can be performed with relative ease (Hillyard & Anllo-Vento, 1998; LaBerge, 1995; Treisman & Gelade, 1980; Van der Heijden, 1992, 1993). Although it appears from these studies that selecting a stimulus based upon a spatial location is easier than a selection based upon nonspatial stimulus features such as color, shape, or intensity, most of these results are obtained from studies presenting stimuli to the visual modality only. Recently more and more studies have focused on attentional processes involved in selecting auditory stimuli (Alho, Töttölä, Reinikainen, Sams, & Näätänen, 1987; Alho, Woods, & Algazi, 1994; Benedict et al., 1998; McDonald & Ward, 1999; Michie, Bearpark, Crawford, & Glue, 1990; Schröger, 1994, 1996; Schröger & Eimer, 1993; Teder-Sälejärvi, Hillyard, Röder, & Neville, 1999; Woldorff, Hackley, & Hillyard, 1993), whereas other studies have addressed the attentional systems involved in selecting stimuli between modalities (Alho et al., 1994; Alho, Woods, Algazi, & Näätänen, 1992; Buchtel & Butter, 1988; De Ruiter, Kok, & Van der Schoot, 1998; Driver & Spence, 2000; Eimer & Driver, 2001; Eimer & Schröger,

1998; Spence & Driver, 1996, 1997a, 1997b; Talsma & Kok, 2001; Teder-Sälejärvi, Münte, Sperlich, & Hillyard, 1999; Woods, Alho, & Algazi, 1992).

Despite the recent advances made in revealing the attentional mechanisms involved in selecting a spatial location, a number of questions remain open. It has been proposed, for example, that attention can be conceived as a set of filters that are selecting information at various levels of detail (see Hansen & Hillyard, 1983; Heslenfeld, 1998). According to this view, the first of these filters selects a relevant domain, which can either be the input, output, or internal processing domain (LaBerge, 1995). Further filters are supposed to select a relevant modality (such as visual, auditory, or tactile), followed by a stimulus dimension (color, brightness), and finally a stimulus feature (e.g., the color purple). However, according to the above-mentioned filter model of attention, selection of a relevant modality will occur before any selection will take place at the level of stimulus dimension or stimulus feature. According to this view, stimuli presented to a modality that is not attended would not be processed after modality selection has taken place. It has been demonstrated, however, that stimuli presented at relevant and irrelevant locations, but to an unattended modality, yield significantly different physiological responses that should be attributed to an attentional selection mechanism operating on stimuli presented to the irrelevant modality (Eimer & Schröger, 1998).

These results raise a number of interesting questions. First, an important issue is whether or not stimuli presented to an unattended modality are processed beyond the point of initial selection. Second, the question remains whether modality specific

Durk Talsma is now at the Center for Cognitive Neurosciences, Duke University.

We wish to thank Jasper Hage and Henric Brouwer for help in collecting the data, Bert Molenkamp, Alfons Hogervorst, and Marcus Spaan for technical assistance, and Heleen Slagter, Tineke 't Jong, Laura Busse, as well as three anonymous reviewers for helpful comments on an earlier version of this article.

Address requests for reprints to: Durk Talsma, Center for Cognitive Neuroscience, Duke University, LSRC Bldg., Box 90999, Durham, NC 27708-0999, USA. E-mail durk@duke.edu.

(sensory) brain areas perform intermodal attention or whether higher-order association areas are involved (Eimer & Driver, 2001). Third is the question of whether the presentation of stimuli to a nonattended modality has an influence on the processing of stimuli to the attended modality. Recent studies have shown, for instance, 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; Vroomen & de Gelder, 2000). Similarly, visual stimuli are known to affect the perceived location of an auditory stimulus (e.g., Pick, Warren, & Hay, 1969; Platt & Warren, 1972). More recently, McDonald and colleagues (McDonald, Teder-Sälejärvi, & Hillyard, 2000; McDonald & Ward, 1999, 2000) have also shown that irrelevant sounds at a given location can enhance the perception of visual stimuli at that same location.

A possible way of studying these questions is to obtain event-related potentials (ERPs) to physically identical stimuli, while attention is selectively focused on specific combination of modality (visual or auditory; hereafter referred to as intermodal attention) and location (left or right; hereafter referred to as intramodal attention). Using this approach, a number of effects on ERP components are typically found. First we will briefly discuss the intramodal effects and then the intermodal effects.

For intramodal (spatial) attention to visual stimuli, these effects comprise a modulation of the early occipital P1 and N1 waveforms for relevant stimuli, contralateral to the location of the stimulus, an effect that is known as "Sensory Gain" (Eason, 1981; Eason, Harter, & White, 1969; Eason, Oakley, & Flowers, 1983; Harter, Aine, & Schroeder, 1982; Hillyard & Münte, 1984; Hillyard, Simpson, Woods, Van Voorhis, & Münte, 1984; Luck & Hillyard, 1994; Mangun & Hillyard, 1988; Neville & Lawson, 1987; Rugg, Milner, Lines, & Phalp, 1987).

The effects of spatial attention following early P1 and N1 modulations are less consistent, however. Some studies have reported modulations of occipital P2 and N2 waves, and sometimes the P3 component (Eason et al., 1969, 1983; Hillyard & Münte, 1984; Hillyard et al., 1984; Mangun & Hillyard, 1988). Other studies have reported slow negative potentials that were recorded over occipital sites (Harter et al., 1982; Hillyard et al., 1984; Rugg et al., 1987), sites located more anteriorly to the occipital ones (Mangun & Hillyard, 1988), or even anterior sites (Neville & Lawson, 1987).

For intramodal attention to auditory stimuli, the most pronounced effects can be found in the N1 latency range. Hillyard et al. (1984) have found that the N1 component was enlarged for attended stimuli. Later studies have shown that the effects of auditory attention comprise a broad negative slow wave, labeled "Processing Negativity," (Näätänen, 1982, 1990), that can partially overlap the N1, but also extend to a later latency range. It has been found that latencies tend to increase with decreasing stimulus discriminability. Unlike visual stimuli, however, no solid distinctions between ERPs for spatial and nonspatial forms of attention have been found for auditory stimuli. In addition, ERPs to auditory stimuli do not generally show a strong contralateral organization, as opposed to the organization of (early) visual ERPs. Although brain activity to monaural stimuli tends to show such a contralateral organization, this is clearly not the case for binaural stimuli, such as those used in the present study (Woldorff et al., 1999).

For intermodal attention Woods et al. (1992) found evidence in favor of the hypothesis that sustaining this type of selection depends primarily on processing in modality-specific brain areas. The main results of these experiments comprised modulations of

the central N1 and P2 components for auditory stimuli when attention was directed to the auditory modality (as compared to auditory ERPs obtained in a condition where attention was directed to the visual modality). For visual stimuli, a negative shift over the contralateral occipital regions was found in the latency range of the N1 and P2 components in conditions where vision was relevant, as compared to conditions where audition was relevant. It should be noted, however, that Woods et al. used a headphone set to deliver their auditory stimuli, whereas a computer screen was used to present the visual stimuli.

Eimer and Schröger (1998) have argued that this procedure introduces a difference in the angle at which visual and auditory stimuli are presented, creating a possible confound, because differences between the processing of visual and auditory stimuli might be traced back to angular differences between visual and auditory stimuli. After Eimer and Schröger (1998) started presenting both visual and auditory stimuli from corresponding locations, they found that auditory attention can influence the processing of visual stimuli and that visual attention can similarly influence auditory attention. The latter result, however, was found only after the angle between the center of fixation and the stimuli was broadened to a sufficiently broad angle and matched for both visual and auditory stimuli (Eimer & Schröger, 1998, Experiment 2). Hillyard et al. (1984) also reported similar results. For auditory stimuli, a broad negativity (Nd) was found for stimuli appearing at attended locations. Moreover, these effects were present both in conditions where audition was the relevant modality and in conditions where vision was relevant. Although effects of auditory attention were observed in the attend visual conditions, it should be noted that the effects in the attend auditory conditions were larger than those in the attend visual condition. A similar pattern of results was found for visual stimuli, in this case consisting of an enhancement of the anterior N170 for stimuli presented at the attended location. Again, this result was found in both attend visual and attend auditory conditions, with effects in attend visual conditions being larger than those in the attend auditory conditions.

The main results of the experiments reported by Eimer and Schröger (1998) comprise significant negative deflections in a difference wave that was obtained when responses to stimuli presented at attended locations were subtracted from responses to stimuli at irrelevant locations. These deflections were found between about 200 and 400 ms and occurred when the stimuli were presented to the attended as well as the unattended modalities. For the visual as well as the auditory stimuli, the differences between relevant and irrelevant locations were present at Fz, Cz, and Pz, even though the exogenous ERPs were the largest at posterior sites for visual stimuli and at fronto-central leads for the auditory stimuli.

In both the studies of Woods et al. (1992) and Eimer and Schröger (1998), visual and auditory ERPs were characterized by distinct scalp topographies. The auditory ERPs revealed mostly a fronto-central scalp distribution, whereas the visual ERPs were mostly distinct on the occipital leads. Hence, it is very well possible that although auditory attention has an influence on the processing of visual stimuli and visa versa, as shown by Eimer and Schröger, the neural processes underlying the intermodal selection process are mediated by the corresponding sensory brain areas for the visual and auditory modality.

The present study has the following aims. First, it will analyze timing differences between intra- and intermodal attention. Second, it will try to answer the question of whether or not location selection processes also take place within unattended modalities.

Third, it will describe scalp potential distributions for both intra- and intermodal attention to visual as well as auditory stimuli. Fourth, it will analyze whether or not stimuli presented to a nonattended modality distract attention from processing stimuli presented to the attended modality.

To reach these goals, we used an intermodal variant of the filter paradigm, in which ERP recordings were made while subjects perceived mixed streams of visual and auditory stimuli. Visual stimuli consisted of white squares that were flashed briefly to the left or right side of a computer display and auditory stimuli consisted of short tones that were presented from a loudspeaker attached to either the left or right side of the display. Prior to each session, participants were instructed to selectively attend to one of the four stimuli while ignoring all others. In addition, a manual response to an occasional target stimulus, which was presented with a longer duration, was required for stimuli within the attended channel. In the present study, participants were instructed to attend to all four stimuli once, whereas in the Hillyard et al. (1984) study, participants attended only to either the visual or the auditory stimuli. In addition, the present study recorded more EEG channels, thus allowing us to investigate the topographies in more detail than what could be established in previous studies.

The present study also differs from the work of Eimer and Schröger (1998) in that we instructed participants before each session which modality and location had to be attended, whereas Eimer and Schröger used a symbolic cue that provided these instructions on a trial-by-trial basis. In one condition, both visual and auditory stimuli were presented regardless of which modality was attended (from here on referred to as the “mixed modality” condition), whereas in the second condition, only stimuli of the attended modality were presented. For instance, when subjects were instructed to attend to a visual stimulus, only visual stimuli were presented. Contrasting responses to relevant modality stimuli in these two conditions allowed us to investigate the potentially distracting influence of stimuli presented to a nonattended modality.

Based on our previous studies on nonspatial intermodal attention (De Ruiter et al., 1998; Talsma & Kok, 2001), we expect that effects of intermodal attention start earlier, and remain significant longer than effects of intramodal attention, to both visual and auditory stimuli. De Ruiter et al. suggested that selecting between two modalities operates on a relatively coarse level of sensory analysis, as compared to selecting among other stimulus features such as color, orientation, and texture within, for example, the visual modality. One could therefore assume that intermodal selection occurs relatively early in the processing hierarchy. Hence, we expect that differences resulting from intermodal attention will manifest themselves on earlier ERPs than those differences resulting from intramodal attention.

Evidence with regard to the brain areas involved in intermodal attention seems currently somewhat inconclusive. Although a number of studies (Hillyard et al., 1984; Eimer & Schröger, 1998; Eimer & Driver, 2001; Spence & Driver, 1996, 1997a, 1997b) have demonstrated significant cross-modal links between visual and auditory attention, it has also been shown (Woods et al., 1992) that visual stimuli elicit a pattern of responses that has a significantly different distribution than the response pattern elicited by auditory stimuli. These observations hold for both the time courses as well as the scalp topographies elicited by visual and auditory ERPs.

A number of studies have shown that prefrontal (e.g., Fuster, 1997) and anterior cingulate (e.g., Gitelman et al., 1999) areas have a supervisory coordinating role in various processes, including attention. Moreover, it is known that these connect to sensory

areas through thalamic pathways, as well as through cortico-cortical pathways (e.g., Brunia, 1999; LaBerge, 1995). It is therefore possible that the observed ERP differences between relevant and irrelevant stimuli are a reflection of differences in neural activity in perceptual areas, but that those areas are controlled by prefrontal or cingulate areas. It is possible that the location selection mechanism operates in parallel on both the visual and auditory modalities and hence it can be expected that differences between relevant and irrelevant locations will be found for the attended as well as the unattended modalities. Because the influence of a higher-order attentional selection mechanism supposedly projects to the sensory brain areas, we therefore also expect that the effects of location selection will be found over the occipital areas for visual stimuli and over the fronto-central areas for the auditory stimuli.

The present design allowed us to investigate the potentially distracting influence of stimuli presented to the irrelevant modality by comparing the processing of relevant and irrelevant stimuli in both the single and mixed modality conditions. As of yet, not much work has been done on this field, although it should be noticed that De Ruiter et al. (1998) and Talsma and Kok (2001) used a similar approach in a nonspatial task. Neither of these studies found any significant evidence, however, for a potentially distracting role of stimuli that were presented to the irrelevant modality. Based on these findings, it is therefore expected that no significant differences will be found between the single and mixed modality conditions.

Method

Participants

Nineteen healthy participants took part in the experiment (age 18–41, mean 22; 9 men and 10 women). All participants had normal or corrected-to-normal vision and normal hearing capabilities. Participants took part in the experiment as part of a requirement for their introductory course in psychology.

Stimuli

The experiment consisted of two different task conditions, denoted as the mixed and single modality task conditions. In the mixed modality task, visual and auditory stimuli were presented in random succession, with a randomly varying interstimulus interval (ISI) between 417 and 817 ms. In this task, streams of two different visual and auditory stimuli were presented. Both visual and auditory standards stimuli were presented with a 50-ms duration. The visual stimuli consisted of white squares¹ subtending a visual angle of about 2.3°, presented randomly and with equal probability to the left or right side of a visual display, at an angle of 15° from the center of fixation. The auditory stimuli consisted of pure tones (sine waves: 65 dB(A), linear rise and fall time 10 ms, frequency 1000 Hz) and were also presented randomly and with equal probability to the left and right hemifields, using a set of loudspeakers that were mounted to the sides of the visual display. Although care was taken to minimize the angle between the location of the visual and auditory stimuli, due to technical limitations, the auditory stimuli were presented at a slightly larger angle of about 16°.

¹To enable a comparison of ERP responses to visual stimuli in attend-visual and in attend-auditory conditions, visual stimuli with a relatively low information load were used. The use of such relatively large stimuli helps counteract potential problems arising when subjects tend to defocus during the attend-auditory part of the experiment.

Visual and auditory stimuli were presented randomly with equal probability. Prior to each block, participants were instructed to attend to stimuli of the designated modality at either the left or right location, to ignore all other stimuli, and to respond by pressing a button to an infrequent target stimulus (25% of all stimuli) in the designated modality and at the designated location. Targets were defined as stimuli with a slightly longer duration. For the visual stimuli, this duration was 200 ms, and for the auditory stimuli, it was 100 ms. The difference in duration for visual and auditory targets was based on prior pilot studies and ensured approximately equal discrimination difficulty for visual and auditory targets. Note that the same durations were also used in the Talsma and Kok (2001) study. Each block of trials consisted of 560 stimuli (140 stimuli per combination of modality and location: 105 of these were presented with the standard duration and 35 were targets presented with the longer duration).

The single modality task differed from the mixed modality task in that only the visual stimuli were presented when participants had to attend to the location of the visual stimuli. Similarly, only auditory stimuli were presented when participants had to attend to the location of the auditory stimuli. In the single modality task, each block of trials consisted of 280 stimuli and the ISI varied between 417 and 1,234 ms. These ISIs were chosen to ensure that the distribution of time intervals between successive stimuli in the single modality task was almost identical to the distribution of intervals that separated consecutive stimuli of the same modality in the mixed modality task condition.

Apparatus

Visual stimuli were presented on a 21-in. computer display, located at a distance of 56 cm, directly in front of the participants' eyes. Stimuli were presented using a 100-MHz pentium personal computer, running an in-house developed application for stimulus presentation. Electroencephalographic (EEG) signals were continuously recorded using 32 electrodes mounted in an electrocap. We used the following locations: FPz, AFz, Fz, Cz, Pz, Oz, FP1, FP2, F7, F3, F4, F8, FC5, FC6, T7, C3, C4, T8, CP5, CP1, CP2, CP6, T5, P3, P4, T6, PO3, PO4, O1, and O2. These electrodes were referenced against the participants' right ear lobe. Impedance was kept below 5 k Ω . Horizontal eye movement measures were obtained by deriving an electro-oculogram (EOG) from two electrodes placed to the outer canthi of the eyes. Vertical eye movements and eye blinks were detected by deriving the EOG from two electrodes placed approximately 1 cm above and below the participant's right eye. The EEGs and EOGs were fed to a Nihon-Kohden Neurotop amplifier, bandpass filtered at 35 Hz using a time constant of 2.5 s and digitized with a sample frequency of 250 Hz. All data were stored digitally for off-line analysis.

Procedure

The order in which the tasks (single versus mixed modality) were presented was balanced randomly across participants, and also the order in which the attention conditions within each task were presented varied randomly. Furthermore, at the beginning of each block of trials, a new stimulus order was randomly generated, thus providing each participant a fresh stimulus sequence. After attachment of the electrodes, participants were given a number of practice trials, to assure they understood the paradigm and make them comfortable with the stimulus material. At the beginning of each block, participants were instructed to attend to one specific combination of modality and location (e.g., the squares presented to the

left side of the screen), and to respond only to the relevant target stimulus, which was presented with a longer duration (e.g., the squares on the left side of the screen that were presented for 200 ms). They were further instructed to respond as quickly as possible while maintaining a low error rate, to minimize blinking, to fixate on a centrally presented fixation cross, and to minimize bodily movements. Eye movements were mainly monitored using the EOG, which was checked for saccadic eye movements (see below). In addition, a closed circuit video camera system was checked by the experimenters at regular intervals to verify that the participants also maintained fixation throughout blocks of trials. To collect a sufficient number of trials, each condition was presented twice to each participant. This resulted in a total of eight blocks of trials for the mixed and eight blocks of trials for the single modality task.

ERP averaging procedure. Time-locked epochs of 1,024 ms (256 samples), including a prestimulus baseline of 100 ms, were selectively averaged. Only those epochs that contained no amplifier saturations, artifacts (50 μ V amplitude deviation per 4 ms), horizontal eye movements, or behavior errors were included. Saccades were detected by computing the first derivative of the HEOG channel and scanning this for voltage changes of at least 10 μ V/ms for two consecutive intervals or more. This procedure was capable of detecting eye movements of about 2.5°. Remaining trials were corrected for eye blinks and small vertical eye movement artifacts (cf. Kenemans, Molenaar, & Verbaten, 1991). Subsequently the averages were corrected for possible overlap between successive trials using the ADJAR level 2 method.² The ADJAR correction was used because ERPs in intermodal experiments are known to be particularly vulnerable to overlap between adjacent trials (Woldorff, 1993).

ERP responses to visual and auditory stimuli were separately averaged for two attentional levels of relevance in the single modality task and four attentional levels in the mixed modality task. For the single modality task, these levels were: attended/unattended location (denoted M+L+ and M+L−, respectively), whereas the mixed modality task comprised the following levels; attended modality/attended location (M+L+), attended modality/unattended location (M+L−), unattended modality/attended location (M−L+), and unattended modality/unattended location (M−L−). ERPs to stimuli presented to the left and right hemifields were collapsed for ipsilateral and contralateral recording sites (i.e., ERPs to a stimulus presented to the right visual field that were recorded at P3 were combined with responses recorded at P4 for stimuli presented to the left visual field).

Topographical analyses. ISO voltage maps were computed using the Brain Electrical Source Analysis (BESA) software package (Scherg & Berg, 1995). These maps were derived from difference waves that were calculated from the grand average ERPs recorded at 30 electrode sites. The difference waves were obtained by subtracting the averaged ERP waveforms associated with the following attention conditions: M+L+ minus M+L− (effect of

²Although the amount of trial overlap is not very large in the present study, during the shortest ISIs, trials partially overlap the negative-going slope of the P300, resulting in a small but consistent negative slope during baseline and the initial phase of mainly the posterior visual ERPs. A comparison of the original and ADJAR corrected ERPs showed small but consistent changes in the slope of the initial phase of the ERP, with the above mentioned slope being greatly reduced.

spatial attention within the attended modality), M–L+ minus M–L– (effect of spatial attention within the unattended modality), M+L+ minus M–L+ (effect of intermodal attention at the attended location), and finally M+L– versus M–L– (effect of intermodal attention at the unattended location). Presentation of the maps was restricted to those time windows that had yielded significant effects of either intra- or intermodal attention. Previous inspection of the data had shown that the effects of attention were always distributed symmetrically or contralaterally with respect to the field of presentation of stimuli. Therefore (to simplify presentation of the data) the maps were collapsed over left and right visual fields. This was done by rearranging the positions of the electrodes in such a way that in the graphical representations, the left side of each map (relative to the scalp midline) always represented the hemisphere ipsilateral, and the right side of the maps (relative to the scalp midline) represented the hemisphere contralateral to the field of presentation of stimuli.

Results

Behavioral Data

Average response times for the correct responses and the percentage of misses and false alarms were calculated for each condition and subjected to a MANOVA analysis, with the following within-subjects factors, modality (two levels: visual versus auditory), location (two levels: left versus right hemifields), and task (two levels: single versus mixed modality). Table 1 shows the average response times and percentages of misses and false alarms in each condition. As can be seen from these data, the observed response times are about equal in each condition. Statistical analysis confirmed this observation by the absence of significant effects on the factors modality, $F(1, 18) = 0.75$, $p > .4$, and location, $F(1, 18) = 2.38$, $p > .15$. Moreover, no interaction between modality and location was found, $F(1, 18) = 0.76$, $p > .3$, indicating that subjects performed equally fast in these conditions. However, for response times, a significant interaction effect occurred between task and modality, $F(1, 18) = 4.65$, $p < .05$, which indicates that responses in the auditory conditions were significantly slower in the mixed than in the single modality task. With regard to accuracy measures (misses and false alarms) no statistically significant differences were found, which can be taken as evidence that subjects performed about equally accurately in all conditions.

Table 1. Mean Response Times, Percentage Misses, and False Alarms

	Visual Stimuli		Auditory Stimuli	
	Attend Left	Attend Right	Attend Left	Attend Right
Response times				
Mixed	501 (67.4)	515 (66.4)	505 (83.7)	515 (74.0)
Single	504 (66.3)	516 (71.7)	491 (45.7)	515 (60.8)
False Alarms				
Mixed	1.90 (1.97)	2.50 (2.74)	3.35 (6.75)	2.52 (2.19)
Single	1.63 (1.94)	1.65 (1.62)	1.92 (2.07)	1.73 (2.58)
Misses				
Mixed	10.6 (11.0)	12.1 (10.7)	12.9 (21.5)	9.28 (5.70)
Single	10.9 (9.28)	13.2 (13.3)	14.9 (14.5)	8.65 (5.62)

Note: Response times in milliseconds. Standard deviations are given in parentheses.

Event-Related Potentials: Visual Stimuli

Mixed modality task. Figure 1 presents the average ERP waveform responses to visual stimuli for the mixed modality task. Inspection of the waveforms suggests that visual attention had the following two effects: (a) a broad negativity that affected N1 and P2 components (around 200 ms after stimulus onset) and (b) a broad positivity affecting P3 (around 400 ms after stimulus onset). The N1 component had a clear contralateral distribution at the posterior parietal scalp sites. Statistical analysis was performed using a within-subjects MANOVA. The analysis was run on consecutive samples of the individual average ERP waveforms, starting at stimulus onset and ending at 908 ms after stimulus onset, that is, the end of the averaging window. Following Guthrie and Buchwald (1991), significant effects are reported only for extended runs of eight samples or more. These analyses were run on data obtained from the P3/4 and F3/F4 electrode pairs. Prior visual inspection of the data had suggested that effects to visual stimuli were most prominent over the P3/P4 electrodes and auditory effects over F3/F4. To differentiate between intra- and intermodal forms of attention, four planned comparisons were run at each consecutive sample. The following comparisons were conducted: M+L+ versus M+L–, M–L+ versus M–L– (intramodal attention), M+L+ versus M–L+, and M+L– versus M–L– (intermodal attention).

Table 2 depicts the time ranges that were found significant in these analyses and Figure 2 presents the difference waveforms that are associated with intra- and intermodal attention. Furthermore, topographical maps associated with time slices that had yielded significant effects of intra- and intermodal attention are displayed in Figure 3.

Effects of intramodal attention. The first comparison concerned the effect of intramodal (spatial) attention when vision was the relevant modality (M+L+ vs. M+L–). At the P3 and P4 locations, significant effects were found for the three latency ranges that corresponded with epochs in which the N1 and P2 components (around 200 ms after stimulus onset) and P3 components (around 400 ms after stimulus onset) were maximally active. Spatial attention caused an increased negativity of the N1 and P2 components, as well as an increased positivity of the P3 (see Figure 1 and Figure 2, upper part). The onset of the early effect was around 180 ms after stimulus onset. Further inspection of Figure 3 shows that the negative modulation of N1 and P2 components had a contralateral posterior distribution around 200 ms after stimulus onset, whereas the P3 enhancement effect had a symmetrical parietal distribution. In addition, a later effect, between about 600 and 730 ms after stimulus onset was found, which corresponds to late slow wave activity (see Figure 1).

Spatial attention also showed an effect at the P3 and P4 locations when vision was not the relevant modality (see Figure 2, upper part). This observation was confirmed by significant effects of intramodal spatial attention for the M–L+ versus M–L– contrasts at the P3 and P4 electrode positions in the 176–236 time area. No significant effects of the same contrasts were obtained in the late (P3) region. To summarize, when vision was not the relevant modality, the statistically significant effects were earlier but smaller, and as we will see below, also more robust at frontal electrodes.

At the F3 and F4 electrode sites, a slightly different pattern of results was found. Here, spatial attention effects started somewhat later when vision was the relevant modality, and affected mainly P2, P3, and late slow wave activity (see Table 2 and Figure 1).

Visual ERPs, Mixed Modality Condition

Figure 1. Event-related potentials to visual stimuli from a selected number of channels in the mixed modality task. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side were combined in one figure. M+L+: Standard duration stimuli that were presented at the attended location, when vision was relevant; M+L-: Standard duration stimuli that were presented to the unattended location when vision was relevant; M-L+: Standard duration stimuli that were presented to the attended location when audition was relevant; M-L-: Standard duration stimuli that were presented to the unattended location when audition was relevant.

However, when vision was not the relevant modality (M-L+ vs. M-L- contrast), effects of attention actually started earlier at the F3 and F4 locations than at P3 and P4 sites (see Table 2). These effects include P1/N1 components as well as late slow wave activity (see Figure 1).

Effects of intermodal attention. At the P3 and P4 electrode positions, intermodal visual attention caused similar effects on ERP components (negative modulation of N1 and P2 and enhancement of P3, respectively) as intramodal visual attention (see Figure 2, lower part). The onset of this effect was slightly earlier then

Table 2. Overview of the Latency Ranges for the Visual Stimuli for which Significant Amplitude Differences Were Obtained in the Mixed Modality Task

Intramodal Attention		Intermodal attention	
M+L+ vs. M+L-	M-L+ vs. M-L-	M+L+ vs. M-L+	M+L- vs. M-L-
F3 and F4			
240–304 ms	52–88 ms	212–300 ms	136–244 ms
432–492 ms	120–216 ms	380–456 ms	
548–888 ms	465–552 ms	564–840 ms	
P3 and P4			
184–272 ms	176–236 ms	176–264 ms	156–220 ms
360–452 ms		360–464 ms	
604–732 ms			

Note: Effects were considered significant when statistical tests (using 1,18 degrees of freedom) showed $\alpha < 5\%$ for runs of eight or more consecutive samples. Results were obtained from F3 and F4 (top) and P3 and P4 electrodes (bottom). M+: relevant modality; M-: irrelevant modality; L+: relevant location; L-: irrelevant location.

Visual Difference Waves

Attended minus Unattended Location

Attended minus Unattended Modality

Figure 2. Difference waves following visual stimuli, obtained by subtracting L- from L+ stimuli (top) and by subtracting M- from M+ stimuli (bottom). Effects are shown for a selected number of locations. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure.

the onset of intramodal spatial attention, namely at 176 ms after stimulus onset. Statistical analysis further verified that the negative modulation of N1 and P2 occurred not only in the condition when visual stimuli were presented at the attended spatial location

(M+L+ vs. M-L+ contrast; see Table 2) but also when these stimuli were presented at the unattended spatial location (M+L- vs. M-L- contrasts, 156-220 ms after stimulus onset). Note also that the onset of the effect of modality relevance was slightly

Visual Topographies

Figure 3. Scalp topographies (top) and C.S.D. maps (bottom) for the attention effects following visual stimuli, obtained at 200 ms (early effect) and 395 ms (late effect) after stimulus onset. Contour spacing is $0.1 \mu\text{V}$ for the potential maps and $0.05 \mu\text{V}/\text{cm}^2$ for C.S.D. maps. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure, where the left side of each map represents the ipsilateral presentation side, and the right side of each map the contralateral.

earlier (around 156 ms after stimulus onset) when visual stimuli were presented at unattended than at attended locations (see also Figure 2, lower part). However, no late (P3 enhancement) effect of intermodal attention was found in the condition when visual stimuli were presented at unattended spatial locations.

At F3 and F4 electrode sites, we found a pattern of results for intermodal attention that was similar to that found at P3 and P4 locations. At attended (relevant) locations, we found that modality relevance affected the N1, P2, and P3 components (with a slightly later onset, as compared to the latency of these effects at the P3 and P4 sites), as well as late negative slow wave activity. At unattended (irrelevant) locations, intermodal attention effects were found in about the same latency range as at the P3 and P4 electrode sites, although this effect started earlier (at 136 ms after stimulus onset) at F3 and F4 and ended later (at 244 ms after stimulus onset) as compared to the effect at P3 and P4 locations.

The scalp distribution of the intermodal attention effect on N1 and P2 components was highly similar to the one that was obtained for intramodal attention (see Figure 2). Notice also that the negative modulation of the N1 and P2 components was again more

pronounced at the posterior scalp sites contralateral to the attended field of presentation of stimuli. In conclusion, the visual ERP results indicated that intramodal (spatial) and intermodal visual attention affected almost the same areas in the ERP waveforms. These effects comprised an early negative modulation of the N1 and P2 components and a late positive enhancement of parietal P3. These effects appeared to be more strongly present in the condition when visual stimuli were presented at attended locations or when vision was the relevant modality. An important additional finding was that the early negative modulation of the N1 and P2 components was also present in ERPs to visual stimuli that were presented at unattended locations and in the condition when vision was not the relevant modality.

Single versus mixed modality conditions. To answer the question to what extent auditory stimuli could have interfered with the selection of location of visual stimuli, ERPs in the single and mixed modality tasks were compared with respect to the effects of spatial attention. Inspection of the average ERP waveforms suggested that the pattern of effects of visual spatial attention in the

single modality task was similar to the effects that were obtained in the mixed modality task. This can be observed clearly in the difference waveforms that were obtained by subtracting ERPs in the M+L- from ERPs in the M+L+ condition (see Figure 4). Notice that the early negative modulation of the N1 and P2 components and the subsequent positive enlargement of the P3 did not differ appreciably between the two tasks. The ERP measures derived from these two tasks were also subjected to a separate SPSS-MANOVA time slice analysis. This analysis was conducted on the ERP waveforms to M+L+ and M+L- stimuli, derived from the single and mixed modality tasks, using ERPs from P3 and P4 as well as F3 and F4 electrode pairs. The analysis contained the following within-subjects factors; task (two levels; single versus mixed) and attention (two levels; designated M+L+ and M+L-). This test was also performed on consecutive samples, from 0 to 908 ms after stimulus onset, and again, effects are reported only when they became significant on eight or more consecutive samples to reduce the possibility of type I errors (see Guthrie & Buchwald, 1991). Notice that especially the Attention $\times$ Task interactions are of theoretical importance here because they index to what extent the effects of spatial attention on the ERP waveforms differed between the single and mixed modality conditions.

Results of the analysis (see Table 3) confirmed the prior observations by showing that visual spatial attention affected approximately the same ERP epochs in the single and mixed modality task, at P3 and P4 locations, although in the single modality task, the P3 effect between about 350 to 450 ms after stimulus onset failed to reach significance.

At the F3 and F4 recording sites, similar effects were found for the M+L+ versus M+L- contrast on the single modality task and the combined single/mixed modality task. These effects consisted of a negative shift in P2 activity as well as a late negative slow wave. Note that these effects are similar to what we found in the mixed modality condition analysis (see above). The single modality analysis as well as the combined single/mixed modality analysis also show an early N1 effect (between about 128 and 168 ms after stimulus onset), which was not significant in the mixed modality task.

Importantly, no interactions between task and attention were found to be significant. This indicates that presentation of auditory stimuli in the mixed task did not interfere with processing of the spatial location of visual stimuli as reflected in the ERP waveforms. Note that these interactions were absent at both the P3 and P4 electrode pairs and at the F3 and F4 electrode pairs.

Visual Difference Waves Single vs. Mixed Modality Tasks

Figure 4. Difference waves for visual stimuli, obtained by subtracting ERPs elicited by M+L- stimuli from ERPs elicited by M+L+ stimuli, in single and mixed modality conditions. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure. A selected number of channels are shown.

Table 3. Overview of the Time Slice Analyses for the Visual Stimuli for which Significant Amplitude Differences Were Obtained in the Single Modality Task. Also Shown Are the Interactions between the Single and Mixed Modality Tasks

Single Modality		Single vs. Mixed Modality	
Attention	Task	Attention	Task $\times$ Attention
F3 and F4			
128–168 ms	96–372 ms	116–156 ms	
224–332 ms	400–512 ms	232–324 ms	
536–908 ms	564–656 ms	536–908 ms	
P3 and P4			
180–280 ms	284–328 ms	180–284 ms	
520–872 ms	404–476 ms	368–436 ms	
		532–872 ms	

Note: Effects were considered significant when statistical tests (using 1,18 degrees of freedom) showed $\alpha < 5\%$ for eight or more consecutive samples. Results were obtained from F3 and F4 electrodes (top) and P3 and P4 electrodes (bottom). Attention: factor attention (M+L+ vs. M+L-); Task: factor task.

Event-Related Potentials: Auditory Stimuli

Mixed modality task. Figure 5 presents the ERPs to auditory stimuli in the mixed modality task. Attention effects had a fronto-central scalp distribution and were reflected in a prolonged negativity. The onset of this negativity was around 100 ms after stimulus onset, where it overlapped with the phasic N1 and P2 components. At posterior-parietal sites, attention became manifest in an enhancement of a late positive wave (P3) that peaked around 380 ms after stimulus onset. It can further be seen that the fronto-central negativity was much more pronounced in the condition when both the modality and location of auditory stimuli were relevant (M+L+) than in the other attention conditions. Attention effects were clearly reflected in two distinct negative phases peaking around 180 to 200 and 400 to 600 ms after stimulus onset (further denoted as Nde and Ndl). The late phase has a more frontal distribution than the early phase. Figure 6 further illustrates that the biphasic morphology was present in the difference waves that reflected effects of intramodal attention (relevant minus irrelevant spatial location; upper figures) as well as in the difference waves that reflected effects of intermodal attention (relevant minus irrelevant modality; lower figures). For intermodal attention, the Nde component seemed to be much more robust than for intramodal attention.

Similar to the analysis of visual ERPs, a series of SPSS-MANOVA time-slice analyses were conducted on auditory ERPs. These analyses were also based on ERPs obtained from the F3 and F4 sites, at which the attention effects were manifested most clearly as determined by visual inspection, and at P3 and P4 to compare auditory analysis to that of the visual stimuli. Lateral scalp sites were selected to verify if manipulations of auditory attention would also become manifest in differences between ERPs at the scalp sites contralateral and ipsilateral to the field of stimulation. Results of the overall MANOVA revealed significant effects of the factor attention (comprising four subconditions) over a very broad latency range (96 to 908 ms after stimulus onset) coinciding with the sustained negativity in the ERP waveforms. The absence of any Attention $\times$ Hemisphere interaction further indicated that the overall effect of auditory attention (as reflected in the Nde and Ndl components) was distributed symmetrically over the left and right fronto-central scalp sites.

Planned comparisons were run on consecutive time slices to analyze effects of intra- and intermodal attention on auditory ERPs. Table 4 presents the significant time ranges that were obtained in these analyses. Topographical maps associated with time slices that had yielded significant effects of intra- and intermodal attention are displayed in Figure 6.

Effects of intramodal attention. Inspection of the difference waves displayed in Figure 6 (upper part) indicates that the effects of spatial attention were much more pronounced when audition was the attended modality. Statistical analyses at the F3 and F4 electrode pair verified the significance of this effect (see Table 4). When audition was the relevant modality, spatial attention significantly affected ERPs in the two time areas, namely around 152 to 300 ms and 312 to 908 ms after stimulus onset (M+L+ vs. M+L- comparison). These latency ranges roughly corresponded to the early and late phases of the sustained negativity (Nde, Ndl) that are also clearly visible in Figure 5 and Figure 6 (upper part). It can further be observed from the topographical maps that the early phase had a symmetrical fronto-central negativity for stimuli presented at the relevant location, whereas the distribution of the late phase was somewhat more frontal (M+L+; see Figure 7).

Significant effects of spatial auditory attention at the F3 and F4 electrode pairs were also found in the condition when audition was not the relevant modality (M-L+ vs. M-L-). Although the attention effects were much smaller than those that were observed in the attend auditory condition, significant effects of spatial attention were obtained in two time areas, around 124 to 172 ms and 284 to 452 ms after stimulus onset.

At the P3 and P4 locations, only late effects were found significant. As can be seen from Table 4, spatial attention to auditory stimuli affected the P3 component (between approximately 320 to 420 ms after stimulus onset). Note that this effect was more pronounced when audition was relevant (M+) than when audition was not relevant (M-). Additionally, late slow wave differences of spatial auditory attention were found in the M+ conditions, but not in the M- conditions.

Effects of intermodal attention. Figure 6 (lower part) illustrates effects of intermodal attention (attended minus unattended modal-

Auditory ERPs, Mixed Modality Condition

Figure 5. Event-related potentials to auditory stimuli from a selected number of channels in the mixed modality task. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure. M+L+: Standard duration stimuli that were presented at the attended location, when audition was relevant; M+L-: Standard duration stimuli that were presented to the unattended location when audition was relevant; M-L+: Standard duration stimuli that were presented to the attended location when vision was relevant; M-L-: Standard duration stimuli that were presented to the unattended location when vision was relevant.

ity difference waves) in the conditions when auditory stimuli appeared at attended and unattended spatial locations. Effects of modality relevance were present in ERPs that were elicited by auditory stimuli at the attended spatial location. These effects became manifest in both an early phase (Nde, between 108 and 312 ms after stimulus onset) and a late phase (Ndl, between 412 and 908 ms after stimulus onset) of the difference waveforms. Notice also that modality relevance had an earlier onset for ERPs to stimuli presented in the auditory than in the visual modality (Figure 2, lower part). A substantial Nde was also elicited by auditory stimuli that appeared at the unattended spatial location. Figure 7 further reveals that effect of intermodal attention on the early phase had a symmetrical fronto-central distribution whereas the distribution of the effect on the late phase was more frontal.

At the F3 and F4 electrode locations, statistical analyses (see Table 4) showed significant effects of intermodal auditory attention for latencies that spanned the areas between 108 to 312 ms (Nde) and 412 to 908 ms after stimulus onset (Ndl), for stimuli presented at the attended locations (M+L+ vs. M-L+ contrast). Intermodal auditory attention also significantly affected the early phase (Nde) in the condition when stimuli appeared at unattended locations (M+L- vs. M-L- contrast, time area 120 to 196 ms

after stimulus onset). No effect of intermodal attention was found for the late phase (Ndl). Thus, in accordance with findings obtained for visual ERPs, the present results indicate that auditory intramodal and intermodal forms of attention became manifest in the same ERP components and were reflected in similar topographical distributions. The effect of intermodal attention had an earlier onset (around 96 ms after stimulus onset) for the auditory modality than for visual modality (around 180 ms after stimulus onset).

At the P3 and P4 electrode pair, intermodal attention to auditory stimuli yielded significant differences as well. At the relevant (attended) location, these effects consisted of increased negativities in the N1/P2 complex latency range (between 128 and 236 ms after stimulus onset), an increased P3 (between 292 and 432 ms), and finally increased negative slow wave activity (between 548 and 812 ms after stimulus onset), when audition was relevant. At the unattended (L-) location, intermodal attention only caused a significant effect between 140 and 180 ms after stimulus onset, in the N1 latency range.

Single versus mixed modality conditions. Figure 8 presents the ERP difference waves that reflected effects of auditory spatial attention (ERPs in the M+L+ minus ERPs in the M+L- condi-

Auditory Difference Waves

Attended minus Unattended Location

Attended minus Unattended Modality

Figure 6. Difference waves following auditory stimuli, obtained by subtracting L- from L+ stimuli (top) and by subtracting M- from M+ stimuli (bottom). Effects are shown for a selected number of channels. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure.

tion) superimposed for the single and mixed modality tasks. Similar effects of attention were found for ERPs elicited in the single and mixed tasks. In both tasks, spatial attention produced a sustained fronto-central negativity at the F3 and F4 electrode loca-

tions with a similar time course that started at around 110 ms after stimulus onset, which had a clear biphasic morphology. At F3 and F4, this effect remained significant throughout the entire remaining part of the interval. It can also be observed that the Nde was more

Table 4. Overview of the Latency Ranges for the Auditory Stimuli for which Significant Amplitude Differences Were Obtained in the Mixed Modality Task

Intramodal Attention		Intermodal attention	
M+L+ vs. M+L-	M-L+ vs. M-L-	M+L+ vs. M-L+	M+L- vs. M-L-
152–300 ms 312–908 ms	124–172 ms 284–452 ms	F3 and F4	
		108–312 ms 412–908 ms	120–196 ms
312–416 ms 552–840 ms	328–452 ms	P3 and P4	
		128–236 ms 292–432 ms 548–812 ms	140–180 ms

Note: Effects were considered significant when statistical tests (using 1,18 degrees of freedom) showed $\alpha < 5\%$ for runs of eight or more consecutive samples. Results were obtained from F3 and F4 (top) and P3 and P4 electrodes (bottom). M+: relevant modality; M-: irrelevant modality; L+: relevant location; L-: irrelevant location.

pronounced in the single than in the mixed task. This became evident in significant Task $\times$ Attention interactions in the statistical analyses (see Table 5) for time slices in the time area between 164 and 264 ms after stimulus onset. In addition, some later interactions occurred between task and attention between 304 and 340 ms after stimulus onset (P3) and between 752 and 796 ms after stimulus onset (late slow wave activity).

At P3 and P4 locations, effects of attention were significant in much shorter intervals, namely between about 115 and 272 ms after stimulus onset (N1 and P2), and from about 500 to 860 ms after stimulus onset (see Table 5 for a detailed overview). In addition, in the combined single/mixed modality analysis, a P3 effect can be observed (between 320 and 404 ms after stimulus onset) that cannot be found in the single modality condition alone. At the P3 and P4 locations, one Task $\times$ Attention interaction was found, between 152 and 200 after stimulus onset.

Discussion

Two attentional processes were investigated in this study, namely attending to the location of a stimulus (intramodal attention) and attending to the modality of a stimulus (intermodal attention). The relevant modality and spatial location of the stimuli were varied within subjects and specified prior to each block of trials (sustained attention).

With respect to intramodal attention, our study attempted to clarify to what extent intramodal spatial attention is based on modality-specific or supramodal mechanisms. It was predicted that if spatial attention is completely modality unspecific, effects of spatial attention should be expressed in the same way in the ERP components regardless of which stimulus modality is attended by the subjects. This would become manifest in (a) strong cross-modal effects of spatial attention, that is, ERP effects that are elicited by stimuli irrespective of whether their modality is attended or not; and (b) identical topographical distributions of ERP effects of spatial attention in vision and audition above the secondary (presumably supramodal) brain areas. A similar question was raised with respect to intermodal attention: Does attention to the modality of a stimulus involve modulation of activity in modality-specific auditory and visual brain areas, the secondary association areas, or a combination of both areas?

Intramodal Attention

The present study supports the view that sustained spatial attention operates primarily on the basis of modality-specific mechanisms. This conclusion was based on the finding that spatial attention was expressed in early ERP components with scalp distributions that are typically associated with the visual and auditory modalities. For the visual modality, this involved modulation of the posterior N1 and P2 components. For the auditory modality, the early phase (Nde) of a fronto-central sustained negativity was enlarged by attentional processes. Note that the contralateral enlarged negativity of the N1 has also been reported in earlier studies on visuo-spatial attention and is believed to be generated by neural sources in extrastriate visual cortex (Eason, 1981; Hillyard & Münte, 1984; Mangun, Hillyard, & Luck, 1992; Wijers, Lange, Mulder, & Mulder, 1997). Such early auditory Nde effects have been reported to originate from the supratemporal auditory cortex (Näätänen, 1992; Woldorff et al., 1993).

In addition, there was also a considerable consistency in the morphology of the ERP waveforms, and the direction of attentional effects at frontal and parietal locations. For example, visual stimuli evoked N1 and P2 components at the (contralateral) P3 and P4 locations similar to those at the F3 and F4 locations. In addition, the negative deflection overlapping these components, related to selection, was similar at these locations. This pattern of results suggests that the effects of visual spatial attention were caused by one set of neural generators. For auditory stimuli, the attention-related negative deflections following the N1 component were also similar at frontal and occipital sites.

One should also notice that spatial attention effects to visual as well as to auditory stimuli were present in the M- conditions, that is, when these modalities were irrelevant. Although these effects of location relevance were, in general, smaller in M- conditions than in M+ conditions, they were consistently present. In this respect, the present results are consistent with prior studies showing that spatial attention maximally facilitates processing of stimuli at the attended location along with a weaker, yet consistent facilitation of early pathways of the unattended modality (Eimer & Schröger, 1998; Hillyard et al., 1984; Teder-Sälejärvi, Münte, et al., 1999).

At first sight, this finding appears to contradict our earlier conclusion that spatial intermodal attention is mainly performed by modality-specific (sensory) areas of the brain. However, it is

Auditory Topographies

Figure 7. Scalp topographies (top) and C.S.D. maps (bottom) for the attention effects following auditory stimuli. Contour spacing is $0.1 \mu\text{V}$ for the potential maps and $0.05 \mu\text{V}/\text{cm}^2$ for C.S.D. maps. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure, where the left side of each map represents the ipsilateral presentation side, and the right side of each map the contralateral.

well possible that the observed differences are originating from a bias in sensitivity in sensory brain areas that is set at the beginning of each task block by control mechanisms that are not specific to one modality (e.g., Brunia, 1999; Gitelman et al., 1999; LaBerge, 1990). See below for a more elaborate discussion of this topic.

Intermodal Attention

The present results are in agreement with the earlier findings of De Ruiter et al. (1998), Eimer and Schröger (1998), Woods et al. (1992), and Talsma and Kok (2001) that manipulation of intermodal attention in the auditory and visual modalities results primarily in modulation of modality-specific areas. The spatiotemporal pattern of ERP modulations in the present study further suggested that this involved the same areas that were affected by manipulation of visual and auditory spatial attention. Intermodal attention

affected both early ERP components (visual: N1 and P2 components, auditory: Nde) and late components (visual: P3, auditory: Ndl). For the early ERP components, facilitatory effects of attending to the modality of the stimulus occurred also for stimuli presented at unattended location. This is an indication that these attention effects reflected a fairly early level of sensory processing. Furthermore, attentional enhancements of the late ERP components (visual modality: parietal P3, auditory modality: Ndl) only occurred for stimuli that shared both the relevant modality and location. This probably indicates that late ERP effects reflected a higher, probably supramodal, level of analysis.

Intra- and Intermodal Interactions

The presence of spatial attention effects in M- conditions can be explained by assuming that, although the expression of attentional

Auditory Difference Waves Single vs. Mixed Modality Tasks

Figure 8. Difference waves for auditory stimuli, obtained by subtracting ERPS elicited by M+L– stimuli from ERPs elicited by M+L+ stimuli, in single and mixed modality conditions. Responses are collapsed across left and right presentation sides, such that the ERPs obtained at the ipsilateral and contralateral presentation side are combined in one figure. A selected number of channels are shown.

Table 5. Overview of the Time Slice Analyses for the Auditory Stimuli for which Significant Amplitude Differences Were Obtained in the Single Modality Task. Also Shown Are the Interactions between the Single and Mixed Modality Tasks

Single Modality Attention	Single vs. Mixed Modality		
	Task	Attention	Task × Attention
120–908 ms	F3 and F4 76–172 ms	112–908 ms	164–264 ms
			304–340 ms
			752–796 ms
112–272 ms 500–860 ms	P3 and P4 88–140 ms	120–272 ms 320–404 ms 508–868 ms	152–200 ms

Note: Effects were considered significant when statistical tests (using 1,18 degrees of freedom) showed $\alpha < 5\%$ for eight or more consecutive samples. Results were obtained from F3 and F4 electrodes (top) and P3 and P4 electrodes (bottom). Attention: factor attention (M+L+ vs. M+L–); Task: factor task.

mechanisms is carried out by neural networks that are located in the sensory brain areas, presumably by modulating the sensitivity of neurons to specific stimulus features, these expression systems are in turn controlled by a higher-order control system (Hopfinger, Buonocore, & Mangun, 2000; LaBerge, 1990, 1995; Posner & Petersen, 1990). This control system, in turn, then probably serves to signal the sensory areas of the visual or auditory system to either (a) boost processing of the whole sensory system (modality) in general, when this system is relevant, or (b) boost processing of a specific stimulus feature within one modality (see also Talsma and Kok (2001) for a similar conclusion). Stimulus location is generally considered to be a primary selection criterion (Hillyard & Anllo-Vento, 1998; LaBerge, 1995; Treisman & Gelade, 1980; Van der Heijden, 1992, 1993). When that is the case, it is possible that attention to a given location of stimuli in one modality activates not only the processing in the modality that is relevant at the moment, but also to the modality that is not relevant at the moment. Thus, one could say that, for spatially relevant stimulus attributes, there exist attentional links between vision and audition (Eimer & Driver, 2001; Eimer & Schröger, 1998; Hackley, Woldorff, & Hillyard, 1990;

Spence & Driver, 1996, 1997a; Teder-Sälejärvi, Münte, et al., 1999). But, as we shall see below, differences in both time course and scalp distribution suggest that mechanisms of intra- and inter-modal attention as well as links between those modalities work differently in visual and auditory modalities.

Finally, the attentive reader should have noticed that, in general, effects of attention to visual and auditory stimuli generally started over sensory-specific brain areas that are followed later by a more distributed activation. That is, effects of attention (intramodal and intermodal) to visual stimuli start over the parietal areas and are followed somewhat later by a more globally distributed pattern of activation. For auditory stimuli, the earliest effects of attention were found only over the frontal areas, whereas a late P3 effect was significant over both frontal and parietal areas.

Note, though, that there is an exception to this pattern, which is the results obtained from visual stimuli in the M–L– condition (intramodal attention) and the L– condition. We found that ERPs in this category show a strong reduction in the size of the frontal N1 component at frontal electrodes as compared to ERPs from the other conditions, but not at parietal locations.

Visual and Auditory Processing: Differences in Time Course

The patterns of ERP results that were obtained for stimuli presented in the visual and auditory modalities further suggested that there were specific differences between the visual and auditory systems in the selective processing of attributes such as the modality and spatial location of stimuli. These differences were related to the timing rather than the locus of stimulus selection process (see also the respective ERP difference waves plotted in Figures 2 and 6). In audition, the facilitatory effect of intermodal attention had a relatively early onset (Nde: around 108 ms after stimulus onset) and equally affected stimuli at attended and unattended locations. Spatial (intramodal) auditory attention, however, had a somewhat later onset (Nde: 150 ms after stimulus onset) and primarily facilitated stimuli at the attended location. Although weaker, this effect was present not only when audition was the attended modality, but also when audition was not the relevant modality. In the visual modality, a somewhat deviant ERP pattern was found. Here, the negative modulatory effect on the N1 and P2 components had a similar onset for intermodal and intramodal (spatial) attention. In addition, the facilitatory effect of intermodal attention on the N1 and P2 was also present at the unattended location. The similar onset latencies of the visual intra- and intermodal ERP effects suggests that, in vision, the processing of stimulus attributes such as modality and spatial location occurred in a more parallel fashion than is the case for auditory stimuli.

These differences in the time course of visual and auditory ERPs seem to support theories claiming that processing of spatial information is qualitatively different in vision than in the audition. For instance, it has been proposed that the visual system is inherently spatially organized and thus more efficient in selection and utilization of spatial information than the auditory system (Buchtel & Butter, 1988; Recanzone, Guard, Phan, & Su, 2000; Spence & Driver, 1997a). Conversely, separating auditory stimuli on the basis of location might be a computationally much more complex operation (Teder-Sälejärvi, Hillyard et al., 1999). The findings of the present study that, for the auditory modality, location selection was hierarchically dependent on selection of the modality of the stimulus, is in line with the earlier hypothesis.

Spatial Selection: Mixed Versus Single Tasks

A final issue that deserves emphasis concerns the interpretation of effects of spatial attention that were observed for stimuli presented to the unattended modality. Were these cross-modal interactions indeed a reflection of an inherent property of spatial attention, or did they perhaps reflect that subjects divided to a certain extent their attention between the two modalities? The interpretation of effects of spatial attention to stimuli presented to the unattended modality becomes problematic if subjects indeed divided their attention between different modalities. The design of the present study made it possible to test this alternative interpretation because visual and auditory stimuli were presented in two versions, a single and a mixed modality version. By contrasting performance measures and ERP difference waves that reflected effects of spatial attention in these two tasks it was possible to verify whether they perhaps divided their attention to some extent between the two modalities.

Inspection of these single/mixed differences indeed raised the suspicion that visual and auditory tasks differed in the amount of interference that was caused by stimuli that were presented in the unattended modality. Two sets of results are worth mentioning here. First, when audition was the relevant modality, subjects were responding significantly more slowly in the mixed than in the single modality condition (see Table 1). Second, a comparison between the ERP difference waves (see Figure 8) indicated that, for auditory ERPs, spatial attention effects as reflected in the early Nde component were significantly smaller in the mixed than single modality condition. These single/mixed differences did not occur in the tasks when vision was the relevant modality (e.g., see Figure 4). The Nde is assumed to reflect early attentional processes such as the tuning of the attentional trace in auditory supratemporal cortex (Näätänen, 1992).

Thus, the attenuation of these processes in the mixed modality condition could mean that subjects found it relatively difficult to focus auditory attention in the mixed condition as compared with the visual task. A possible mechanism that could explain such a difference was already described in the previous section. If selection of spatial attributes is indeed more difficult to accomplish in the auditory than in the visual system, then it is also plausible that subjects suffered from more interference by stimuli presented to the other modality in the mixed version of the auditory task.

General Conclusions

To summarize, the present study has been successful in confirming the existence of symmetrical links between attention to vision and audition. That is, effects of location relevance were obtained in both visual and auditory modalities when these modalities were not attended. However, our results suggest that for visual stimuli these effects were somewhat larger than for auditory stimuli. The latter pattern suggests that the visual system is capable of selecting both relevant modality and location in parallel, whereas the auditory system selects modality prior to location, in a more serial and hierarchical fashion. This discrepancy could be explained by the fact that the visual system is inherently spatially organized, whereas this is not the case for the auditory system. Similarly, this difference in architecture of the visual and auditory systems can also explain why intermixing visual and auditory stimuli affected performance and ERPs following auditory stimuli, whereas performance and ERPs following visual stimuli were not affected when auditory stimuli were intermixed.

Scalp topographies of the attention effects showed distinct differences for visual and auditory selection. These topographies suggest that the selection processes are most likely occurring within the sensory visual and auditory brain areas. The existence of cross-modal links between visual and auditory attention, however,

suggest that the selection processes in the visual and auditory cortices are under the control of a higher-order supramodal control process (see, e.g., LaBerge, 1990), which controls the sensitivity of the sensory areas.

REFERENCES

- Alho, K., Töytäjä, L., Reinikainen, K., Sams, M., & Näätänen, R. (1987). Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 68, 458–470.
- 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 auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, 82, 356–368.
- Benedict, R. H. B., Lockwood, A. H., Schucard, J. L., Schucard, D. W., Wack, D., & Murphy, B. W. (1998). Functional neuroimaging of attention in the auditory modality. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 9, 121–126.
- Brunia, C. H. M. (1999). Neural aspects of anticipatory behavior. *Acta Psychologica*, 101, 213–242.
- Buchtel, H. A., & Butter, C. M. (1988). Spatial attentional shifts: Implications for the role of polysensory mechanisms. *Neuropsychologia*, 26, 499–509.
- De Ruiter, M. B., Kok, A., & Van der Schoot, M. (1998). Effects of inter- and intramodal selective attention to non-spatial visual stimuli: An event-related potential analysis. *Biological Psychology*, 49, 269–294.
- Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*, 10, 731–735.
- Eason, R. G. (1981). Visual evoked potential correlates of early neural filtering during selective attention. *Bulletin of the Psychonomic Society*, 18, 203–206.
- Eason, R. G., Harter, M. R., & White, C. (1969). Effects of attention and arousal on visually evoked cortical potentials and reaction times in man. *Physiology and Behavior*, 4, 283–289.
- Eason, R. G., Oakley, M., & Flowers, L. (1983). Central neural influences on the human retina during selective attention. *Physiological Psychology*, 1, 18–28.
- Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, 25, 497–511.
- Eimer, M., & Schröger, E. (1998). Effects of intermodal attention and cross-modal attention in spatial attention. *Psychophysiology*, 35, 313–327.
- Fuster, J. M. (1997). The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe (3rd ed.). New York: Lippincott-Raven.
- Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., Kim, Y.-H., Meyer, J. R., & Mesulam, M. M. (1999). A large-scale distributed network for covert spatial attention: Further anatomical delineation based on stringent behavioural and cognitive controls. *Brain*, 122, 1093–1106.
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, 28, 240–244.
- Hackley, S. A., Woldorff, M. G., & Hillyard, S. A. (1990). Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. *Psychophysiology*, 27, 195–208.
- Hansen, J. C., & Hillyard, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 1–19.
- Harter, M. R., Aine, C. J., & Schroeder, C. (1982). Hemispheric differences in the neural processing of stimulus location and type: Effects of selective attention on visual evoked potentials. *Neuropsychologia*, 20, 412–438.
- Heslenfeld, D. J. (1998). Features and attention in vision: An analysis of electromagnetic brain responses. Unpublished doctoral dissertation, University of Amsterdam.
- Hillyard, S. A., & Anillo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the USA*, 95, 781–787.
- 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.
- Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature neuroscience*, 3, 284–291.
- Kenemans, J. L., Molenaar, P. C. M., & Verbaten, M. N. (1991). Models for estimation and removal of artifacts in biological signals. In R. Weiskunat (Ed.), *Digital biosignal processing*. New York: Elsevier.
- LaBerge, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience*, 2, 358–372.
- LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31, 291–308.
- 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., & Luck, S. J. (1992). Electrocortical 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, MA: 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.
- Michie, P. T., Bearpark, H. M., Crawford, J. M., & Glue, L. C. (1990). The nature of selective attention effects on auditory event-related potentials. *Biological Psychology*, 30, 219–250.
- Näätänen, R. (1982). Processing negativity: An evoked potential reflection of selective attention. *Psychological Bulletin*, 92, 605–640.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, 13, 201–288.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Neville, H. J., & Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task: An event-related potential and behavioral study. I. Normal hearing adults. *Brain Research*, 405, 253–267.
- 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.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25–42.
- Recanzone, G. H., Guard, D. C., Phan, M. L., & Su, T. I. K. (2000). Correlation between the activity of single auditory cortex neurons and sound-localization behavior in the macaque monkey. *Journal of neurophysiology*, 83, 2723–2739.
- Rugg, M. D., Milner, A. D., Lines, C. R., & Phalps, R. (1987). Modulation of visual event-related potentials by spatial and non-spatial visual selective attention. *Neuropsychologia*, 25, 85–96.

- Scherg, M., & Berg, P. (1995). BESA brain electric source analysis. User manual, version 2.1. Munich: Megis Software GmbH.
- Schröger, E. (1994). Human brain potential signs of selection by location and frequency in an auditory transient attention situation. *Neuroscience Letters*, 173, 163–166.
- Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, 8, 528–539.
- Schröger, E., & Eimer, M. (1993). Effects of transient spatial attention on auditory event-related potentials. *Neuroreport*, 4, 588–590.
- 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. (1997a). Audiovisual links in exogenous covert spatial attention. *Perception and Psychophysics*, 59, 1–22.
- Spence, C. J., & Driver, J. (1997b). On measuring selective attention to an expected sensory modality. *Perception and Psychophysics*, 59, 389–403.
- 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.
- Teder-Sälejärvi, W. A., Hillyard, S. A., Röder, B., & Neville, H. J. (1999). Spatial attention to central and peripheral auditory stimuli as indexed by event-related potentials. *Cognitive Brain Research*, 8, 213–227.
- Teder-Sälejärvi, W. A., Münte, T. F., Sperlich, F.-J., & Hillyard, S. A. (1999). Intra-modal and cross modal spatial attention to auditory and visual stimuli: An event-related brain potential study. *Cognitive Brain Research*, 8, 327–343.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Van der Heijden, A. H. C. (1992). *Selective attention in vision*. London, UK: Routledge.
- Van der Heijden, A. H. C. (1993). The role of position in object selection in vision. *Psychological Research*, 56, 44–58.
- Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1583–1590.
- Wijers, A. A., Lange, J. J., Mulder, G., & Mulder, L. J. M. (1997). An ERP study of visual spatial attention and letter target detection for isoluminant and nonisoluminant stimuli. *Psychophysiology*, 34, 553–565.
- 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., Hackley, S. A., & Hillyard, S. A. (1993). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, 28, 30–42.
- Woldorff, M. G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., Heinze, H.-J., & Scheich, H. (1999). Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Human Brain Mapping*, 7, 49–66.
- Woods, D. L., Alho, K., & Algazi, A. (1992). Intermodal selective attention. I: Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalography and Clinical Neurophysiology*, 82, 341–355.

(RECEIVED October 30, 2000; ACCEPTED March 30, 2002)