Time course of allocation of spatial attention by acoustic cues in non-human primates

Luis C Populin and **Abigail Z Rajala**
Depts of Anatomy and Psychology, Neuroscience Training Program, and Eye Research Institute, University of Wisconsin-Madison, Madison, WI 53706

Abstract

Spatial attention mediates the selection of information from different parts of space. When a brief cue is presented shortly before a target (cue to target onset asynchrony = CTOA) in the same location, behavioral responses are facilitated, a process called attention capture. At longer CTOAs responses to targets presented in the same location are inhibited, called inhibition of return (IOR). In the visual modality, these processes have been demonstrated in both humans and non-human primates, the latter allowing for the study of the underlying neural mechanisms. In audition, the effects of attention have only been shown in humans when the experimental task required sound localization. Studies in monkeys using similar cues but lacking a sound localization requirement have produced negative results.

We have studied the effects of predictive acoustic cues on the latency of gaze shifts to visual and auditory targets in monkeys experienced in localizing sound sources in the laboratory with the head unrestrained. Both attention capture and IOR were demonstrated using acoustic cues though on a faster time course compared to visual cues. Additionally, the effect was observed across sensory modalities (acoustic cue to visual target) suggesting that the underlying neural mechanisms of these effects may be mediated within the superior colliculus, a center where inputs from both vision and audition converge.

Keywords

spatial attention; auditory; gaze shifts; gaze latency; and reaction time

Introduction

Spatial attention is a process that mediates the selection of information from specific locations in space. Most of what is known about its effects on behavior and its underlying neural mechanisms has been learned from studies in which attention was directed by visual cues to locations in the surrounding space (Colby and Goldberg, 1999; Desimone and Duncan, 1995; Kinchla, 1992; Müeller and Findlay, 1988; Pashler, 1998; Posner, 1978). The sequence of events that characterizes the allocation of spatial attention by a visual signal, spatial visual attention, is well understood. Shortly after attention is captured (Jonides, 1981) by a visual event in the periphery, the processing of additional stimuli is facilitated at that location and hindered elsewhere (Richard et al., 2003; Wright and Richard, 2000). After attention is directed away from where it was originally captured, the processing of information from additional stimuli at the original location is hindered, a process called inhibition of return (IOR) (Posner and Cohen, 1984; Posner et al. 1985).

Luis C Populin, Dept of Anatomy, B 385 MSC, University of Wisconsin-Madison, Madison, WI 53706, Lpopulin@wisc.edu, Office: 608-265-6451, Fax: 608-262-7306.

Spatial attention can also be allocated to locations in the surrounding space by acoustic signals, but the resulting behavioral consequences and underlying neural mechanisms are not well understood as compared to the visual modality. Acoustic signals, unlike visual signals which are inherently spatial, may or may not convey spatial information depending on their bandwidth (Frens and van Opstal 1995). Furthermore, they may be processed in a non-spatial manner as in experimental tasks that require only the detection of an event or when they are presented diotically or dichotically over earphones. Studies of auditory attention in humans have used experimental tasks that had different requirements and, consequently, have produced contradictory results. Posner (1978), for instance, concluded that auditory detection was unaffected by spatial attention because it occurred before localization could take place whereas Buchtel and Butter (1988) concluded that the lack of effect of auditory attention was due to the auditory system's lack of a fovea or finger tip. Rhodes (1987), who demonstrated an effect of acoustic signals on orienting, suggested on the other hand that localization of sound sources might be necessary to reveal the effects of attention because completion of the task would require the use of auditory representations of space. This view has been substantiated by the results of Spence and Driver (1994), McDonald and Ward (1999), and Roberts et al. (2009). Acoustic cues have also been shown to alter the responses of humans to visual stimuli under various experimental conditions (McDonald et al. 2000; Schmitt et al. 2000; Spence and Driver, 1997).

Studies in non-human primates, on the other hand, have produced negative results. Bell et al. (2004) reported no effect of acoustic cues on the latency of saccadic eye movements of rhesus monkeys to visual targets; the same monkeys showed a significant effect of visual cues on saccade latency to visual targets. Some aspects of Bell et al.'s (2004) experimental task might explain their negative results in the acoustic cue condition. For example, the acoustic cue did not have behavioral significance suggesting the possibility that their subjects did not use them. In addition, their subjects were tested with their heads restrained, a condition known to affect the localization of sound sources (Populin, 2006).

Considering Rhodes' (1987) results and proposition and Spence and Driver's (1994) findings, Bell et al.'s (2004) negative results in the acoustic cueing condition are not totally unexpected. They are troubling, however, because as the animal model of choice for studies of the neural mechanisms underlying spatial attention, rhesus monkeys are expected to be able to process information from different sensory modalities. Accordingly, the present study sought to determine if non-human primates could covertly deploy spatial attention in response to predictive broadband acoustic cues. Evidence that spatial attention had been successfully allocated was expected in the form of changes in the latency of gaze shifts to acoustic and visual targets. Two monkeys with experience in localizing sound sources in a laboratory setting (Populin, 2006) served as subjects. The results unequivocally show that attention was allocated about the space in response to broadband acoustic cues. As in the visual modality, attention capture and IOR were observed but with a distinct, much shorter time course.

Materials and Methods

Subjects and surgery

Two adult male rhesus monkeys that had participated in a previous study of sound localization behavior (Populin, 2006) served as subjects. Each monkey was implanted with scleral search coils (Judge et al. 1980) to record gaze, and a titanium head post to which could be attached a spout to deliver water rewards after successful trials with the head unrestrained and which also served to restrain the head for care of the implant area. Further details about subject preparation and the experimental paradigm have been presented (Populin, 2006). All surgical and experimental procedures were approved by the University

of Wisconsin Animal Care and Use Committee and were in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals*.

Gaze movement measurements

Gaze position was measured with the scleral search coil technique (Robinson, 1963) using a phase angle system (CNC Engineering, Seattle, WA) calibrated to provide a linear output over a 270° range on the horizontal plane. The term gaze is commonly used to refer to eye position recorded under head unrestrained conditions. Head movement data were measured with a similar coil embedded in the acrylic that formed the head cap supporting the head-post. Those data were taken into account for analysis to determine if the animal had moved after the presentation of the cue before target onset. Horizontal and vertical gaze and head position signals were low-pass filtered at 250 Hz (Krohn-Hite, Brockton, MA) and digitally sampled at 500 Hz with Tucker Davis Technologies System 2 (TDT, Alachua, FL). Data acquisition was performed with custom software. The digitized gaze and head position signals were stored in a computer hard drive for off-line analysis.

Experimental tasks and experimental sessions

Central to this work was the cueing task, an audio-oculomotor adaptation (Bell et al. 2004) of Posner's task (Posner et al. 1973, Posner, 1978), shown schematically in Figure 1. First, a visual fixation event, a red light emitting diode (LED) was presented at the straight-ahead position. During fixation, an exogenous acoustic cue was presented from one of two locations (±30°,0°) at a variable time after the onset of the fixation LED. The subject was required to maintain fixation on the LED straight ahead until it was turned off and not to respond overtly to the cue. A (2°, 2°) electronic window was set around the fixation LED. Trials in which the subjects eyes or head moved outside the window during the fixation were immediately discarded.

At a variable time after the presentation of the cue (1-800 msec) the fixation LED was turned off and at the same time either a visual or a broadband acoustic target was presented. The time elapsed between the onset of the cue and the onset of the target, defined as the cue to target onset asynchrony (CTOA), was randomly varied across trials. The subject was expected to orient to the location of the target after the fixation LED was turned off.

The experimental sessions were organized as in Populin (2006). Trials of various types (fixation, standard-saccade, and cueing), involving visual and acoustic targets at 24 locations in the frontal hemifield, were presented in random order to minimize the chances that the subject would be able to anticipate the location or modality of the upcoming target. Cue trials constituted approximately 25% of the total number of trials in a session. In approximately 75% of trials of this type the cue predicted the location of the ensuing target: the valid cue condition. In the remaining 25% of cue trials the target was presented in the opposite hemifield: the invalid cue condition. One of the major disadvantages of this arrangement was that a small number of trials was obtained in every session. This was preferred, however, over experimental sessions with a larger proportion of cue trials, or even sessions with just cue trials, in order to maintain unequivocally the requirement that the subject produced accurate, not speedy responses.

It is important to note that the proportion of trials corresponding to each condition was approximate because trials were selected in random order for presentation and the duration of the experimental sessions was determined by the animals stopping work due to satiety; however, a minimum of 100 trials was collected for each condition. The selection of the predictive form of presentation of the cue was based on Spence and Driver's (1994) findings

in humans that the effects of auditory attention were larger and longer lasting if the location of the cue predicted the location of the target.

Stimulus presentation

Acoustic stimuli were generated with Tucker Davis Technologies System 3 (TDT, Alachua, FL) hardware and presented with Morel MDT 20-28 mm soft dome tweeters located 84 cm away from the head of the subject in the frontal hemifield. The acoustic cue consisted of a single 100 μsec click and the acoustic target of a 100 msec broadband (0.1-25 KHz) noise burst. Visual targets, also 100 msec in duration, were presented with red LEDs mounted in front of some of the speakers; they subtended visual angles of 0.2°.

Basis for selecting target locations and stimulus type—A major consideration for the selection of cue and target location was the manner in which gaze shifts are executed. Specifically, for visual targets of up to approximately 20° of eccentricity, the contribution of the head to the gaze shift is typically negligible (Freedman and Sparks, 1997), whereas for acoustic targets within such a range of eccentricities the head does reliably contribute (Biesiadecki and Populin, 2005). Thus, to be able to compare the latencies of equivalent motor actions, the ±30° eccentricity location was selected for cue trials. This ensured that gaze shifts to targets of both modalities, acoustic and visual, would include head movements.

Stimulus type—Broadband noise bursts were selected as targets because both subjects had demonstrated that they were able to orient overtly to the sources from which stimuli of this type were presented in the frontal hemifield (Populin, 2006).

For the selection of the acoustic cue, however, the major consideration was to avoid presenting stimuli in a configuration that could result in the perception of some of the auditory phenomena (echo suppression, localization dominance, and summing localization) known as the precedence effects (Blauer, 1983). The two subjects in this study had shown that they could perceive acoustic stimuli in phantom locations when presented with pairs of noise bursts with interstimulus intervals (0-1000 μsec) in the summing localization range (Populin, 2006). In addition, account was also taken of the fact that uncertainty about sound source location, as it is experienced at the echo threshold period of the precedence effect, can result in longer latencies (see e.g., Fig. 7 of Tollin and Yin, 2003).

Accordingly, a 100 μsec click was selected as the cue stimulus because it could be presented at short CTOAs before the noise target without temporal or spatial summation. No significant differences in localization accuracy were found between the 100 msec broadband noise stimulus, which was used as a target, and the single 100 μsec click stimulus, which was used as a cue, in control experiments. In addition, the differences between the click and the noise signals further ensured that precedence effect phenomena would not be perceived. Consistent with previous reports (Frens and Van Opstal, 1998), the data showed that the perception of acoustic and visual target locations was unaffected by the 100 μsec click cue.

Dependent variable and data analysis

The dependent variables were gaze latency and final gaze position. Gaze latency was defined as the time elapsed between the onset of the acoustic or visual target and the initiation of the gaze shift. The time of presentation of the target, available to the acquisition software in the form of a hardware signal from the logic circuit that powered the LEDs or the TDT system used to present the acoustic stimuli, was saved as a time stamp. The time of initiation of the gaze shift was determined offline with interactive graphics software written in Matlab (The MathWorks, Natick, MA) using a velocity criterion (Populin and Yin, 1998).

Briefly, a mean velocity baseline was computed separately for the horizontal and vertical channels representing gaze position from an epoch starting 100 msec before to 10 msec after the presentation of the target during which the eyes were expected to be stationary. The onset of a gaze shift, defined as the end of fixation, was the time at which the velocity signal exceeded two standard deviations of the mean baseline. Data from only the first gaze shift in a trial were included in the analysis. Trials in which the monkey moved his eyes or head during the CTOA period, irrespective of the magnitude of the movement (i.e., movements smaller than 2° that would not have resulted in rejection of the trial during data acquisition), were excluded from analysis. An example of a trial in which the subject moved both his eyes and head during the CTOA period is shown in Figure 2. The difference in time between the hardware signals corresponding to the onset of the target (visual or acoustic) and the cue was monitored to ensure that CTOAs of proper duration were presented. Final gaze position was also determined in interactive graphics software. It was defined as the position of the eye in space at the time the velocity of the eye returned to within two standard deviations of the mean baseline.

An ANOVA revealed no differences between the subjects ($F_{1,0.784} = 7.39$, $P = 0.276$) and between the two types of targets ($F_{1,1} = 1.4$, $P = 0.447$), therefore, the data were grouped accordingly. Subsequently, gaze latency data from the valid condition from each CTOA were converted to z-scores and used as a standard against which gaze latency of the invalid condition was compared. The conversion to z-scores was performed because of the large difference in the number of trials, 75% from the valid and 25% from the invalid condition, that resulted from the design of the experiments. The differences between the conditions were then tested against the test value of zero, indicating that there was no effect of attention, at each CTOA using a 2-tailed t-test. This analysis was done separately for each of the two experiments, with a level of significance set at p < 0.05.

Further considerations about the use of the saccade cueing task

The monkeys were rewarded for localizing the sources of sounds and visual stimuli accurately, not for responding quickly. A reward window of (8°, 8°) was used in trials involving acoustic targets and a (5°, 5°) window in trials involving visual targets. A discussion of the rationale used to determine the dimensions of the reward window for acoustic targets is found Populin and Yin (1998). Briefly, the dimensions selected for the reward window represent a compromise between a window that is too small, thus guiding the subject to a particular area of space to obtain rewards regardless of perceived sound location, and a window that is too large, thus rewarding lackadaisical performance. In short, rewards were delivered for responding accurately, not quickly!

Results

The goal of this study was to determine if rhesus monkeys could allocate attention to locations in the surrounding space guided by presentation of predictive broadband acoustic cues. Evidence that attention had been allocated was sought in the form of changes in the latency of gaze shifts directed to acoustic and visual targets presented after the acoustic cue at various CTOAs. Two experiments were conducted. Experiment 1 studied the effects of the acoustic cue at CTOAs ranging from 50 to 800 msec, similar to studies of visual spatial attention (e.g., Fecteau et al. 2004), whereas Experiment 2, which was prompted by unexpected results from Experiment 1, studied the effects of the acoustic cue at CTOAs ranging from 1 to 50 msec.

Central to accomplishing the goals of the study was the correct execution of the cueing task, which required no overt responses to the acoustic cue during the CTOA period. Figure 2 illustrates the typical oculomotor behavior of one of the subjects in the condition that

imposed the longest CTOA among those tested, 800 msec, highlighted with a gray background. The experimental task is shown schematically in Figure 2B; note that neither the timing nor the amplitude of the traces representing the stimuli are drawn to scale. In separate panels, horizontal gaze (Fig. 2A) and head (Fig. 2C) position, and the corresponding velocity profiles used for analysis (Fig. 2B and 2D), are plotted as a function of time and synchronized to the onset of the acoustic broadband noise target illustrated with a vertical broken line at time 0 msec. The vertical component of the movements was omitted for clarity. The horizontal position of the acoustic target (-30°,0°) is illustrated with the contour of a speaker (Fig. 2A).

Both monkey Shepard and monkey Glenn performed the cueing task without difficulty. Overt responses to the targets consisted of coordinated movements of the eyes and head. The head component of the response typically consisted of a single movement (Fig. 2C,D), whereas the initial eye component was in some instances followed by corrections (Fig. 2A,C). Only approximately 1 percent of cueing trials in a typical experimental session had to be discarded due movement of the eyes or head during the CTOA period. An example of a discarded trial is plotted in red among several successfully completed trials plotted in blue. All trials in which movement was detected were excluded from analysis ensuring, therefore, that any orienting that might have taken place as a result of the presentation of the cue was covert not overt.

Experiment 1

The effects of the single 100 μsec click cue presented at CTOAs ranging from 50 to 800 msec are illustrated in Figure 3. The average latency of gaze shifts to visual and acoustic targets from the valid (filled symbols) and invalid (hollow symbols) conditions are plotted as a function of CTOA. Data from the two subjects for the two types of targets, acoustic and visual, are presented together because an initial ANOVA reveal no significant differences. The average latency for each type of target changed from session to session due to factors such as the motivation of the subjects to orient for rewards, but consistently maintained the same relationship: latencies to the acoustic targets at (30°,0°) were slightly shorter than those to visual targets at the same eccentricities, but not significantly different.

At CTOAs ranging from 50-150 msec the valid cue resulted in shorter latencies, whereas the invalid cue resulted in longer latencies revealing facilitation, i.e., capture. Note, however, the overlap between the 95% confidence intervals indicating that the differences did not reach statistical significance (Fig. 3A). The pattern reversed at the 400 and 600 msec CTOAs with the valid condition resulting in longer latencies compared to the invalid. At 800 msec CTOA the acoustic cue did not affect gaze latency.

It is the differences between the latency from the invalid and valid conditions that should reveal if attention capture and IOR took place (Fecteau et al. 2004). These differences between gaze latency in the invalid minus the valid conditions are plotted in Figure 3B as a function of CTOA. Proper comparison, however, could not be performed on the raw latency data because of the large differences in number of trials; the valid condition comprised 75% of the data. Accordingly, data from the valid condition were transformed into z-scores and used as a standard against which gaze latency of the invalid condition was compared using a 2-tailed t-test with a level of significance of $p < 0.05$. The results of the statistical analysis are shown in Table 1.

At the 50, 100, and 150 msec CTOAs the differences between the invalid and valid conditions fell above the no difference line revealing attention capture (Fig. 3B). However, as shown in Table 1, the effect did not reach statistical significance for these three CTOAs. It must be noted, nevertheless, that the 50 msec CTOA barely missed reaching significance

(p < 0.5). At the 200-600 msec CTOAs the differences between the invalid and valid conditions fell below the no difference line revealing IOR, which as shown in Table 1 reached significance. Therefore, the IOR documented under the conditions of this experiment was a much more robust effect than the capture documented at the shorter CTOAs. As demonstrated by the data in Figure 3A, the cue was ineffective in modifying gaze latency at the 800 msec CTOA.

Experiment 2

The rationale for conducting this experiment was based on two observations from the previous. First, the significant IOR documented at the 200, 400, and 600 msec CTOAs in the previous experiment (Fig. 3, Table 1) indicated that attention had to have been captured earlier, given that for IOR to take place attention has to first be allocated to a specific spatial location. Second, because attention capture did not reach significance at the 100 and 150 msec CTOAs but had just missed significance at the shortest CTOA tested, 50 msec, it was hypothesized that attention capture by the acoustic cue could take place at shorter CTOAs than previously tested.

Accordingly, Experiment 2 studied the effects of the 100 μsec single acoustic click on gaze latency at CTOAs ranging from 1-50 msec. Across all CTOAs tested the latency of gaze shifts to both visual and acoustic targets was consistently longer in the invalid cue condition compared to the valid (Fig. 4A).

As for the previous experiment, the difference between the latencies measured in the invalid and valid conditions were computed and plotted as a function of CTOA (Fig. 4B). All data points fell above the no difference line indicating that facilitation took place at all CTOAs tested. The same statistical analysis used in Experiment 1 revealed that the differences between the invalid and valid conditions were significant (Table 2); no data were recorded for the 20 and 40 msec CTOAs. Thus, the results supported the hypothesis that attention capture does take place in the auditory modality, but at shorter CTOAs than those documented with visual signals. It must be noted that although the results were not identical to those of Experiment 1 at the 50 msec CTOA, they were nevertheless consistent.

Accuracy

Accuracy is another component of the gaze response that could have been affected by processing the acoustic cue, which conveyed either valid or invalid information about the position of the ensuing target. Figure 5 shows average data from the two subjects in the 30 and 400 msec CTOAs. Data from these two CTOAs, one showing facilitation (Fig. 4 and Table 2) and the other showing IOR (Fig. 3 and Table 1), are representative of all CTOAs tested; the standard bars represent 95% confidence intervals.

Final gaze position from the no-cue condition are included for illustration (open round symbols). The data show no significant differences in final gaze position between the valid and invalid conditions, demonstrating, therefore, that the information provided by the cue affected only the timing of execution of the gaze shifts and not the perception of target location.

Discussion

The data demonstrate that non-human primates are able to allocate attention covertly about the surrounding space in response to broadband acoustic cues. Gaze shifts were facilitated at short CTOAs indicating capture and inhibited at longer CTOAs indicating IOR. In addition, the acoustic cue was equally effective in changing the latency of gaze shifts to acoustic and visual targets.

While similar to the results of some human studies that demonstrated capture (McDonald and Ward, 1999; Spence and Driver, 1994) and IOR (McDonald and Ward, 1999; Reuter-Lorenz and Rosenquist, 1996; Tassinari et al. 2002) in response to acoustic cues, the present results also stand in contrast to those of a similar study in monkeys by Bell et al. (2004) who found that an acoustic cue had no effect on saccadic latency. The lack of agreement in the results of these studies may reflect either fundamental differences between species, a potentially troubling finding being that macaque monkeys are the animal model closest to humans available for behavioral and electrophysiological studies, or differences in methodology, the understanding of which could shed much needed light for the design of future physiological studies of the underlying neural mechanisms. The present behavioral results indicate the latter.

The acoustic cue: predictive vs non-predictive

To start, it is important to consider the extent to which the cue predicts the location of the ensuing target. The non-predictive configuration of the cue is thought to trigger an automatic, bottom up process that results in the allocation of attention to the cued location. The predictive configuration of the cue, on the other hand, incorporates the effects of the subject's expectations, knowledge that an event is likely to occur at the cued location, to the sensory triggered aspect of this process.

The predictive configuration of the cue was used in the present study whereas Bell et al. (2004) used the non-predictive. It is, therefore, possible that the difference in the results of the two monkey studies could be simply explained by the predictive nature of the cue presented. Based on data from human studies, it is clear that this is not the case. First, pilot work carried out before this study suggested that both types of acoustic cues, predictive and non-predictive, could lead to the allocation of spatial attention. Second, Spence and Driver (1994) showed that both predictive and non-predictive acoustic cues were effective in driving the allocation of spatial attention, albeit the effects of the predictive configuration were larger and longer lasting. Furthermore, McDonald and Ward (1999) demonstrated both facilitation and IOR using non-predictive acoustic cues. Therefore, the predictive or non-predictive nature of the cue cannot account for the difference in the results between Bell et al. (2004) and the present study.

Time course of allocation and potential mechanisms

The time course of the effects of the acoustic cue documented in this study is consistent with the time course of allocation of spatial attention with visual cues in both humans (Posner, 1978) and monkey studies (Bell et al. 2004; Fecteau et al. 2004). At short CTOAs the latencies of gaze shifts directed at visual and acoustic targets were shorter in the valid condition suggesting attention capture (Jonides, 1981), whereas at longer CTOAs the latencies of gaze shifts to both types of targets were longer in the valid condition suggesting IOR (Posner and Cohen, 1984).

Attention capture—Distinctively, behavioral facilitation in this study was documented at CTOAs ranging from 1-50 msec, revealing a much faster and more transient process than the equivalent documented in humans. Spence and Driver (1994) and McDonald and Ward (1999), reported behavioral facilitation by acoustic cues at CTOAs ranging from 100-1000 msec and 100-500 msec, respectively. These differences, spanning several hundred milliseconds are difficult to explain in terms of a single neural mechanism and are likely the result of the very different requirements of the experimental tasks used in different studies. Spence and Driver (1994), for instance, used a manual response based on orthogonal spatial mapping whereas McDonald and Ward (1999) used a single key to measure reaction time.

Attention capture is thought to be a sensory-driven process (Fecteau et al. 2004) that occurs independently of the intentions of the observer (Remington et al. 1992), although it can be modulated by other variables (Fecteau et al. 2004; Santangelo and Spence, 2008). The short duration of the cue to target intervals and the effectiveness of the cue to alter the latency of the gaze shifts to targets of both modalities, acoustic and visual, raise the question of whether the behavioral facilitation was the result attention allocation or bimodal integration. The answer to this complex question cannot be clearly dictated by an arbitrary rule of minimum separation between the presentation of the cue and the target, e.g., 100 msec (Macaluso et al. 2001; McDonald et al. 2001), particularly when the argument is based on neuronal latencies recorded in the superior colliculus (SC) of anesthetized animals, which are significantly altered by the effects of the drugs (Populin, 2005).

Physiological recordings from the SC of behaving monkeys indicate that behavioral facilitation driven by a visual cue is associated with stronger sensory responses to the visual target brought about by summation with the responses evoked by the valid visual cue. Consistent with this argument, Bell et al. (2004) reported that responses evoked by the presentation of the acoustic cue were too weak and transient to influence the sensory responses evoked by the presentation of the ensuing target.

Although it is possible that Bell et al. (2004) did not observe significant facilitation because the CTOAs they studied were too long for the sensory responses evoked by the cue to sum with the sensory responses evoked by the target (in this study significant facilitation was obtained at CTOAs ranging from 1-50 msec), this explanation is unlikely because they did not observe IOR at longer CTOAs tested. Clearly, for attention to have been captured in the present study, proper sensory responses had to have been produced. The question that remains, however, is how?

Inhibition of return—The existence of inhibition of return in the auditory modality has been the subject of much controversy (Reuter-Lorenz and Rosenquist, 1996; Spence et al. 2000). It has been demonstrated in some human studies by requiring subjects to perform eye movements to the location of the cue (Reuter-Lorenz and Rosenquist, 1996) or by presenting re-orienting stimuli at the straight ahead location (Spence and Driver, 1998). Surprisingly, a strong inhibitory effect lasting hundreds of milliseconds was readily observed in the present study while requiring subjects not to move in response to the cue during the CTOA period. Complementing the facilitation observed at short CTOAs, this inhibition demonstrates that the subjects in this study allocated attention about the space in response to the acoustic cue.

The similarity of the effects between the auditory and visual target conditions, in terms of both magnitude and time course, lends supports the hypothesis that IOR is the result of a generalized mechanism (Spence et al. 2000). Because subjects were required to respond to the presentation of the target with a gaze shift, the present results also lend support to the hypothesis that the mechanism underlying IOR involves the oculomotor system (Rafal et al. 1989; Tassinari et al. 1989).

The SC, a midbrain structure known to play a fundamental role in the generation of gaze shifts and saccadic eye movements (Sparks, 1999; Wurtz and Optican, 1994) is thought to be intimately involved in the generation of IOR, as revealed by neuropsychological studies in patients affected by supra nuclear palsy, a condition that eventually results in the loss of the ability to make saccadic eye movements (Posner et al. 1985), and single unit recordings in behaving monkeys that revealed that IOR is associated with a weaker neural response evoked by the presentation of the target (Bell et al. 2004, Fecteau et al. 2004). Thus, a clear correlate was observed between IOR and neural responses in the SC for visual cues.

Whether the mechanism underlying IOR is in the SC or in some of its inputs is not entirely clear. The fact that similar reductions in the magnitude of visual responses have been documented in earlier stages of the visual pathway such as the superficial layers of the SC (Dorris et al. 2002; Robinson and Kertzman, 1995) and parietal cortex (Robinson et al. 1995), and that electrical stimulation applied to the SC instead of the presentation of a visual target revealed facilitation not inhibition, argues against the hypothesis that the mechanism actually resides in the SC. On the other hand, the present results support the hypothesis that IOR is generated in the SC as suggested by Posner et al. (1985) because convergence of inputs from different sensory modalities takes place here (Huerta and Harting, 1984) and not in earlier visual structures.

In the auditory modality Bell et al. (2004) also observed a correlation between behavior and the properties of single unit recordings in the SC, albeit of a negative nature. In the same subjects in which they demonstrated a correlation between changes in saccade latency and the magnitude of single unit responses to visual targets due to the presentation of visual cues, they did not observe any effects of acoustic cues on gaze latency or neural responses to the visual targets. They concluded that the neural responses to the acoustic cue were too small and transient to interact with the sensory responses evoked by the presentation of the ensuing target. The question remains, therefore, as to what specific aspect of the experimental conditions of the present study facilitated the occurrence of sensory responses to the acoustic cue strong and long lasting enough to interact with the responses to the ensuing target.

Sound localization and behavioral relevance

The results of studies in humans indicate that sound localization requirements, as first postulated by Rhodes (1987), are essential to reveal the effects of attention allocation in the auditory modality (e.g., McDonald and Ward, 1999; Spence and Driver, 1994). The basis for this hypothesis is the manner in which auditory information is encoded in the periphery. Unlike in vision, which inherently encodes spatial information at the periphery because of the two-dimensional arrangement of the photoreceptors on the retina, in audition the location of sound sources must be computed at a later stage because the hair cells are arranged linearly according to frequency along the cochlea. The location of sound sources is computed from interaural differences of time and level (Yin and Kuwada, 1984) with the results of such computations integrated as representations of space at higher levels, such as the map of auditory space of the SC. Rhodes (1987) suggested that the use of such representations, achieved by having to localize the sources of sounds, is required to reveal the effects of attention.

The behavioral data from humans are clear. Experimental tasks that require some form of sound localization, which attributes relevance to the location of acoustic stimuli, reveal the effectiveness of acoustic cues to drive the process of attention allocation (McDonald and Ward, 1999; Spence and Driver, 1994). Conversely, experimental tasks that can be completed without having to localize the sources of sounds, requiring simple detection for example, do not seem to be affected by spatial attention because detection of an acoustic stimulus, according to Posner (1978), takes place before sound localization. Regardless of the hypothesized functional order of events, it is clear that auditory information can be used distinctively in a spatial or non-spatial manner.

Behavioral data from monkeys, on the other hand, show no effects of the acoustic cue on saccade latency (Bell et al. 2004). Based on the results from humans studies that used a similar acoustic cue (e.g., McDonald and Ward, 1999; Spence and Driver, 1994), it is reasonable to hypothesize that the negative result of Bell et al. (2004) could be explained by a lack of sound localization requirements in their experimental task. Indeed, inspection of

the behavioral requirements imposed by their version of the acoustic cue reveals that the task of orienting to the visual target could have been performed successfully without regard for the acoustic cue. However, this cannot be the explanation for the difference in the results because the experimental task in the present study could also be completed successfully without regard for the acoustic cue.

It could be argued that the subjects of this study had experience in localizing sound sources in the laboratory setting, which indeed they had. At the time this study was conducted each subject had performed at least 300,000 gaze shifts to acoustic targets. However, this explanation is also unlikely because the subjects had never been exposed to the single 100 μsec click stimulus that was used as a cue. It is also unlikely that a cognitive function such as the allocation of covert attention about the space in response to a sound could be developed *de novo* by training over a relatively short time in the life of adult animals. The present study, however, included a condition in which the target was acoustic, thus in those trials the subjects needed to localize the source of the sounds to obtain a reward. It is likely, therefore, that the subjects treated all sounds as potential targets for gaze shifts, the accurate execution of which could lead to reward. Thus, in the context of the cueing task, as used in this study, the localization requirement was achieved by including a condition with acoustic targets and attributing behavioral significance to their localization by the administration of rewards. In addition, it must be noted that the subjects were allowed to orient with their heads unrestrained, which is necessary for accurate sound localization (Populin, 2006).

Summary

Future physiological studies aimed at unraveling the neural mechanisms underlying the effects documented in this study will have to reconcile two apparently conflicting sets of facts. On the one hand, it is clear from single unit recordings in the SC using visual cues that interactions between sensory responses evoked by the cue and target are essential for behavioral effects - facilitation and IOR - to take place (Bell et al. 2004; Fecteau et al. 2004). On the other hand, data from humans (Spence and Driver, 1994; McDonald and Ward, 1999) and this study indicate that for the effects of spatial attention to be revealed in the auditory modality, the experimental task must require some aspect of sound localization. It is, therefore, hypothesized that the properties of auditory responses in the SC, their magnitude in particular, should be a function of the behavioral requirements of the experimental task. Accordingly, single units recordings from monkeys performing a cueing task with sound localization requirements should reveal cue-target interactions similar to those demonstrated with visual cues.

Acknowledgments

References

Bell AH, Fecteau JH, Munoz DP. Using auditory and visual stimuli to investigate the behavioral and neuronal consequences of reflexive covert orienting. J Neurophysiol. 2004; 91:2172–2184. [PubMed: 14702335]

Biesiadecki MG, Populin LC. Effects of target modality on primate gaze shifts. Soc Neurosci Abstr. 2005; 35:858.15.

Blauer, J. Spatial hearing: the psychophysics of human sound localization. MIT Press; Cambridge, MA: 1983.

Buchtel HA, Butter CM. Spatial attentional shifts: implications for the role of polysensory mechanisms. Neuropsychologia. 1988; 26:499–509. [PubMed: 3405396]

Colby CL, Goldberg ME. Space and attention in parietal cortex. Annu Rev Neurosci. 1999; 22:319–349. [PubMed: 10202542]

Desimone R, Duncan J. Neural mechanisms of selective visual attention. Annu Rev Neurosci. 1995; 18:193–222. [PubMed: 7605061]

Dorris MC, Kein RM, Everling S, Munoz DP. Contribution of the primate superior colliculus to inhibition of return. J Cog Neurosci. 2002; 14:1256–1263.

Fecteau JH, Bell AH, Munoz DP. Neural correlates of the automatic and goal-driven biases in orienting spatial attention. J Neurophysiol. 2004; 92:1728–1737. [PubMed: 15115792]

Freedman EG, Sparks DL. Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys. J Neurophysiol. 1997; 77:2328–2348. [PubMed: 9163361]

Frens MA, Van Opstal AJ. A quantitative study of auditory-evoked saccadic eye movements in two dimensions. Exp Brain Res. 1995; 107:103–117. [PubMed: 8751068]

Frens MA, Van Opstal AJ. Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. Brain Res Bull. 1998; 46:211–224. [PubMed: 9667814]

Huerta, MF.; Harting, JK. The mammalian superior colliculus: studies of its morphology and connections. In: Vanegas, H., editor. Comparative neurology of the optic tectum. Plenum; New York: 1984. p. 687-773.

Jonides, J. Voluntary vs automatic control over the mind's eye's movement. In: Long, JB.; Baddeley, AD., editors. Attention and performance IX. Erlbaum; Hillsdale, NJ: 1981. p. 187-203.

Judge SJ, Richmond BJ, Shu FC. Implantation of magnetic search coils for measurement of eye position: an improved method. Vis Res. 1980; 20:535–537. [PubMed: 6776685]

Kinchla RA. Attention. Annu Rev Psychol. 1992; 43:711–742. [PubMed: 1539951]

Macaluso E, Frith C, Driver J. Multisensory integration and crossmodal attention effects in the human brain. Science. 2001; 292:1791. [PubMed: 11397913]

McDonald JJ, Ward LM. Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. J Exp Psychol Hum Percept Perform. 1999; 25:1234–1252.

McDonald JJ, Teder-Salejarvi WA, Hillyard SA. Involuntary orienting to sound improves visual perception. Nature. 2000; 407:906–908. [PubMed: 11057669]

McDonald JJ, Teder-Salejarvi WA, Ward LM. Multisensory integration and crossmodal attention effects in the human brain. Science Science. 2001; 292:1791.

Müeller HJ, Findlay JM. The effect of peripheral discrimination thresholds in single and multiple element displays. Acta Psychol. 1988; 69:129–155.

Pashler, H. The psychology of attention. MIT Press; Cambridge, MA: 1998.

Populin LC. Anesthetics change the excitation/inhibition balance that governs sensory processing in the cat superior colliculus. J Neurosci. 2005; 25:5903–5914. [PubMed: 15976079]

Populin LC. Monkey sound localization: head-restrained vs head-unrestrained orienting. J Neurosci. 2006; 26:9820–9832. [PubMed: 16988053]

Populin LC, Yin TCT. Behavioral studies of sound localization in the cat. J Neurosci. 1998; 18:2147–2160. [PubMed: 9482800]

Posner, MI. Chronometric explorations of mind. Erlbaum; Hillsdale, NJ: 1978.

Posner, MI.; Cohen, Y. Components of visual orienting. In: Bouma, H.; Bouwhuis, D., editors. Attention and performance X. Lawrence Erlabum; London: 1984. p. 531-556.

Posner MI, Klein R, Summers J, Buggie S. On the selection of signals. Mem and Cognition. 1973; 1:2–12.

Posner MI, Rafal RD, Choate LS, Vaughan J. Inhibition of return: Neural basis and function. Cogn Neuropsychol. 1985; 2:211–228.

Rafal R, Calabresi P, Brennan CW, Sciolto TK. Saccade preparation inhibits reorienting to recently attended locations. J Exp Psychol Hum Percept Perform. 1989; 15:673–685. [PubMed: 2531204]

Remington RW, Johnston JC, Yantis S. Involunary attentional capture by abrupt onsets. Percept Psychophys. 1992; 51:279–290. [PubMed: 1561053]

Reuter-Lorenz PA, Rosenquist JN. Auditory cues and inhibition of return: the importance of oculomotor activation. Exp Brain Res. 1996; 112:119–126. [PubMed: 8951414]

Rhodes G. Auditory attention and the representation of spatial information. Percept Psychophys. 1987; 42:1–14. [PubMed: 3658631]

Richard CM, Wright RD, Ward LM. Goal-driven modulation of stimulus-driven attentional capture in multiple-cue displays. Percept Psychophys. 2003; 65:939–955. [PubMed: 14528901]

Roberts KL, Summerfield AQ, Hall DA. Covert auditory spatial orienting: an evaluation of the spatial relevance hypothesis. J Exp Psychol: Hum Percept Perform. 2009; 35:1178–1191. [PubMed: 19653757]

Robinson DA. A method of measuring eye movement using a sclera search coil in a magnetic field. IEEE Biomed Eng. 1963; 10:137–145.

Robinson DL, Kertzman C. Covert orienting of attention in macaques. III. Contributions of the superior colliculus. J Neurophysiol. 1995; 74:713–721. [PubMed: 7472376]

Robinson DL, Bowman EM, Kertzman C. Covert orienting of attention in macaques. II. Contributions of the parietal cortex. J Neurophysiol. 1995; 74:698–712. [PubMed: 7472375]

Santangelo V, Spence C. Is exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. Conscious Cog. 2008; 17:989–1015.

Schmitt M, Postma A, De Haan E. Interactions between exogenous auditory and visual spatial attention. Quart J Exp Psychol. 2000; 53A:105–130.

Sparks DL. Conceptual issues related to the role of the superior colliculus in the control of gaze. Curr, Opin Neurobiol. 1999; 9:698–707. [PubMed: 10607648]

Spence CJ, Driver J. Covert spatial orienting in audition: exogenous and endogenous mechanisms. J Exp Psychol Hum Percept Perform. 1994; 20:555–574.

Spence CJ, Driver J. Audiovisual links in exogenous covert spatial orienting. Percept Psychophys. 1997; 59:1–22. [PubMed: 9038403]

Spence CJ, Driver J. Inhibition of return following an auditory cue. The role of central reorienting events. Exp Brain Res. 1998; 118:352–360. [PubMed: 9497142]

Spence C, Lloyd D, McGlone F, Nicholls MER, Driver J. Inhibition of return is supramodal: a demonstration between all possible pairings of vision, touch, and audition. Exp Brain Res. 2000; 134:42–48. [PubMed: 11026724]

Tassinari G, Biscaldi M, Marzi CA, Berlucchi G. Ipsilateral inhibition and contralateral facilitation of simple reaction time to non-foveal visual targets from non-informative visual cues. Acta Psychol. 1989; 70:267–291.

Tassinari G, Campara D, Benedetti C, Berlucchi G. The contribution of general and specific motor inhibitory sets to the so-called auditory inhibition of return. Exp Brain Res. 2002; 146:523–530. [PubMed: 12355281]

Tollin DJ, Yin TCT. Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect. J Neurophysiol. 2003; 90:2149–2162. [PubMed: 12801901]

Wright RD, Richard CM. Location cue validity affects inhibition ofreturn of visual processing. Vision Res. 2000; 40:2351–2358. [PubMed: 10927120]

Wurtz RH, Optican LM. Superior colliculus cell types and models of saccade generation. Curr Opin Neurobiol. 1994; 4:857–861. [PubMed: 7888769]

Yin, TCT.; Kuwada, S. Neuronal mechanisms of binaural interaction. In: Edelman, GM.; Gall, WE.; Cowan, WM., editors. Dynamic aspects of neocortical function. John Wiley and Sons; New York: 1984. p. 263-313.
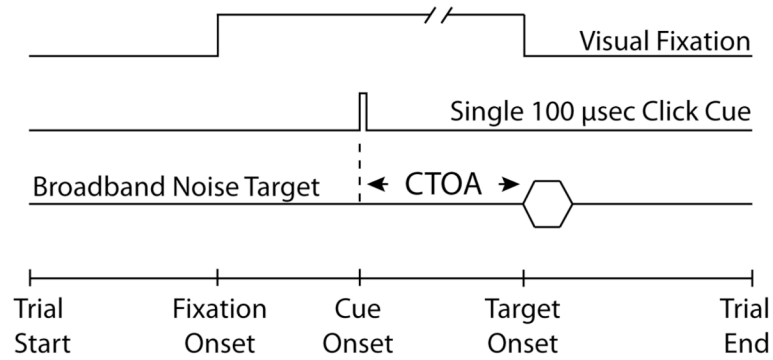
Figure 1.
Schematic representation of the cueing task. A fixation LED was presented at the straight-ahead position for a variable time (1000-1700 msec). The monkey was expected to look at the stimulus and to maintain fixation until it was extinguished. During the period of fixation a 100 μsec click cue was presented from transducers located at (±30°,0°). In the valid condition, which comprised ~75% of trials of this type, the cue predicted the location of the ensuing target, whereas in the invalid condition, which comprised the remaining 25% of trials of this type, it did not. The time between the onset of the cue and onset of the target, the cue to target onset asynchrony (CTOA), was randomly varied between 1-800 msec. Successful completion of a trial required accurate orienting, not speedy responses.
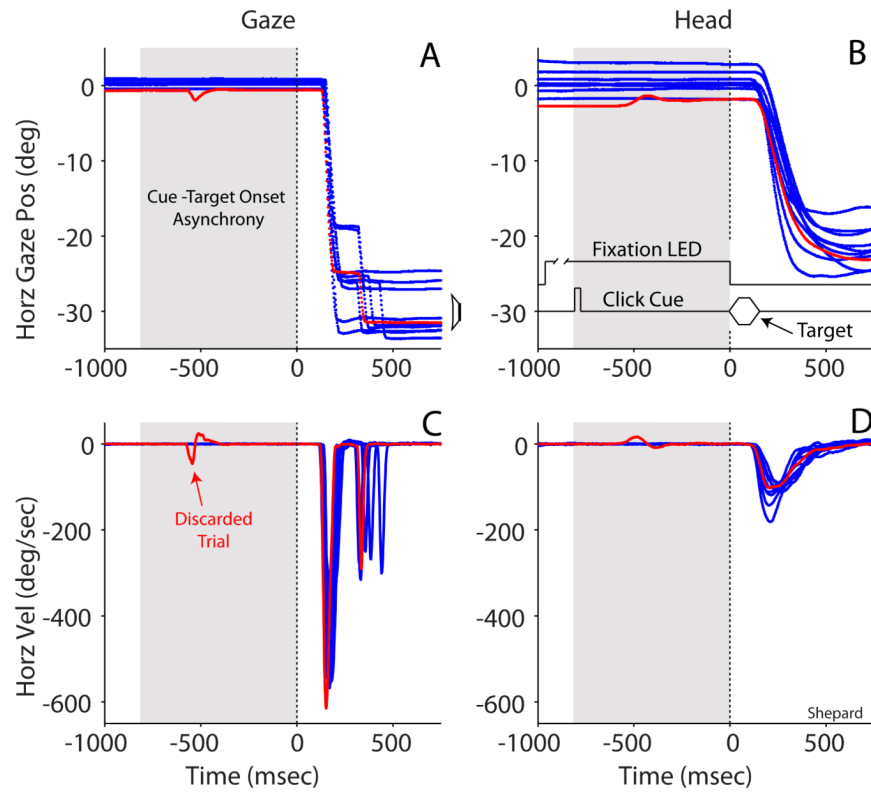
Figure 2.
Typical oculomotor behavior in the cueing task illustrated with data from the 800 msec cue to target onset asynchrony (CTOA) valid cue condition. Horizontal gaze (A) and head (B) position and their corresponding first derivates used for analysis (C and D) are plotted as a function of time and synchronized to the onset of the broadband noise target at time 0 msec. The gray background illustrates the CTOA period. Traces representing the vertical component of the movements are omitted for clarity. The horizontal position of the speaker, which was located at (30°,0°), is illustrated with a drawing to the right of the gaze plot (A). Traces plotted in blue represent successful trials. A single trial depicting movement during the CTOA period is plotted in red. The experimental task is shown in schematic form in (B). The acoustic cue was a 100 μsec click and the acoustic target a 100 msec broadband noise burst (10 msec rise/fall); the duration of these events is not accurately drawn to scale.
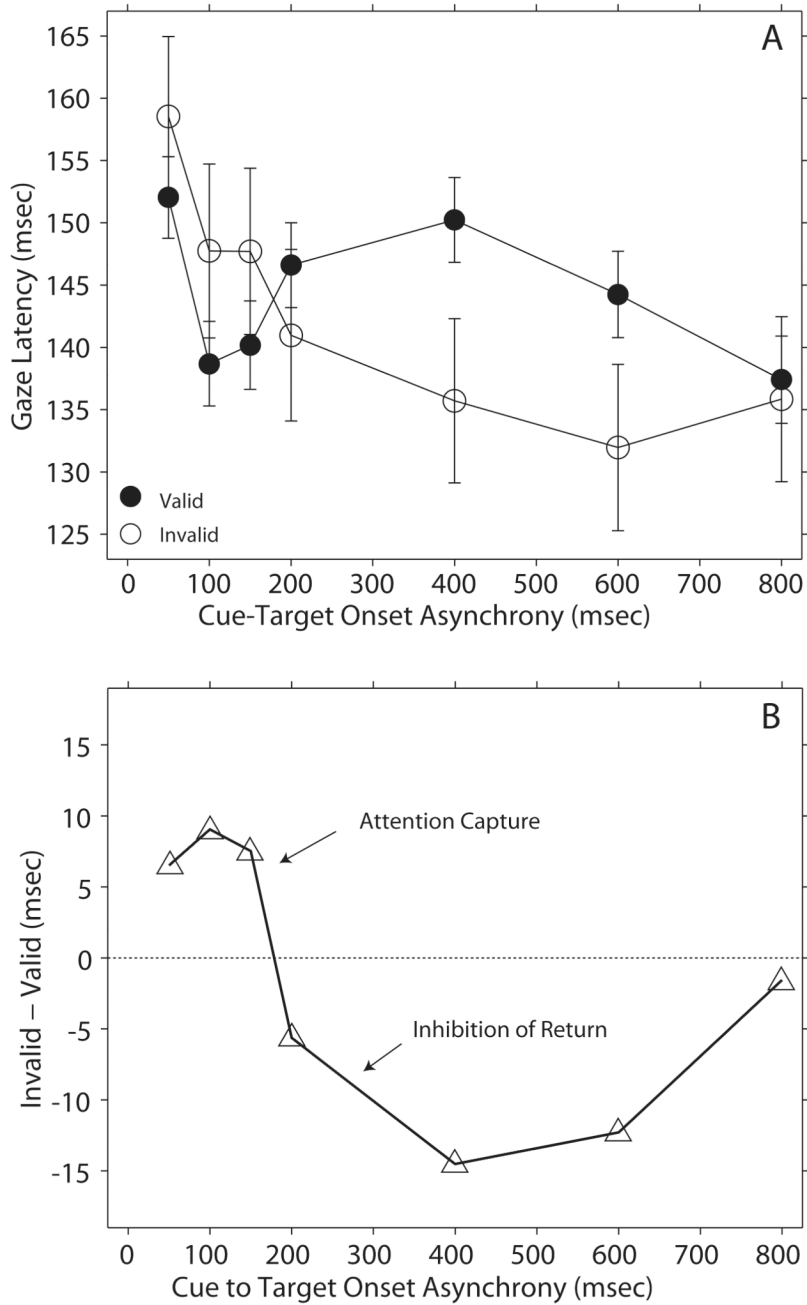
Figure 3.
Effect of the acoustic cue on gaze latency - Experiment 1. (A) Average data from gaze shifts to acoustic and visual targets from two subjects plotted as a function of cue to target onset asynchrony (CTOA). The errors bars represent 95% confidence intervals. (B) Difference in latency between the invalid and valid conditions plotted as a function of CTOA.
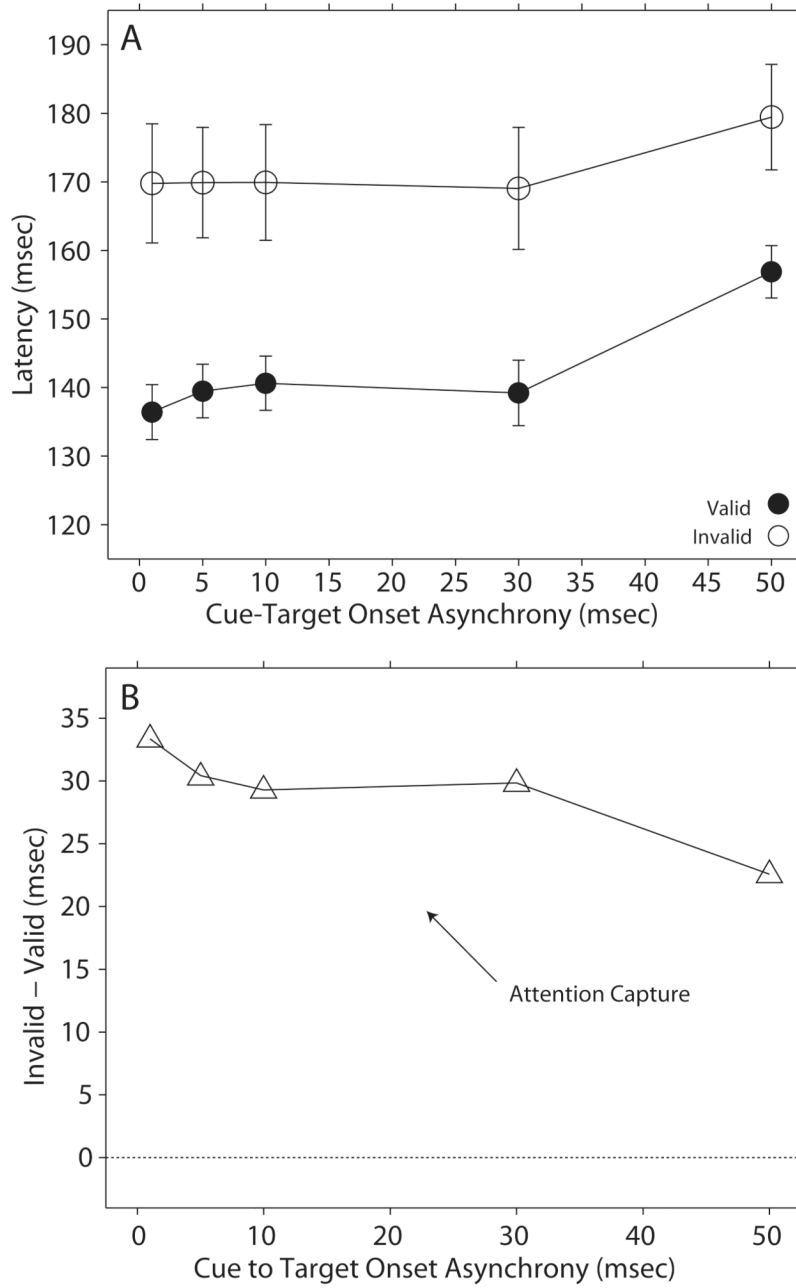
Figure 4.
Effect of the acoustic cue on gaze latency - Experiment 2. The data in this figure are presented as in Figure 3.
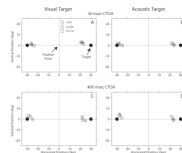
Figure 5.
Average final gaze position from the two subjects of the study recorded in the 30 and 400 msec cue to target onset asynchronies. Data from the no cue, valid, and invalid conditions are plotted separately for visual and acoustic targets. The standard bars represent 95% confidence intervals. The fixation point was located at the straight ahead position (0,0).

Table 1

T-test results from Experiment 1

CTOA (msec)	df	t value	Significance (2-tailed test)
50	76	1.9	0.062
100	70	1.14	0.257
150	70	1.32	0.191
200	68	-2.13	0.036
400	72	-6.20	0.000
600	69	-4.31	0.000
800	72	-0.58	0.561

Table 2

T-test results from Experiment 2

CTOA (msec)	df	t value	Significance (2-tailed test)
1	45	5.035	0.000
5	64	4.816	0.000
10	49	4.490	0.000
20	--	--	--
30	62	5.346	0.000
40	--	--	--
50	60	2.485	0.016