

Active Listening Impairs Visual Perception and Selectivity: An ERP Study of Auditory Dual-task Costs on Visual Attention

Elena Gherri^{1,2} and Martin Eimer²

Abstract

■ The ability to drive safely is disrupted by cell phone conversations, and this has been attributed to a diversion of attention from the visual environment. We employed behavioral and ERP measures to study whether the attentive processing of spoken messages is, in itself, sufficient to produce visual-attentional deficits. Participants searched for visual targets defined by a unique feature (Experiment 1) or feature conjunction (Experiment 2), and simultaneously listened to narrated text passages that had to be recalled later (encoding condition), or heard backward-played speech sounds that could be ignored (control condition). Responses to targets were slower

in the encoding condition, and ERPs revealed that the visual processing of search arrays and the attentional selection of target stimuli were less efficient in the encoding relative to the control condition. Results demonstrate that the attentional processing of visual information is impaired when concurrent spoken messages are encoded and maintained, in line with cross-modal links in selective attention, but inconsistent with the view that attentional resources are modality-specific. The distraction of visual attention by active listening could contribute to the adverse effects of cell phone use on driving performance. ■

INTRODUCTION

When two tasks have to be performed simultaneously, performance decrements are frequently observed, and these dual-task costs are often ascribed to attentional limitations (e.g., Pashler & Johnston, 1998). Such limitations may be much less relevant when the two tasks involve different sensory modalities. For example, it has been argued that attention can be efficiently divided between visual and auditory tasks, because the mechanisms that underlie the processing of visual and auditory events draw upon separate and independent modality-specific attentional resources (e.g., Wickens, 1984). More recent studies that have investigated cross-modal links in attention have found results that are inconsistent with such multiple-resource models of dual-task performance. For example, Spence and Driver (1997) found that visual-spatial discrimination performance was reduced when participants had to simultaneously attend to audition in order to detect occasional auditory targets, suggesting shared attentional resources between vision and audition. In addition, a number of behavioral and electrophysiological studies have demonstrated the existence of strong cross-modal links in spatial attention between vision, audition, and touch, and have shown that links can affect early perceptual processing stages (e.g., Eimer, Van Velzen, & Driver, 2002;

Spence, Pavani, & Driver, 2000; Spence & Driver, 1997; see also Eimer & Driver, 2001 for a review).

This debate about modality-specific versus supramodal attentional resources, cross-modal links in attention, and the consequences of dividing attention between different sensory modalities is not just of theoretical interest, but also has important practical implications. It is well known that using a mobile or cellular phone while driving substantially increases the risk of a motor-vehicle accident (e.g., Redelmeier & Tibshirani, 1997). There is little evidence that this type of dual-task interference is linked to peripheral factors such as sensorimotor processes that are associated with holding and operating the phone (e.g., Strayer & Johnston, 2001; Redelmeier & Tibshirani, 1997). It has therefore been argued that cell phone conversations impair driving performance because they distract selective visual attention: When drivers are engaged in such conversations, attention is partially withdrawn from relevant aspects of the visual scene (Strayer & Johnston, 2001). Such a diversion of visual attention can even result in “inattentive blindness,” where objects in the driving environment are no longer registered (e.g., Strayer & Drews, 2007; Strayer, Drews, & Johnston, 2003). This was demonstrated by showing that recognition memory for visual objects encountered during simulated driving is reduced when drivers had been using a cell phone relative to a single-task condition (Strayer et al., 2003).

To gain more precise insights into the conditions under which cell phone conversations result in driving-related

¹City University, London, UK, ²Birkbeck College, University of London, London, UK

performance deficits, Strayer and Johnston (2001) employed a pursuit tracking task to simulate the visual demands of driving. Participants used a joystick to continuously track a target that moved unpredictably on a computer screen. They had to press a button whenever the target flashed red, while ignoring green flashes. The detection of red target flashes was slower and less reliable when participants used a cell phone relative to a single-task control condition. This performance impairment was attributed to a diversion of attention from the visual task toward the cognitively engaging cell phone conversation. Interestingly, even though this dual-task deficit was larger when participants were talking, it was still present during intervals where they were listening, suggesting that the attentive processing of spoken messages might, in itself, interfere with visual task performance. To underline the important role of the attentional demands of an auditory task, Strayer and Johnston (2001) demonstrated that visual tracking performance was impaired when participants had to generate new words in response to words communicated over the cell phone, but not in a simpler shadowing task where these words only had to be repeated (see also Kunar, Carter, Cohen, & Horowitz, 2008, for similar findings).

It is important to note that engaging in a conversation does not inevitably affect driving performance. Cell phone conversations impair the quality of driving much more than engaging in a conversation with a passenger (Drews, Pasupahti, & Strayer, 2008), presumably because driver and passenger are both exposed to the same road traffic situation, and are therefore able to adjust the speed and complexity of their conversation to the demands of the driving environment. Such a flexible real-time allocation of attentional resources across tasks is more difficult when the communication partner has no access to the current driving situation. For similar reasons, simply listening to a radio broadcast does not have adverse effect on performance in a simulated driving task (Strayer & Johnston, 2001), as drivers are free to withdraw their attention from the auditory message and direct it toward the visual environment whenever this is required by the traffic situation.

In summary, recent evidence suggests that driving is impaired by a cognitively engaging auditory task such as a cell phone conversation because such a task diverts attention from the visual environment, and will do so, in particular, when it is temporally decoupled from the on-line demands of the driving situation. This apparent attentional interference between concurrent visual-spatial and auditory-verbal processing is problematic for models that assume the existence of independent modality-specific attentional resources (e.g., Wickens, 1984). In contrast, these dual-task costs are consistent with the alternative view that there are strong and reciprocal links between visual and auditory attention (e.g., Spence & Read, 2003; Eimer & Driver, 2001). However, more detailed insights into the mechanisms that underlie the visual processing

costs induced by simultaneous auditory tasks have yet to be obtained. For example, it is not yet clear which stages of visual processing are modulated by the supposed diversion of attention from the visual environment toward a concurrent auditory task. The adverse effects of an auditory task on visual performance could manifest themselves already at early sensory stages of visual perception, might affect the attentional selection of task-relevant visual events, or could be confined to later postperceptual stages associated with response selection and execution. Another important question is whether some of the adverse effects of cell phone use on driving are associated purely with the attentional demands of speech reception, that is, with processes involved in the perception, analysis, and storage of vocally communicated information. Alternatively, such dual-task costs might only emerge in situations that emphasize speech production, as is the case in an active and reciprocal phone conversation. Results from two previous experiments (Kunar et al., 2008; Strayer & Johnston, 2001) suggest that speech production could indeed be critical. In these studies, visual tracking performance was unimpaired when participants attentively listened to spoken text, but was adversely affected by the requirement to actively generate verbal responses.

In the present study, we combined behavioral and ERP measures of visual processing in attentional selection tasks that were performed with or without a simultaneous auditory task that involved the active processing of spoken text. One aim was to further investigate whether the requirement to encode and maintain spoken auditory messages is, in itself, sufficient to adversely affect visual processing and visual selectivity in a concurrent attention task. Another aim was to study which stages of visual-attentional processing, if any, are affected when participants have to simultaneously process verbally communicated information. ERPs provide a continuous on-line measure of perceptual and cognitive processes during single- and dual-task performance, and are therefore ideally suited to track the impact of auditory task demands on the attentional processing of visual information. Two different visual search tasks were employed. In Experiment 1, participants searched for targets that differed from distractors in terms of their unique color (singleton search). In Experiment 2, targets were defined by a combination of color and shape (conjunction search). Throughout all visual search blocks, a stream of auditory events was presented, and the critical manipulation concerned the task-relevance and the semantic content of these events. In the encoding condition, participants listened to spoken passages from an audiobook ("stories") where a narrator described different travel adventures. They were instructed to attentively process these stories in order to answer two test questions about their content at the end of each block, without sacrificing speed or accuracy in the concurrent visual search task. In the control condition, these stories were played backward, which made them incomprehensible. Participants had to

ignore the auditory events and focus fully on the visual search task.

We opted for conventional visual search tasks instead of more direct approximations of real-life driving situations such as pursuit tracking (e.g., Strayer & Johnston, 2001) or measuring visual performance in a driving simulator (e.g., Strayer et al., 2003), in order to prevent the contamination of ERP correlates of visual-attentional processing by artifacts resulting from eye, head, and body movements that are inevitable in more naturalistic task settings. However, because visual search engages perceptual and attentional processes that are also involved in the control of driving, such as the detection and selection of spatially unpredictable salient visual events (e.g., a pedestrian suddenly crossing the road), or the identification of stimuli that are defined by specific combinations of color and shape (e.g., road signs), studying the impact of a concurrent auditory task on ERP correlates of visual search performance can provide new insights into the nature of this type of dual-task interference that may also be applicable to more ecological situations.

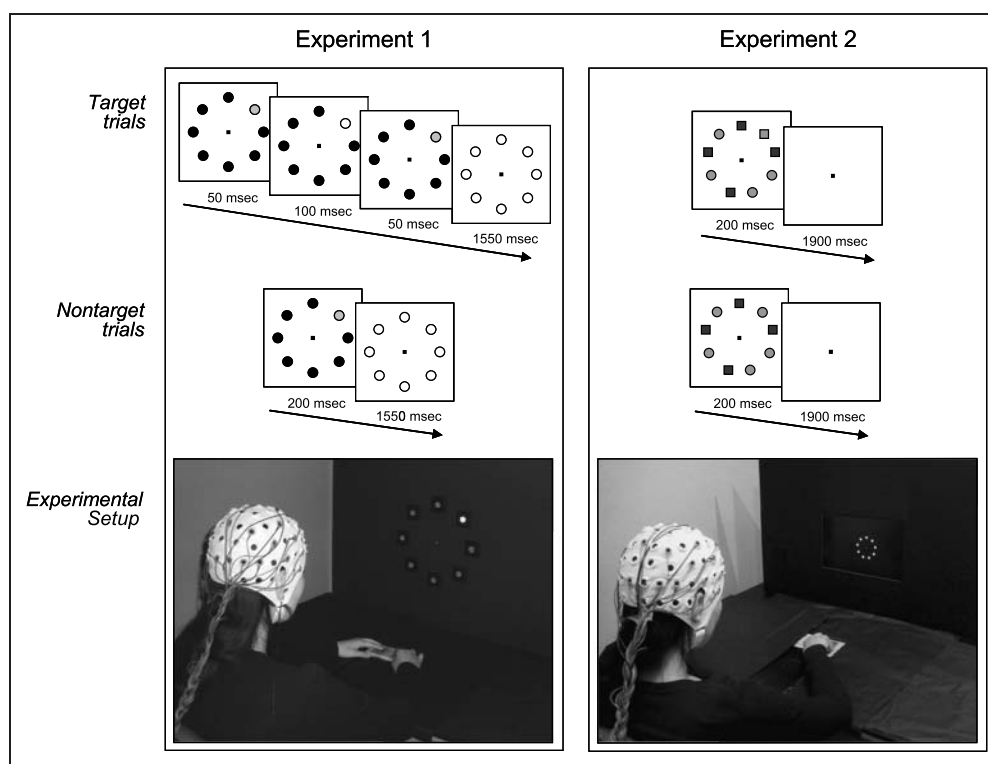
Because the low-level acoustic features of the auditory stimulation were equivalent in the encoding and control conditions, differences in the quality of visual processing between conditions, as reflected by performance and ERP measures, cannot be attributed to the presence of auditory stimulation as such, but instead to the additional demands of the auditory task in the encoding condition. If visual selective attention is impaired only when the concurrent auditory task involves speech production, but not under conditions where only the attentive encoding and maintenance of verbal material is required (as suggested by the behavioral results of Kunar et al., 2008 and Strayer & Johnston, 2001), there should be no systematic performance or ERP differences between the encoding and control conditions. Alternatively, if the attentive processing of a story impaired visual search, RTs to visual search targets should be slower in the encoding relative to the control condition. In this case, visual ERPs obtained in these two task conditions could provide additional new insights into which stages of visual processing are responsible for these behavioral costs. To investigate this, three sets of ERP analyses were conducted. First, we studied whether visual ERP components (P1, N1, P2, and N2) triggered by the search arrays are modulated by the attentionally demanding auditory task in the encoding condition. Previous studies have shown that these visual ERP components are sensitive to variations in selective attention, with enhanced early P1 and/or N1 components and a subsequent sustained attentional selection negativity for attended as compared to unattended visual stimuli (e.g., Hillyard & Anllo-Vento, 1998; Eimer, 1994; Mangun & Hillyard, 1991). If the attentional processing of visual events is compromised by a concurrent auditory task, ERP differences between the two task conditions should be similar to the effects found in previous studies where focal attention was explicitly manipulated, thereby demon-

strating that attentive listening can affect early stages of visual processing.

The other two analyses focused on ERPs triggered in response to visual search arrays that contained target stimuli, and investigated whether the detection and processing of visual target events is impaired by a concurrent cognitively engaging auditory task. To study the attentional selection of visual targets, we measured the N2pc component. The N2pc is an enhanced negativity over posterior scalp electrodes contralateral to the side of task-relevant visual stimuli that is elicited between 180 and 300 msec after the onset of a visual search array, and is assumed to reflect the spatial selection of candidate target items among distractors (e.g., Woodman & Luck, 1999; Eimer, 1996; Luck & Hillyard, 1994). If the efficiency of selecting visual search targets is adversely affected by the concurrent auditory task, this should be reflected by an increase in the latency and/or a decrease in the amplitude of the N2pc to visual target events in the encoding relative to the control condition. Finally, we measured the P3 component in response to visual search targets at centroparietal electrodes. The P3 is known to be sensitive to the amount of attention that is allocated to a specific task, with larger P3 amplitudes under focal as compared to divided attention conditions (e.g., Sirevaag, Kramer, Coles, & Donchin, 1989; Wickens, Kramer, Vanasse, & Donchin, 1983). In a previous study that used ERP measures to study how cell phone use impairs attentional processes in a driving-related visual task, Strayer and Drews (2007) observed reduced P3 amplitudes to visual target events (red brake lights) for participants that were engaged in a cell phone conversation relative to a single-task control group. If the processing and encoding of spoken text diverted attention from the concurrent visual search task, this should be reflected in a reduction of P3 amplitudes to visual targets in the encoding relative to the control condition.

These predictions were tested in two experiments that differed with respect to visual presentation conditions and the properties of visual search targets. In Experiment 1, visual stimuli were presented in the free field via eight sets of LEDs that were arranged in a virtual circle (see Figure 1, left). One of these stimuli was a color singleton (a red LED among green LEDs, or vice versa). On target trials, this singleton had a gap (i.e., it was rapidly switched off and then on again), and participants had to report its color. Free-field visual stimulation was employed in order to approximate the spatial parameters of visual selectivity during driving. In Experiment 2, visual stimuli were presented on a computer screen, and targets were now defined by a specific color-shape combination (see Figure 1, right). All distractors shared one of the target-defining features so that targets could not be found on the basis of one feature alone. Comparing a task where candidate target events were color singletons and a task where target detection required feature integration makes it possible to determine whether the presence of auditory dual-task costs in visual

Figure 1. Task setup and visual stimulation procedures used in Experiments 1 and 2. In Experiment 1 (left), visual stimuli were presented via LEDs, and participants searched for color singleton targets that were defined by a 100-msec gap. On nontarget trials, color singletons had no gap. To visualize the position of a color singleton, the bottom left photograph shows the singleton as brighter than the distractors, although, in reality, stimuli were equiluminant. In Experiment 2 (right), visual stimuli were presented on a computer screen. Targets were defined by a specific combination of a color and shape, and distractors shared one of these features with the target (see text for more details).



search tasks depends on the specific attentional demands of these tasks.

EXPERIMENT 1

Methods

Participants

Sixteen paid volunteers participated in this experiment (10 women, 6 men; age = 19–36 years, average age = 25.9 years). Three were left-handed, 13 were right-handed, and all had normal or corrected-to-normal vision. Written consent was obtained from all participants. The experiment was approved by the local ethics committee and was conducted following the guidelines of the Declaration of Helsinki.

Stimuli

Participants sat in a dimly lit sound-attenuated cabin in front of a black cardboard panel (90 × 60 cm) and fixated a small white circle point at the center of the panel at a viewing distance of 57 cm. Eight LED ensembles were mounted on the panel and were arranged in a virtual circle at 10° visual angle from the fixation point (see Figure 1, left). LED ensembles were composed of six LEDs arranged in a circle plus one central LED (size of each LED: 0.4 cm; circle diameter: 2.4 cm).

Visual search displays were presented by illuminating the eight LED ensembles for 200 msec. One of them was a color singleton (red or green) that was presented among seven uniformly colored LED ensembles (green or red).

Singletons were presented with equal probability and in random order at one of the six positions to the left and right of fixation, but never at the top or bottom positions of the virtual circle. On nontarget trials, singleton and distractor LEDs were illuminated for 200 msec. On target trials, the singleton LED was illuminated for 50 msec, turned off for 100 msec, and turned on again for 50 msec (gap stimulus), whereas distractor LEDs remained on for 200 msec. Red and green LED ensembles were approximately equiluminant (13.7 and 14.1 cd/m², respectively).

Auditory stimuli consisted of 16 passages (“stories”) of a continuous spoken text that was digitally recorded from an audiobook (*Around the World in 80 Days*, narrated by Michael Palin; Palin, 2003). Each passage was approximately 2.5 min long. Eight different stories were played during the experimental blocks, and four others were used for the training blocks. In the encoding condition, the recordings of the eight stories were presented normally. The same eight stories were played backward in the control condition, which rendered them incomprehensible but maintained the same level of auditory stimulation. Stories were presented via a speaker centrally located behind the cardboard panel and controlled by the sound card of a laptop PC. Each story started 2000 msec before the first visual stimulus presentation, and continued until after the last visual search array in a given block was presented.

Procedure

The experiment consisted of 16 experimental blocks with 72 trials per block. Target color singletons (gap stimuli)

were presented on one third of the trials (24 trials per block) with equal probability and in random order at one of the six lateral positions on the left or right side. On the remaining 48 trials, a nontarget singleton (nongap stimulus) was presented at one of these positions. Trials with red singletons among green distractors or green singletons among red distractors were presented with equal probability and in random order. Each search array was presented for 200 msec, and the interval between stimulus offset and the onset of the search array on the next trial was 1550 msec.

Participants' task was to detect singleton targets (gap stimuli) and to report their color (red or green) with a left-hand or right-hand button press. They were instructed to keep central fixation throughout each trial, to respond to target singletons as fast and accurately as possible, and to refrain from responding on nontarget trials (trials with color singletons without gap). Response keys were arranged vertically, and participants pressed the top or bottom key with their left or right index finger to indicate target color. The mappings between response hand and target color, and between response hand and response key, were counterbalanced across participants.

Participants performed two different task conditions, each presented in eight successive experimental blocks (resulting in a total of 576 trials per condition), with the order in which these two conditions were presented counterbalanced across participants. In the encoding condition, they were instructed to concentrate on the visual task, but at the same time to carefully listen to the story in order to answer two multiple-choice questions that were asked by the experimenter at the end of each block. These questions were related to the content of the story presented during the preceding experimental block, and participants had to choose one of three possible answers. For example, following a story describing a journey in America, one of the two questions was "Which river is Michael travelling down?" (possible answers: Hudson, Amazon, Mississippi). In the control condition, where stories were played backward, participants were asked to concentrate on the visual task and to ignore the auditory stimulation. The sequence in which the eight stories were presented was randomly determined for each participant, but they were always presented in the same order in the encoding and control conditions. Two training blocks of 72 trials per block were run prior to the start of the first experimental block in each condition.

EEG Recording and Data Analysis

The EEG was DC-recorded with a low-pass filter of 40 Hz and a sampling rate of 200 Hz from 23 Ag–AgCl electrodes mounted in an elastic cap according to the extended International 10–20 System at Fpz, F7, F3, Fz, F4, F8, FC5, FC6, T7, C3, Cz, C4, T8, CP5, CP6, P7, P3, Pz, P4, P8, PO7, PO8, and Oz. The left earlobe was used as on-line reference, and EEG data were re-referenced off-line relative the average

across left and right earlobes. Horizontal eye movements (HEOG) were measured bipolarly from a pair of electrodes placed at the outer canthi of the eyes. All electrode impedances were kept below 5 k Ω . The EEG was epoched from 100 msec prior to 600 msec after the onset of the visual display. Epochs containing blinks (Fpz exceeding ± 60 μ V), horizontal eye movements (HEOG exceeding ± 25 μ V), or movement artifacts (voltage exceeding ± 80 μ V at all other electrodes) were eliminated from further analyses.

Statistical analyses were conducted on ERP mean millisecond amplitudes obtained within successive measurement windows relative to a 100-msec prestimulus baseline. The first set of analyses included EEG data from target and nontarget trials. ERP mean amplitudes obtained within the P1 (100–130 msec), N1 (140–190 msec), P2 (200–240 msec), and N2 (250–300 msec) time windows at lateral posterior (P3/4, P7/8, PO7/8), lateral central (C3/4, CP5/6, T7/8), and midline electrodes (Cz, Pz, Oz) were analyzed with repeated measures ANOVAs for the factors task condition (encoding vs. control), trial type (target vs. nontarget), electrode (P3/4 vs. P7/8 vs. PO7/8 for lateral posterior sites; C3/4 vs. CP5/6 vs. T7/8 for lateral central sites; Cz vs. Pz vs. Oz, for midline electrodes), and recording hemisphere (left vs. right, for lateral electrodes only). Analyses of N2pc and P3 components were conducted for target trials only. The N2pc was quantified on the basis of mean amplitudes obtained at lateral posterior electrodes PO7/PO8 in two successive time windows (210–250 msec and 255–310 msec after search array onset). Analyses were conducted the factors task condition, target side (left vs. right), and contralaterality (electrode contralateral vs. ipsilateral to the side of the target). P3 mean amplitudes were measured in the 400–550 msec interval after search array onset at centro-parietal electrodes (C3/4, CP5/6, P3/4, Cz, Pz), and were analyzed for the factors task condition, target side, and electrode site. For all analyses, Greenhouse–Geisser adjustments to degrees of freedom were applied where appropriate.

Results

Behavior

RTs to color singleton targets (gap stimuli) were slower in the encoding condition than in the control condition (537 msec vs. 512 msec), resulting in a main effect of task condition [$F(1, 15) = 7.6, p < .015$]. This analysis excluded the 0.4% of all target trials with RTs longer than 1000 msec. Incorrect responses (i.e., reports of the wrong target color) were observed on 6.2% of all target trials. Participants failed to respond to 0.5% of all targets, and the false alarm rate on nontarget trials was 0.4%. Error rates did not differ between the two task conditions [$F(1, 15) < 1$]. In the encoding condition, participants answered 78% of all multiple-choice questions correctly, which was significantly above chance [33.3%; $t(15) = 12.4, p < .001$].

Event-related Potentials

P1, N1, P2, and N2. Figure 2 shows visual ERPs triggered during the 400-msec interval after search array onset in the encoding and control conditions at lateral posterior, lateral central, and midline electrodes, collapsed across target and nontarget trials. ERPs were more negative in the control relative to the encoding condition. This differential effect started about 140 msec after stimulus onset, and remained present in a sustained fashion for about 200 msec, overlapping with the N1, P2, and N2 components. This is further illustrated in Figure 2 (right), which shows a topographical map of the distribution of ERP amplitude differences between the encoding and control conditions obtained between 140 and 300 msec after search array onset, and demonstrates that the sustained enhanced negativity for the control relative to the encoding condition was most pronounced over posterior areas.

Statistical analyses found no effects of task condition in the P1 time window [100–130 msec; all $F(1, 15) < 1$]. In contrast, the sustained negativity for the control relative

to the encoding condition resulted in reliable effects of task condition at lateral posterior and central sites and at midline electrodes Cz, Pz, and Oz in the subsequent N1, P2, and N2 time intervals between 140 and 300 msec after search array onset [all $F(1, 15) > 9.0$; all $p < .01$]. There were no significant interactions between task condition and trial type during any of these time windows, indicating that this differential modulation of visual ERPs in the encoding and control tasks was elicited regardless of whether or not search arrays contained a target stimulus.

N2pc. Figure 3 (top) shows ERP waveforms elicited at electrodes PO7/8 contralateral (dashed lines) and ipsilateral (solid lines) to the location of a target singleton, separately for the encoding and control conditions. As expected, an N2pc component was elicited by targets in both conditions, and the scalp distribution of this component is shown in the topographical maps (Figure 3, top left). N2pc amplitude differences between the encoding and control

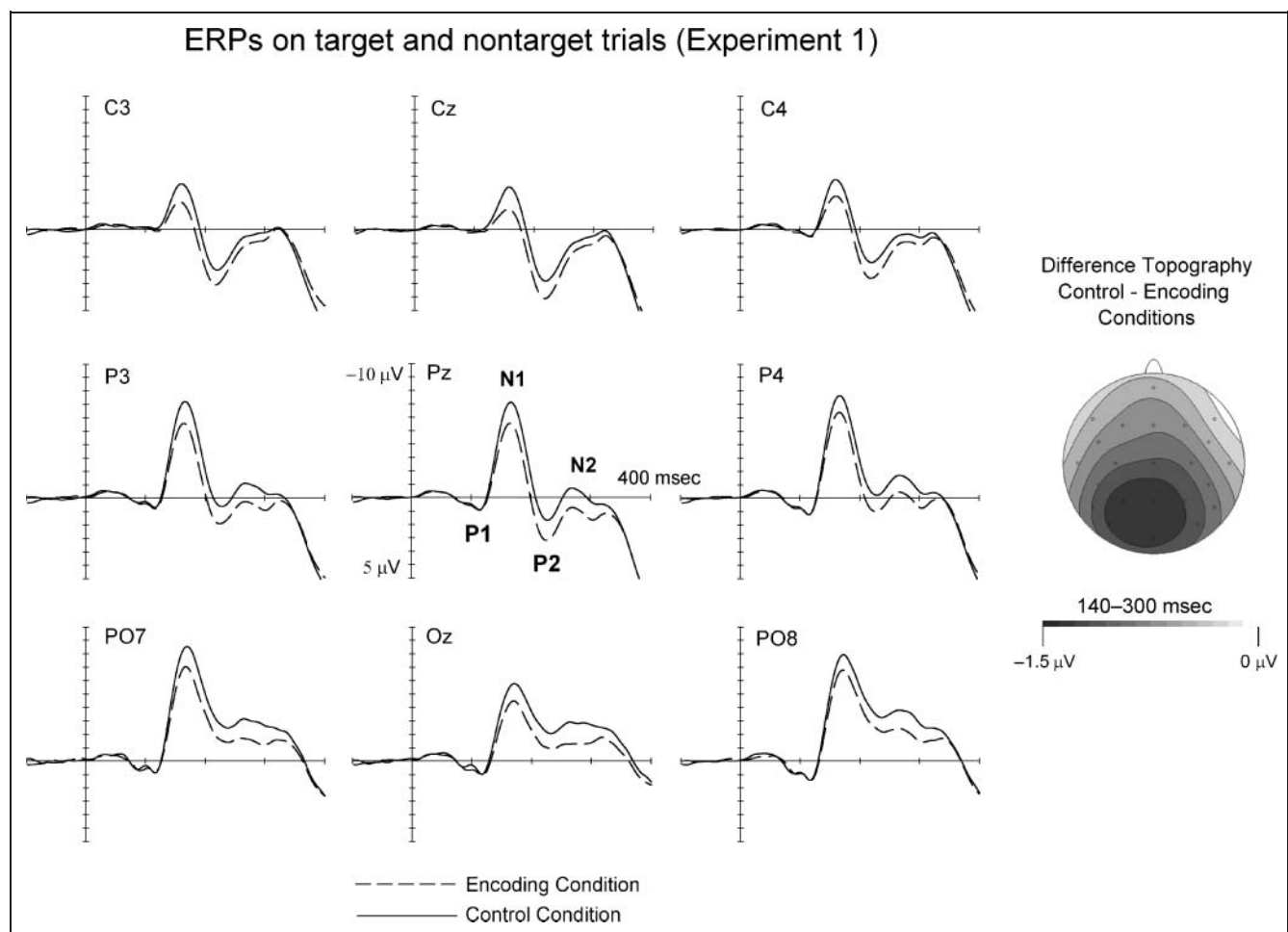
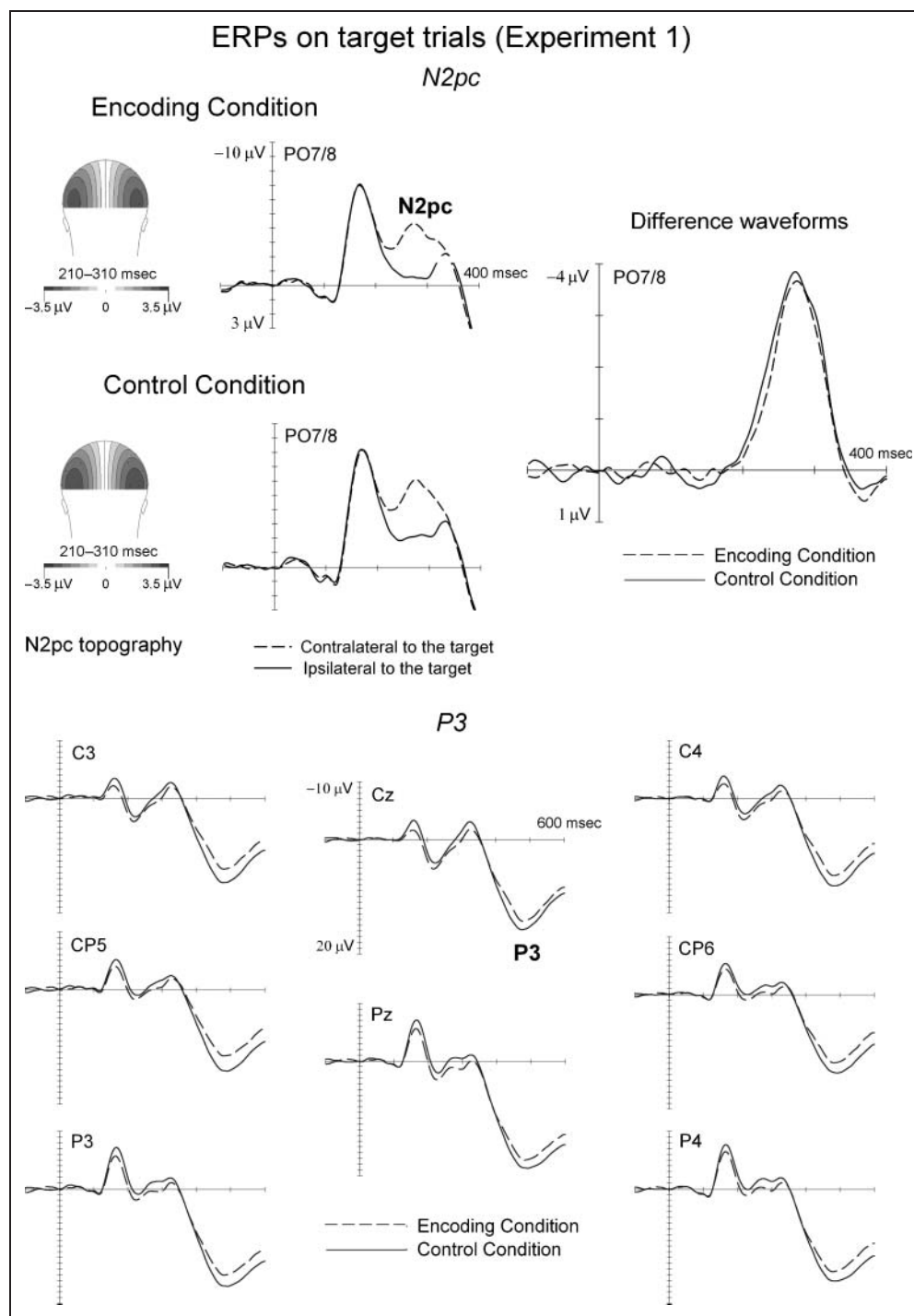


Figure 2. Grand-averaged visual ERPs elicited in the 400-msec interval after search array onset in the encoding condition (dashed lines) and control condition (solid lines) of Experiment 1. ERPs are averaged across target and nontarget trials. The topographic map on the right shows the scalp distribution of ERP amplitude differences between task conditions across the N1, P2, and N2 time windows. Difference amplitude values were obtained by subtracting ERP mean amplitudes measured in the 140–300 msec interval after search array onset in the encoding condition from mean amplitudes in the control condition.

Figure 3. Grand-averaged ERPs elicited on target trials in Experiment 1. Top: ERPs triggered in the 400-msec interval after search array onset in the encoding condition (top) and control condition (bottom) at posterior electrodes PO7/8 contralateral (dashed lines) and ipsilateral (solid lines) to the visual hemifield of the target. The N2pc component is shown in the difference waveforms on the right that were obtained by subtracting ipsilateral from contralateral ERPs for the encoding condition (dashed line) and control condition (solid line). The topographic maps (back views) on the left side show the scalp distribution of the N2pc in the 210–310 msec interval after search array onset for the encoding and control conditions. Maps were constructed by spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989) after mirroring the ipsilateral–contralateral difference waveforms to obtain symmetrical voltage values for both hemispheres. Bottom: ERPs triggered in the 600-msec interval after search array onset at centro-parietal electrodes in the encoding and control conditions, showing the P3 component in response to visual search targets.



conditions can be seen more clearly in the difference waveforms in Figure 3 (top right) that were obtained by subtracting ERPs elicited at electrodes ipsilateral to the target from contralateral ERPs. During the early phase of the N2pc, amplitudes were larger in the control condition. This was confirmed by analyses of N2pc amplitudes obtained between 210 and 250 msec after search array onset. Here, a main effect of contralaterality [$F(1, 15) = 13.9$, $p < .002$], reflecting the presence of an N2pc to targets, was accompanied by an interaction between task condi-

tion and contralaterality [$F(1, 15) = 5.8$, $p < .03$] due to the fact that the early phase of the N2pc was more pronounced in the control condition as compared to the encoding condition.¹ In contrast, during the later phase of the N2pc (255–310 msec after search array onset), there was still a main effect of contralaterality [$F(1, 15) = 32.5$, $p < .001$], but no interaction between task condition and contralaterality [$F(1, 15) = 1.2$, $p = .29$], indicating that the auditory task did not affect N2pc amplitudes to visual targets during this time range.² There were also main effects

of task condition during both N2pc time intervals [both $F(1, 15) > 14.4$, both $p < .002$], as nonlateralized ERPs elicited at PO7/8 in the 210–310 msec time window were generally more negative in the control task relative to the encoding task (as also shown in Figure 2).

P3. Figure 3 (bottom) shows ERPs triggered on target trials in the encoding and control conditions at centroparietal electrodes. P3 components elicited by targets were larger in the control condition. This was substantiated by the analysis of mean amplitudes in the P3 time window (400–550 msec after search array onset), which revealed a main effect of task condition [$F(1, 15) = 16.1$, $p < .001$]. There was no interaction between task condition and electrode site [$F(1, 15) = 1.7$].

EXPERIMENT 2

Methods

Participants

Nineteen paid volunteers participated in Experiment 2. Three had to be excluded because of excessive alpha activity or poor signal-to-noise ratio after artifact rejection, leaving 16 participants in the sample (11 women, 5 men; age = 19–20 years; average age = 24.9 years). All were right-handed and had normal or corrected-to-normal vision. Written consent was obtained from all participants. The experiment was approved by the local ethics committee, and was conducted following the guidelines of the Declaration of Helsinki.

Stimuli

Auditory stimuli were identical to Experiment 1. Visual stimuli were presented on a computer screen that was visible through a 20 × 20 cm hole in the center of the cardboard panel in front of the participants (see Figure 1, right). A light gray fixation point was continuously present throughout each block. Search displays consisted of nine stimulus elements that were presented at equidistant positions from central fixation (2.5° visual angle) and were arranged in a virtual circle centered on the fixation point. Stimuli were circles or squares that subtended 0.5° of visual angle. They were either green or blue (CIE 1931 x/y coordinates: 0.290/0.420 and 0.228/0.250 for green and blue stimuli, respectively), and these two colors were approximately equiluminant (35.2 and 33.7 cd/m², respectively). Participants performed a conjunction visual search, where search targets were defined by one specific combination of shape and color (e.g., blue squares), and each distractor shared one feature (either shape or color) with the target. On target trials, search displays contained eight distractors and one target. Targets were never presented at the top position of the virtual circle. On nontarget trials, nine distractors were presented.

Procedure

The experiment consisted of 16 experimental blocks with 72 trials per block. The auditory task manipulation (encoding vs. control conditions, both presented in 8 successive blocks) was identical to Experiment 1. In the conjunction visual search task, targets were presented on 48 trials per block, with equal probability at one of the eight lateral positions to the left or right of fixation, but never at the top location. The remaining 24 trials per block were nontarget trials. Each search display was presented for 200 msec, and the interval the offset of a search display and search display onset on the next trial was 1900 msec. Participants were instructed to keep central fixation throughout each trial, to detect target stimuli, and to report their location (left or right) as fast and accurately as possible by pressing a left or right response key with two fingers of the same hand. Response hand was changed every two blocks. Half of all participants started the experiment with left-hand responses, whereas the other half started with using their right hand. Target identity remained constant throughout the experiment for each participant, and was counterbalanced across participants, such that each of the four possible color–shape combinations served as target for four participants.

EEG Recording and Data Analysis

EEG recording and analysis procedures were identical to Experiment 1, with one exception. Because Experiments 1 and 2 employed different visual stimulation procedures (free-field LEDs vs. computer presentation) and target-defining features, latencies of visual ERP components and target-elicited components (N2pc and P3) differed between these two experiments. Therefore, slightly different ERP analysis windows were used in Experiment 2. Poststimulus measurement windows were now 110–130 msec (P1), 160–200 msec (N1), 210–250 msec, 260–310 msec (N2), and 450–600 msec (P3). For the N2pc analysis, the two successive time windows used were 250–295 and 300–340 msec.

Results

Behavior

RTs to targets were slower in the encoding condition than in the control condition (526 msec vs. 495 msec), resulting in a main effect of task condition [$F(1, 15) = 5$, $p < .041$]. This analysis excluded trials where RTs were longer than 1000 msec (0.9% of all target trials). Incorrect responses (i.e., reports of targets in the incorrect visual hemifield) were observed on 1.7% of all target trials. Participants failed to respond to 1.8% of all targets, and the false alarm rate on nontarget trials was 4.1%. Error rates did not differ between the two task conditions [$F(1, 15) < 1$]. In the encoding condition, participants answered 77% of all multiple-choice questions correctly, which was significantly above chance [33.3%; $t(15) = 12.1$, $p < .001$].

Event-related Potentials

P1, N1, P2, and N2. Figure 4 shows ERPs triggered in the encoding and control conditions at lateral posterior, lateral central, and midline electrodes in the 400-msec interval after search array onset, collapsed across target and nontarget trials. Confirming the observations of Experiment 1, a sustained negativity for visual ERPs was observed in the control relative to the encoding condition. This effect overlapped with N1, P2, and N2 components, and its posterior distribution is shown in the scalp map (Figure 4, right) of ERP differences between the control and encoding condition amplitudes across the N1, P2, and N2 time windows (160–310 msec after search array onset).

There were no effects of task condition in the P1 time window [100–130 msec; all $F(1, 15) < 1.2$]. In contrast, the sustained negativity for the control relative to the encoding condition was present in the N1, P2, and N2 time windows, resulting in significant effects of task condition at lateral posterior and central sites and at midline electrodes between 160 and 310 msec after search array onset [all $F(1, 15) > 8.7$, all $p < .01$].³ There were no significant in-

teractions between task condition and trial type during any of these time windows, demonstrating that these ERP differences between the encoding and control tasks were similar on target and nontarget trials.

N2pc. Figure 5 (top) shows ERP waveforms at electrodes PO7/8 contralateral (dashed lines) and ipsilateral (solid lines) to the location of a visual search target, separately for the encoding and control conditions, together with N2pc scalp topographies (left panels). As in Experiment 1, an N2pc component was elicited by targets in both conditions, but its amplitude was again larger in the control condition. This is illustrated in the difference waveforms obtained by subtracting ipsilateral from contralateral ERPs (Figure 5, top left), which show that N2pc peak amplitudes triggered around 300 msec after search array onset were larger in the control condition. For the early phase of the N2pc (250–295 msec), a main effect of contralaterality was present [$F(1, 15) = 19.5$, $p < .001$], reflecting the presence of the N2pc, but there was no Task condition \times Contralaterality interaction [$F(1, 15) < 1$], indicating that the auditory task did not affect N2pc amplitudes during

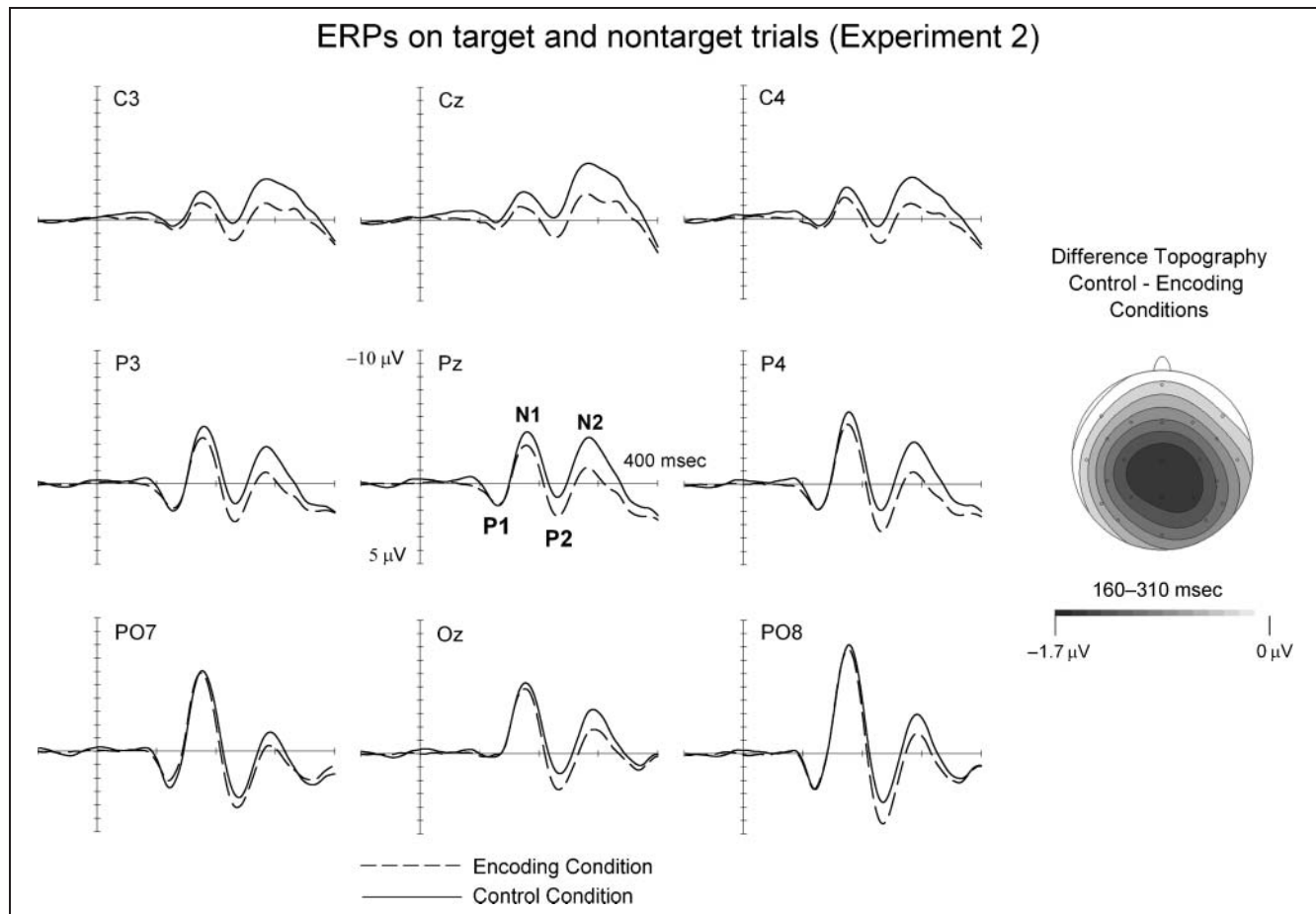
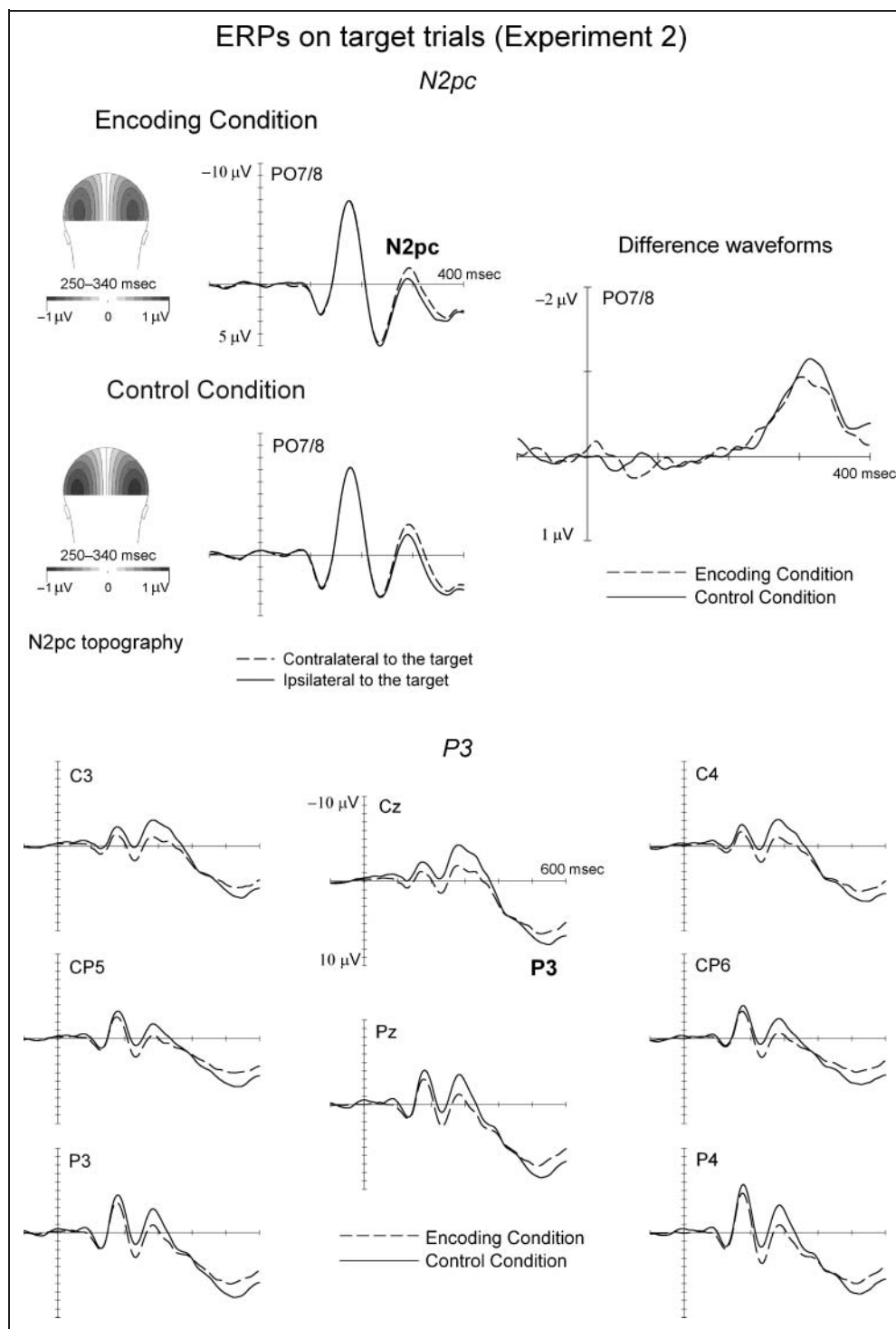


Figure 4. Grand-averaged visual ERPs elicited in the 400-msec interval after search array onset in the encoding condition (dashed lines) and control condition (solid lines) of Experiment 2. ERPs are averaged across target and nontarget trials. The topographic map on the right shows the scalp distribution of ERP amplitude differences between task conditions across the N1, P2, and N2 time windows. Difference amplitude values were obtained by subtracting ERP mean amplitudes measured in the 140–300 msec interval after search array onset in the encoding condition from mean amplitudes in the control condition.

Figure 5. Grand-averaged ERPs elicited on target trials in Experiment 2. Top: ERPs triggered in the 400-msec interval after search array onset in the encoding condition (top) and control condition (bottom) at posterior electrodes PO7/8 contralateral (dashed lines) and ipsilateral (solid lines) to the side of the target. The N2pc component is shown in the difference waveforms on the right that were obtained by subtracting ipsilateral from contralateral ERPs for the encoding condition (dashed line) and control condition (solid line). Topographic maps (back views) on the left show the N2pc scalp distribution in the 250–340 msec interval after search array onset for the control and encoding conditions. Bottom: ERPs triggered in the 600-msec interval after search array onset at centro-parietal electrodes in the encoding and control conditions, showing the P3 component in response to visual search targets.



this time interval. In contrast, during the later N2pc analysis window (300–340 msec), a main effect of contralaterality [$F(1, 15) = 39.5, p < .001$] was accompanied by an interaction between task condition and contralaterality [$F(1, 15) = 5.9, p < .028$], thus substantiating the observation that N2pc amplitudes were larger in the control relative to the encoding condition. In the early N2pc time window (250–295 msec), a main effect of task condition [$F(1, 15) = 9.4, p < .008$] confirmed that nonlateralized

ERPs at PO7/8 were generally more negative in the control task relative to the encoding task (see also Figure 4). This differential effect was no longer present in the 300–340 msec time window [$F(1, 15) = 1.4, p = .26$].

P3. Figure 5 (bottom) shows ERPs elicited on target trials in the encoding and control conditions at centro-parietal electrodes. As in Experiment 1, P3 amplitudes were larger in the control condition. This was confirmed by the presence

of a main effect of task condition [$F(1, 15) = 5.3, p < .036$] in the analysis of ERP mean amplitudes in the P3 time window (450–600 msec after search array onset). There was no interaction between task condition and electrode site [$F(1, 15) = 1.2, p = .3$].

DISCUSSION

We investigated whether and how the semantic processing and maintenance of verbal information impairs perception and attentional selectivity in visual search tasks which engage processes that are also involved in the control of driving. Participants searched for visual targets defined by unique features (Experiment 1) or feature conjunctions (Experiment 2) while listening to narrated text passages that had to be recalled later (encoding condition), or to backward-played speech signals (control condition). In both experiments, RTs to visual search targets were about 30 msec slower in the encoding relative to the control condition.⁴ These behavioral dual-task costs strongly suggest that the requirement to encode and maintain verbal stimuli diverted attention from the visual environment, and this conclusion was further supported by the ERP results. ERPs triggered in response to visual search arrays were systematically affected by a concurrent auditory task. Visual ERPs were more negative in the control as compared to the encoding condition. This differential effect started about 150 msec after search array onset, and remained present in a sustained fashion for about 150 msec, overlapping with the N1, P2, and N2 components. As can be seen in Figures 2 and 4, these ERP differences between the two task conditions were very similar across the two experiments in terms of their onset latency, duration, and posterior scalp distribution. The presence of a sustained negativity between 150 and 300 msec after stimulus onset is a common finding in ERP studies of visual attention (see Hillyard & Anllo-Vento, 1998 for a review). This selection negativity is typically found at posterior electrodes in response to currently task-relevant as compared to irrelevant visual stimuli, and is assumed to reflect the focal attentional processing of these stimuli. The sustained enhanced negativity observed for visual ERPs in the control relative to the encoding condition was very similar to the attentional selection negativities observed in previous studies, which strongly suggests that the focal attentional processing of visual search arrays was less efficient in the encoding condition, where participants had to perform a concurrent auditory task, than in the single-task control condition. The fact that these ERP differences emerged at relatively short poststimulus latencies (about 150 msec after visual stimulus onset) indicates that the presence of an auditory task can affect relatively early perceptual stages of visual processing.

Sustained modulations of posterior visual ERPs were equally present for target and nontarget visual search arrays, which demonstrates that this effect reflects a modula-

tion of visual–perceptual processing at stages that precede the detection and spatial selection of target stimuli. To investigate whether the attentional demands of the auditory task in the encoding condition also affect the spatial selection and processing of visual target events, N2pc and P3 components in response to visual search arrays that contained a target were compared across the two task conditions. In both experiments, N2pc amplitudes were reduced in amplitude in the encoding relative to the control condition. The N2pc is an electrophysiological marker of the spatial selection of target events in visual search tasks (e.g., Eimer, 1996; Luck & Hillyard, 1994). N2pc amplitudes reflect the difference in brain activity between visual areas contralateral and ipsilateral to a target, and therefore, provide a measure of the relative distribution of attention in the visual field: Large N2pc amplitudes indicate fully focused attention, whereas smaller N2pc components suggest a more diffuse attentional state (see also Kiss et al., 2007). The reduction of N2pc amplitudes in the encoding condition thus provides evidence that the efficiency of attentional target selection is adversely affected by the concurrent auditory task. The fact that the effects of task condition on N2pc amplitudes were relatively small in both experiments, and were not accompanied by corresponding N2pc onset latency differences, indicates that attentional target selection was only moderately affected by the auditory task, and suggests that participants attached sufficient attentional priority to the visual search task to successfully detect and select such targets.

Analyses of P3 amplitudes further supported the conclusion that the auditory task interfered with attentional target processing. In both experiments, P3 components to search targets were reduced in amplitude in the encoding as compared to the control condition. Modulations of P3 amplitudes in response to target events are assumed to reflect the degree to which attention is allocated to one specific task: Target P3s are typically larger when attention is focused exclusively on one task than when it is divided across different tasks (e.g., Sirevaag et al., 1989; Wickens et al., 1983). Thus, the observed P3 reduction on target trials in the encoding relative to the control condition provides additional electrophysiological evidence that the requirement to process and maintain auditory information in the encoding condition resulted in a diversion of attention from the primary visual search task, thereby reducing the efficiency of target-related processing. Similar effects of a concurrent auditory task on P3 amplitudes to visual target stimuli were reported by Strayer and Drews (2007), who compared a group that was actively engaged in a cell phone conversation and a single-task control group. In our encoding condition, participants only had to process and maintain auditory information while performing the visual search task. The fact that reduced P3 amplitudes to visual targets were still found in this condition suggests that an active and reciprocal interaction with a conversation partner may not be necessary to divert attention from a visual task.

The conclusion that the active processing of speech produces dual-task costs in visual attention tasks appears inconsistent with the findings of two previous behavioral experiments (Kunar et al., 2008; Strayer & Johnston, 2001). In these studies, target detection performance in visual tracking tasks was equally good when this task was performed while participants listened to spoken text that had to be maintained for subsequent recall and when it was performed in isolation. This suggests that the mere processing of spoken verbal material by itself does not impair attentional performance. According to these authors, such dual-task costs are produced only when an auditory task also includes the requirement to actively produce speech. The results of the present study indicate that this conclusion may be premature, as both behavioral and ERP measures were consistent in suggesting that visual-attentional processing was impaired when participants had to simultaneously process spoken auditory information, and no speech production was involved. One obvious difference between the present study and the experiments by Kunar et al. (2008) and Strayer and Johnston (2001) is the nature of the visual attention task. In these previous studies, participants performed continuous visual tracking tasks, which might have required a stronger and more sustained focus of visual attention than the visual search tasks employed in the present experiments, where search arrays were discrete events. Furthermore, the visual target events in the study of Strayer and Johnston were highly salient red flashes which might have been more efficient in attracting attention away from a concurrent auditory task than the target stimuli used in the present study. Future studies will have to investigate in more detail how auditory dual-task costs are affected by the features of specific visual-attentional tasks. In any case, the current results demonstrate that the need to process and maintain auditory information can, in itself, impair concurrent visual processing, even when speech production is not required. It is, however, likely that such dual-task deficits, and their ERP correlates, will be even more pronounced when an auditory task includes the active engagement in a reciprocal conversation (Strayer & Johnston, 2001). This possibility should be investigated in future studies.

The impairment of visual perception and selective attention found in the encoding condition where participants were engaged in the active encoding and maintenance of verbal information does not necessarily imply that merely listening to verbal material such as radio programs will affect visual processing, and thus, interfere with driving performance. In the present experiment, visual processing costs emerged when participants were instructed to continuously encode and maintain auditory information while searching for visual targets. Although these costs could be exclusively due to the demands imposed by listening and encoding, it is more likely that the additional requirement to maintain the spoken text for subsequent recall is the more important factor. Previous behavioral studies have shown that passive listening to radio programs does not

affect performance in simulated driving tasks such as pursuit tracking (e.g., Strayer & Johnston, 2001), suggesting that attentional dual-task costs of the sort observed in the present study are eliminated when observers are free to flexibly withdraw their attention from auditory messages in order to focus fully on a concurrent visual task (see also Drews et al., 2008). This could be demonstrated in a future experiment where ERP correlates of perceptual and attentional processing in visual search obtained in the encoding condition are directly compared to a condition where auditory stimulation is identical, but participants are not required to maintain spoken message for later recall.

In summary, the results of the current study provide new behavioral and electrophysiological evidence that the encoding and maintenance of verbal material can have adverse effects on the attentional processing of visual information and impair performance in concurrent visual search tasks. The ERP differences observed between the encoding and the control conditions suggest that these dual-task costs include both a general decrease in the efficiency of visual-perceptual processing stages that precede target detection and identification, as well as impairments in the attentional selection of target events. The fact that a very similar results was obtained in Experiment 1, where visual stimuli were free-field LEDs, and candidate target events were color singletons, and in Experiment 2, where visual stimuli were presented on a computer screen, and target detection required the conjunction of color and shape information indicates that such detrimental effects of a concurrent auditory task on attentive visual processing can be observed across different visual stimulation procedures, and different attentional task sets. Overall, the current findings provide new evidence for the proposal of Strayer and Drews (2007) and Strayer and Johnston (2001) that the deficits of driving-related performance associated with cell phone conversations are closely linked to the diversion of visual selective attention.

On a more general theoretical level, the observation that the active processing and maintenance of auditory information produced behavioral and electrophysiological costs for concurrent visual search tasks also has implications for psychological models of dual-task performance. If visual and auditory tasks engaged entirely separate modality-specific attentional resources, as proposed by the multiple-resource view (e.g., Wickens, 1984), there should have been little, if any, interference from auditory on visual processing in the encoding condition, with visual search performance and visual ERPs very similar to the control condition. The fact that systematic behavioral and ERP differences between these two conditions were observed is not consistent with the hypothesis that attentional resources are fully modality-specific. These results provide new evidence for the existence of cross-modal attentional links between audition and vision (see also Eimer & Driver, 2001; Spence & Driver, 1996), which can result in dual-task costs when attention has to be divided between auditory and visual tasks.

Acknowledgments

This research was supported by a grant from the Medical Research Council (MRC), UK. M. E. holds a Royal Society-Wolfson Research Merit Award. We thank Sue Nicholas for technical assistance, and David Strayer and an anonymous referee for helpful comments.

Reprint requests should be sent to Elena Gherri, Department of Psychology, City University, Northampton Square, London, EC1V 0HB, UK, or via e-mail: Elena.Gherri.1@city.ac.uk.

Notes

1. Although the difference waveforms in Figure 3 (top right) suggests that the onset of the N2pc might be earlier in the control relative to the encoding condition, analyses of N2pc onset latencies yielded no reliable differences between conditions.
2. For completeness, an additional analysis was also conducted for the N2pc in response to nontarget singletons (not shown in Figure 3). As expected, N2pc amplitudes were smaller than on-target trials, with N2pc mean amplitudes measured between 210 and 310 msec after search array onset of 0.7 μ V, as compared to 2.7 μ V in response to target singletons. However, main effects of contralaterality during the 210–250 msec and 255–310 msec time intervals [both $F(1, 15) = 18.4$, both $p < .001$] confirmed the presence of a reliable N2pc to nontarget singletons. There were no significant interactions between task condition and contralaterality for either time window [both $F(1, 15) < 1.8$, both $p > .2$].
3. The only exception was that no significant effect of task condition was found at lateral posterior electrodes in the N1 time window [160–200 msec; $F(1, 15) = 2.7$, $p = .12$].
4. It may seem surprising that RTs to targets were similar in both experiments, even though targets were color singletons in Experiment 1, whereas target identification required the conjunction of a specific color and shape in Experiment 2. This is due to the fact that in Experiment 1, singleton detection was, in itself, not sufficient for response selection: Only singletons with a gap were designated as targets, and the response to these targets was determined by their color.

REFERENCES

- Drews, F. A., Pasupathi, M., & Strayer, D. L. (2008). Passenger and cell phone conversation in simulated driving. *Journal of Experimental Psychology: Applied*, *14*, 392–400.
- Eimer, M. (1994). "Sensory gating" as a mechanism for visuospatial orienting: Electrophysiological evidence from trial-by-trial cuing experiments. *Perception & Psychophysics*, *55*, 667–675.
- Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*, 225–234.
- Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and endogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, *25*, 497–511.
- 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.
- Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 781–787.
- Kiss, M., Goolsby, B. A., Raymond, J. E., Shapiro, K. L., Silvert, L., Nobre, A. C., et al. (2007). Efficient attentional selection predicts distractor devaluation: ERP evidence for a direct link between attention and emotion. *Journal of Cognitive Neuroscience*, *19*, 1316–1322.
- Kunar, M. A., Carter, R., Cohen, M., & Horowitz, T. S. (2008). Telephone conversation impairs sustained visual attention via a central bottleneck. *Psychonomic Bulletin & Review*, *15*, 1135–1140.
- Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1000–1014.
- 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.
- Palin, M. (2003). *Around the world in 80 days* [Audiobook CD]. Written and narrated by Michael Palin. BBC Worldwide Ltd.
- Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155–189). Hove, UK: Psychology Press.
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography & Clinical Neurophysiology*, *72*, 184–187.
- Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular–telephone calls and motor vehicle collisions. *New England Journal of Medicine*, *336*, 453–458.
- Sirevaag, E., Kramer, A. F., Coles, M. G. H., & Donchin, E. (1989). Resource reciprocity: An event-related brain potentials analysis. *Acta Psychologica*, *70*, 77–97.
- Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1005–1030.
- Spence, C., & Driver, J. (1997). On measuring selective attention to an expected sensory modality. *Perception & Psychophysics*, *59*, 389–403.
- Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1298–1319.
- Spence, C., & Read, L. (2003). Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *Psychological Science*, *14*, 251–256.
- Strayer, D. L., & Drews, F. A. (2007). Cell-phone-induced driver distraction. *Current Directions in Psychological Science*, *16*, 128–131.
- Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, *9*, 23–32.
- Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*, 462–466.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). Orlando, FL: Academic Press.
- Wickens, C. D., Kramer, A., Vanasse, L., & Donchin, E. (1983). The performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information processing resources. *Science*, *221*, 1080–1082.
- Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature*, *400*, 867–869.