

Spatial and temporal factors determine auditory–visual interactions in human saccadic eye movements

M. A. FRENS, A. J. VAN OPSTAL, and R. F. VAN DER WILLIGEN
University of Nijmegen, Nijmegen, The Netherlands

In this paper, we show that human saccadic eye movements toward a visual target are generated with a reduced latency when this target is spatially and temporally aligned with an irrelevant auditory nontarget. This effect gradually disappears if the temporal and/or spatial alignment of the visual and auditory stimuli are changed. When subjects are able to accurately localize the auditory stimulus in two dimensions, the spatial dependence of the reduction in latency depends on the actual radial distance between the auditory and the visual stimulus. If, however, only the azimuth of the sound source can be determined by the subjects, the horizontal target separation determines the strength of the interaction. Neither saccade accuracy nor saccade kinematics were affected in these paradigms. We propose that, in addition to an aspecific warning signal, the reduction of saccadic latency is due to interactions that take place at a multimodal stage of saccade programming, where the *perceived* positions of visual and auditory stimuli are represented in a common frame of reference. This hypothesis is in agreement with our finding that the saccades often are initially directed to the average position of the visual and the auditory target, provided that their spatial separation is not too large. Striking similarities with electrophysiological findings on multisensory interactions in the deep layers of the midbrain superior colliculus are discussed.

Humans, as well as other animals, are equipped with various specialized senses that provide them with information about their environment. Several of these sensory systems represent the spatial location of an object on the basis of the received sensory input. This information about stimulus location can already be present at the level of the sensory organ, as is the case in the visual and somatosensory systems, or it can be neurally derived on the basis of indirect cues, as in the auditory system. Many of the objects that surround an organism, however, provide it with sensory information through various modalities at the same time.

In the literature, there is accumulating evidence that multimodal information about an object's location can lead to a reduction of the response latency and to an improvement of localization accuracy. For example, it has been shown that a motor response toward a visual target can be made with a shorter latency when this target is accompanied by an auditory signal at the same location. Simon and Craft (1970) have investigated this effect for arm movements, and Lee, Chung, Kim, and Park (1991) report qualitatively similar findings for saccadic eye move-

ments. In both studies, this effect was not present when the visual and the auditory stimulus were presented at opposite sides of a central fixation point. Perrott, Saberi, Brown, and Strybel (1990) showed that the time to foveate and identify one of two visual symbols was markedly reduced when an auditory costimulus was spatially aligned with the visual target, not only if the targets were presented in the far periphery but even if the stimuli were presented within the subject's parafoveal visual field.

Stein, Hunneycutt, and Meredith (1988) reported that, under near-threshold conditions, cats are able to localize combined audiovisual targets more accurately than visual targets. In their study, cats were trained to make whole body movements toward dimly lit visual targets, thereby learning that the presence of an auditory stimulus was irrelevant. Nevertheless, the animals performed better when an audiovisual cue was presented in spatial alignment. Performance dropped dramatically when the auditory and the visual stimulus were spatially disparate.

It should be noted that such auditory–visual interactions pose far from trivial problems to the nervous system, since the different sense organs initially encode the outside world in very different ways. First, the visual world is encoded *retinocentrically*, whereas auditory cues are represented with respect to the pinnae. For humans, this results in a *craniocentric* code, since the pinnae are immobile with respect to the head. Second, the retina is a spatially organized structure, whereas the cochlea has a *tonotopic* organization. From various bin-

This work was supported by the University of Nijmegen and the Mucom II program (6615) of the European ESPRIT initiative (M.A.F. and A.J.V.O.). We acknowledge the valuable technical assistance of H. Kleijnen and T. Van Dreumel. Correspondence should be addressed to M. A. Frens, Department of Medical Physics and Biophysics, University of Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands (e-mail: maarten@mbfys.kun.nl).

aural and monaural cues, present in the acoustical signal at the eardrums, the azimuth and elevation of a sound source appear to be derived through separate neural pathways, involving binaural and monaural processes (see Blauert, 1983; Irvine, 1986). It is generally recognized that in order to respond to a multimodal (e.g., an audiovisual) target, the different modalities must, at some stage in the neural programming, be merged into a single frame of reference (e.g., Stein & Meredith, 1993).

We wondered which rules underlie the generation of multimodally evoked targeting movements in human subjects. We have chosen the saccadic eye movement system as a model system to study this problem, since it is a very precise natural orienting system in primates and much has been learned about the neural pathways that are involved in these movements. We therefore have investigated whether gradual changes in audiovisual spatial and temporal alignment result in systematic changes of relevant saccade parameters.

Some of the results described in this paper have been presented previously in abstract form (Van Opstal, Frens, & Van der Willigen, 1993).

EXPERIMENT 1

Experiment 1 was designed to study the spatial factors that rule multimodal interactions in saccadic eye movements.

Method

Subjects

Three male volunteers, 24, 27, and 36 years of age, participated in this experiment. All subjects were without any known uncorrected visual, auditory, or oculomotor deficits, except Subject J.O., who has a strong dominance of one eye and is basically monocular. From this subject, movements of the amblyopic eye were measured.

Experimental Setup

Experiments were performed in a completely dark, sound-attenuated room ($3 \times 3 \times 3$ m) in which acoustic reflections above 500 Hz were strongly reduced by means of sound-absorbing foam that covered walls, ceiling, and floor. The average background noise level was 30 dB (SPL). The subjects were comfortably seated in a chair with a head support that prevented them from making head movements. Stimulus presentation as well as data acquisition were controlled by a PC-386 equipped with a data acquisition board (Metrabyte Das16).

Auditory noise stimuli were generated by a white-noise generator (Hewlett-Packard HOI-3722a), band-pass filtered (150–20 kHz, Krohn-Hite 3343), amplified (Luxman 58A), and presented through a speaker (Philips, AD44725). The speaker was mounted on a two-joint robot arm that was controlled by a second computer (PC-486). The robot enabled rapid positioning of the speaker anywhere on the surface of a virtual sphere with a radius of 0.90 m, centered at the subject's head. Between trials, the speaker was first moved to a randomly chosen peripheral location. In this way, sounds produced by the two stepping motors did not provide the subjects with any cues about the speaker's position.

Visual targets were red LEDs (radius 0.3°), mounted on an acoustically transparent wire frame, which constituted a spherical surface just proximal to the range of the robot ($r = 0.85$ m). Viewing was binocular.

Movements of the right eye were measured in two dimensions by means of the scleral coil technique (Collewijn, Van der Mark, & Jansen, 1975). In short, the subject was seated in a rapidly oscillating horizontal and vertical magnetic field (30 and 50 kHz), generated through two orthogonal coils (3×3 m). The coils did not obstruct the visual field of the subjects, nor did they disturb the sound field.

A scleral search coil (Skalar Instruments, Delft) was placed on the subject's right eye. The magnetic induction voltage in this scleral coil was directly proportional to its orientation with respect to the magnetic fields. In this way, eye position could be accurately measured with a resolution of about 0.25° in all directions.

In order to decompose the signal from the scleral coil into a horizontal and a vertical component of eye position, it was passed through two phase-lock amplifiers (PAR 128A) that used the driving signals for the horizontal and the vertical field, respectively, as a reference signal. The resulting position signals were then low-pass filtered at 150 Hz, before being collected by the data acquisition board. The sampling rate was 500 Hz for both the horizontal and the vertical components of eye position. Each trial consisted of 2 sec of recording time, starting 400 msec before presentation of the peripheral stimuli.

To calibrate the recorded eye movements, we asked our subjects to foveate visual targets on the horizontal axis and on the vertical axis at eccentricities of 2° , 5° , 9° , 14° , 20° , 27° , and 35° from straight ahead. The signals that were thus obtained were used to calculate off-line linear regression lines between the target coordinates and horizontal and vertical eye position signals. This method provided accurate calibration for all directions.

Experimental Protocol

The subjects were first required to foveate a visual fixation spot straight ahead. After a random period of 0.5 to 2.0 sec, the fixation spot was extinguished and a visual target was presented in the periphery. Synchronous with the onset of the visual target, an auditory stimulus was presented in 80% of the trials. Duration of each stimulus was 500 msec. The subjects were instructed to redirect their gaze as quickly and accurately as possible toward the visual target and were explicitly told to ignore the auditory nontarget.

One of four different visual targets was presented at spherical polar coordinates $R = 27^\circ$, and $\phi \in [60^\circ, 120^\circ, 240^\circ, 300^\circ]$. In this coordinate system, R is the distance from the central fixation spot and ϕ is the direction of the target, where $\phi = 0^\circ$ is to the right and $\phi = 90^\circ$ is upward (Figure 1A). Possible positions of a synchronous auditory stimulus were at these same locations. Thus, combined visual and auditory stimuli could be presented in one of four spatial configurations: (1) spatially coincident (*coincident*, for short), (2) diametrically opposed to each other with respect to the fixation spot (*opposite*), (3) horizontally aligned but vertically opposite (*horizontally aligned*), and (4) vertically aligned but horizontally opposite (*vertically aligned*).

During each experiment, all possible visual/auditory stimulus combinations were presented in random order, randomly interleaved with visual-only trials, in which the auditory stimulus was not presented (20% of the trials). Each of the 20 different stimulus configurations was presented at least 8, and—if time permitted—12, times in one experimental session.

In separate experimental sessions, the intensities of the visual and auditory noise stimuli were set either at $0.15 \text{ cd} \cdot \text{m}^{-2}$ (measured with the Minolta LS 100 luminance meter) and 70 dB (SPL) (here denoted as *high intensities*) or at $0.015 \text{ cd} \cdot \text{m}^{-2}$ and 45 dB (SPL) (*low intensities*), respectively. Note that all intensities were well above detection threshold.

In a recent study (Frens & Van Opstal, 1994), we have shown that the composition of the auditory spectrum has a strong influence on a sound's localizability. Therefore, in two separate high-intensity experimental sessions, we also selected a different spectral content of the auditory stimulus. In one session, the spectrum

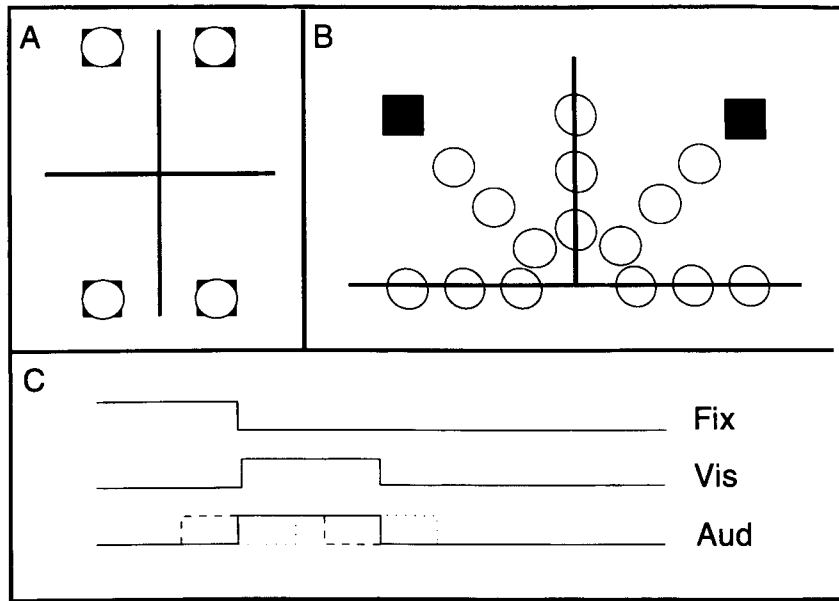


Figure 1. Schematic stimulus configuration of the three experiments described in this paper. (A) Experiment 1. Visual targets are represented as open circles; auditory stimuli are represented as filled squares. Note that the possible positions of visual and auditory stimuli coincide, but that each visual target could be presented in combination with any one of the four auditory stimuli or as a unimodal stimulus. (B) Experiment 2. A larger set of visual target positions was chosen in combination with one of two possible auditory stimuli. (C) Experiment 3. In all experiments, one of two possible visual targets was lit as soon as the central fixation spot was extinguished. The onset of the auditory stimulus (at either the same or the opposite position) could appear within a range of -50 (dashed signal) to $+100$ msec (dotted signal) with respect to the visual target onset. Neither timing nor target sizes have been drawn to scale.

was broad-band noise (see above); in the second session, it was a sharply peaked harmonic spectrum, having its most prominent component at 700 Hz (58 dB). Higher harmonics in this signal had an intensity that was at least 20 dB lower. This stimulus will be referred to as the *tone stimulus* in the rest of this paper.

At the end of an experimental session, the applied auditory stimuli were presented in a separate series, in which they served as targets. This auditory experiment aimed at measuring the latency, accuracy, and kinematic properties of auditory-evoked eye movements of our subjects. If time permitted, auditory-evoked saccades were also measured toward stimuli that were presented for 500 msec at random positions throughout the oculomotor range. These experiments served to assess the ability of our subjects to accurately localize acoustic targets.

Data Analysis

From the calibrated eye position signal, the onset and offset of saccadic eye movements were detected by the computer on the basis of velocity and mean acceleration criteria. All detection markings were visually checked by the experimenters. Subsequently, saccadic latency L (defined as the time interval between target onset and saccade onset, in milliseconds), overall saccade direction ϕ (in degrees), amplitude R (in degrees), and maximum velocity V_{\max} (in degrees/second) were determined from the calibrated eye position signals. Trials in which the primary saccade had a reaction time outside the 100–300-msec interval were discarded from further analysis. For the analysis of latencies, saccades with a direction that deviated more than 30° from the direction of the visual target were excluded.

The radial distance, ΔR , between the visual stimulus position, \vec{V} , and the auditory stimulus, \vec{A} , was defined as

$$\Delta R = \sqrt{(V_h - A_h)^2 + (V_v - A_v)^2}, \quad (1)$$

in which V_h , V_v , A_h , and A_v are the horizontal and vertical coordinates of the visual and the auditory target positions (in degrees), relative to the straight ahead fixation point.

Results

Effect on Saccade Trajectories and Kinematics

When the subjects made saccades toward well-lit unimodal visual targets, the trajectories of the movements were approximately straight in all four directions. When high-intensity noise stimuli were presented in combination with the visual targets, the trajectories of the saccades did not change with respect to the visually elicited saccades in any of our subjects. However, when the intensity of both stimuli was decreased (see Method section), the saccade trajectories depended strongly on the spatial configuration of the visual and auditory stimuli for 2 of the 3 subjects (M.F. and J.O.) in the following way (see Figure 2, for data from Subject J.O.): When the auditory noise stimulus was spatially coincident with the visual target, no systematic change with respect to the unimodal visual condition was obtained in the sac-

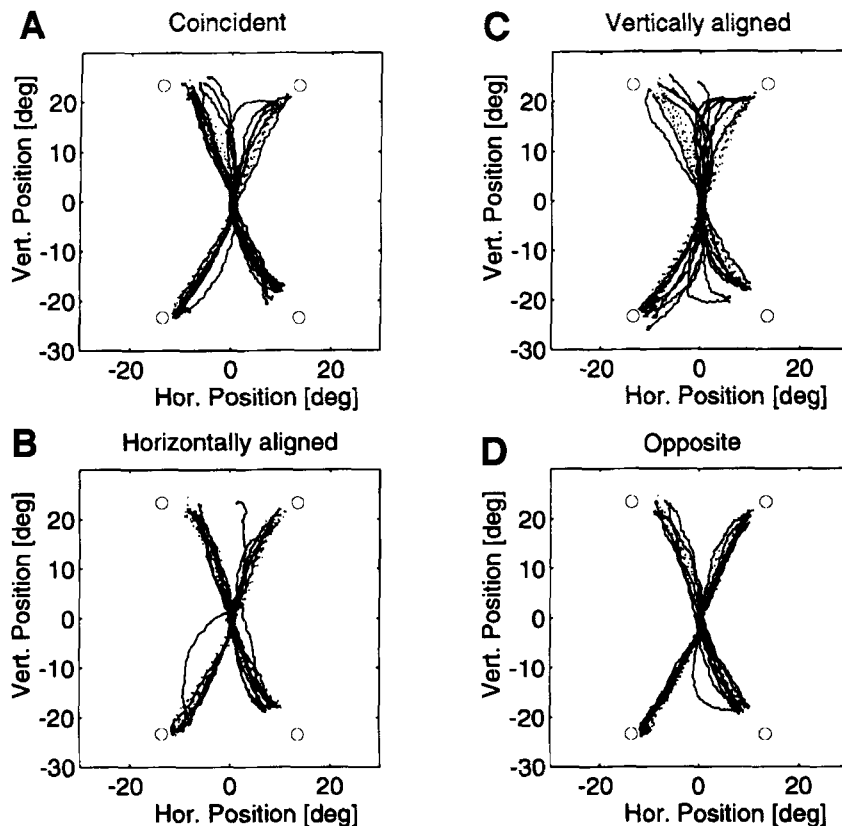


Figure 2. Experiment 1: Primary saccade trajectories (solid lines) of Subject J.O. under low-intensity bimodal stimulation. (A) Spatially coincident visual and auditory noise stimuli. (B) Horizontally aligned stimuli. (C) Vertically aligned stimuli. (D) Oppositely positioned stimuli. Open circles indicate the visual target positions (see Figure 1). For reference, in all panels, the trajectories of the unimodal visually elicited saccades are indicated by dotted lines. Note that all primary saccades undershoot the target position.

cade trajectories (Figure 2A). However, when the visual and auditory stimuli were presented vertically aligned ($\Delta\phi = 60^\circ$), the saccades typically started in a direction that was between the two stimuli. Subsequently, the movement curved in midflight toward the visual stimulus (Figure 2C). Within this population of saccades, no significant correlation between the initial direction and the latency of the responses was found ($p > .05$). Increasing the angle between the two stimuli to 120° (horizontally aligned) or 180° (opposite) resulted in straight saccades that were correctly directed toward the visual stimulus (Figures 2B and 2D). Whenever the trajectory of the saccade was not changed by the presence of the auditory target, the velocity profile and the duration of the movements were always indistinguishable from the unimodal visually driven saccades. The trajectories of Subject J.G. were straight and goal-directed under all conditions.

Effect on Latency

High-intensity stimuli. The reaction time results of this experiment are summarized in Figure 3. In Figure 3A, saccadic latencies of a representative subject (M.F.) are shown for visual targets in combination with a noise stimulus. Both stimuli have the highest applied

intensities (see Method section). One can see that the presentation of a spatially coincident auditory stimulus ("Coinc" column) reduces the latency of the response to the visual targets by about 50 msec ($p < .01$) with respect to the values found for purely visual stimuli. Note that this figure shows the pooled data of the four visual targets. This pooling was allowed since no significant differences were obtained for the latency distributions of responses toward the four visual targets ($p > .1$) in any of the spatially similar multimodal configurations.

Increasing the distance of the auditory target with respect to the visual target reduces this latency facilitation. In the most extreme case tested, in which the auditory target was positioned opposite to the visual target at a distance $\Delta R = 54^\circ$ ("Opposite" column), the effect was absent. This reduction of facilitation was significant for all subjects tested. Spearman's rank order correlation coefficients (Press, Flannery, Teukolsky, & Vetterling, 1992) between ΔR and saccade latency for all 3 subjects are given in Table 1.

For all subjects, the latency distribution of coincident audiovisual targets (second column from left in Figure 3A) was similar to the latencies of unimodal auditory-evoked saccades (right column), which could, at least in

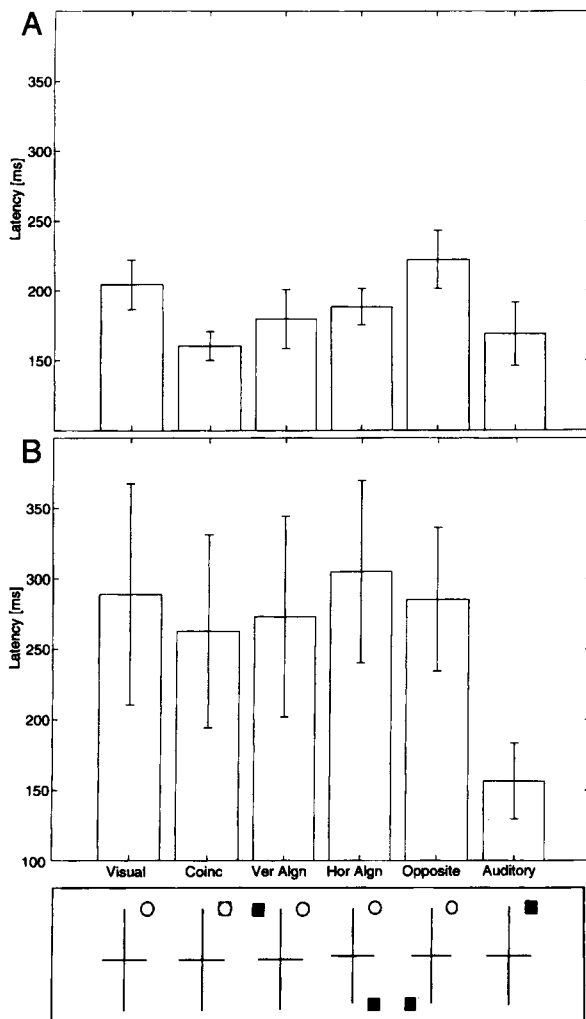


Figure 3. Experiment 1: Latencies (Subject M.F.). (A) Saccadic latencies (mean \pm SD) for the five different stimulus conditions in Experiment 1 (schematically indicated in the bottom panel for one of the visual targets) with high stimulus intensities and broad-band noise as the auditory nontarget (see Method section). Data are pooled over the four visual targets. Visual = unimodal visual stimulus; Coinc = coincident stimuli ($\Delta R = 0^\circ$); Ver Algn = stimuli are vertically aligned but horizontally opposite ($\Delta R = 27^\circ$); Hor Algn = stimuli are horizontally aligned but vertically opposite ($\Delta R = 39^\circ$); Opposite = stimuli are oppositely positioned with respect to the central fixation point ($\Delta R = 54^\circ$); Auditory = unimodal auditory stimulus. Note that the multimodal conditions are represented in ascending order of vectorial distance. The symbols at the bottom end of this figure exemplify each stimulus condition for the visual target at $(R, \Phi) = (27^\circ, 60^\circ)$. The visual target is represented as an open circle, whereas the auditory stimulus is a filled square. (B) Same format for low-intensity stimuli.

principle, be explained if, under this condition, the saccades were acoustically triggered. However, several arguments can be raised against this hypothesis. First, the accuracy of the responses to the coincident targets was much higher than was found for the auditory-guided saccades (compare, e.g., Figures 4A and 4B). In addition, the velocity profiles of audiovisual and visual saccades were stereotyped, in the sense that they obeyed the same

relations between amplitude and duration, maximum velocity or skewness. In contrast, auditory-guided saccades were generally slower and followed more curved trajectories. Finally, the results of the experiments in which low stimulus intensities were employed (see below) show that there can be substantial differences between the auditory and the coincident audiovisual latency distributions.

Low-intensity stimuli. In the low-intensity experiment, the latencies of the visually triggered saccades increased considerably with respect to those in the high-intensity condition (see Figure 3B, left-hand column). Nevertheless, the mean latency of auditory-evoked saccades remained approximately the same, which resulted in a much larger difference between the means of both unimodal latency distributions (for Subject M.F., high intensity, $\Delta L = 35$ msec, low intensity, $\Delta L = 132$ msec). Notwithstanding, the absolute effect of spatial alignment with the auditory stimulus was of the same order of magnitude as obtained in the high-intensity experiment (see also Table 1).

Note that, due to the decrease of intensities, the scatter in latencies of eye movements toward the visual targets increased, which reduced the correlations but kept the slope of the fitted linear relation intact. In contrast, the variability in the auditory saccades was not affected. This suggests that the visual stimuli were closer to the perceptual threshold than were their acoustical counterparts.

Tone stimuli. If the observed auditory-visual interactions could be attributed to a level where both modalities are represented in a common frame of reference, the localizability of the auditory target should influence the properties of the spatial interaction. Toward that end, we presented a tonal acoustic stimulus (see Method section). Figure 4 (data of Subject M.F.) shows that, compared with the broad-band sound (B), saccade accuracy to a 700-Hz tone was markedly reduced (C). In both panels, the trajectories of saccades toward the same four auditory target positions are shown. The data of Subject M.F. are representative for all subjects tested.

Table 1
Fit Parameters Between the Radial Interstimulus Distance, ΔR , and the Latency, L , of the Primary Saccades in Experiment 1

| Condition | Subject | r | N | L_{\max} | σ_L | a | σ_a |
|----------------------|---------|-----|-----|------------|------------|------|------------|
| High-intensity noise | M.F. | .72 | 172 | 156.5 | 3.2 | 0.96 | 0.09 |
| | J.G. | .33 | 134 | 176.9 | 5.0 | 0.46 | 0.13 |
| | J.O. | .53 | 176 | 177.2 | 3.0 | 0.62 | 0.08 |
| Low-intensity noise | M.F. | .29 | 136 | 262.0 | 9.2 | 0.63 | 0.25 |
| | J.G. | .33 | 171 | 264.1 | 6.8 | 0.75 | 0.18 |
| | J.O. | .31 | 173 | 271.3 | 7.2 | 0.57 | 0.21 |
| Tone | M.F. | .26 | 125 | 166.4 | 3.0 | 0.23 | 0.08 |
| | J.G. | .11 | 131 | 174.8 | 4.2 | 0.16 | 0.10 |
| | J.O. | .01 | 183 | 190.5 | 2.8 | 0.04 | 0.07 |

Note— r = Spearman's rank order correlation coefficient; L_{\max} and a = the offset and slope, respectively, of a straight line fit through the data; σ_L and σ_a = standard deviations of the parameters, which were determined by the bootstrap method (see Note 1). L_{\max} is given in milliseconds; a is given in milliseconds/degree.

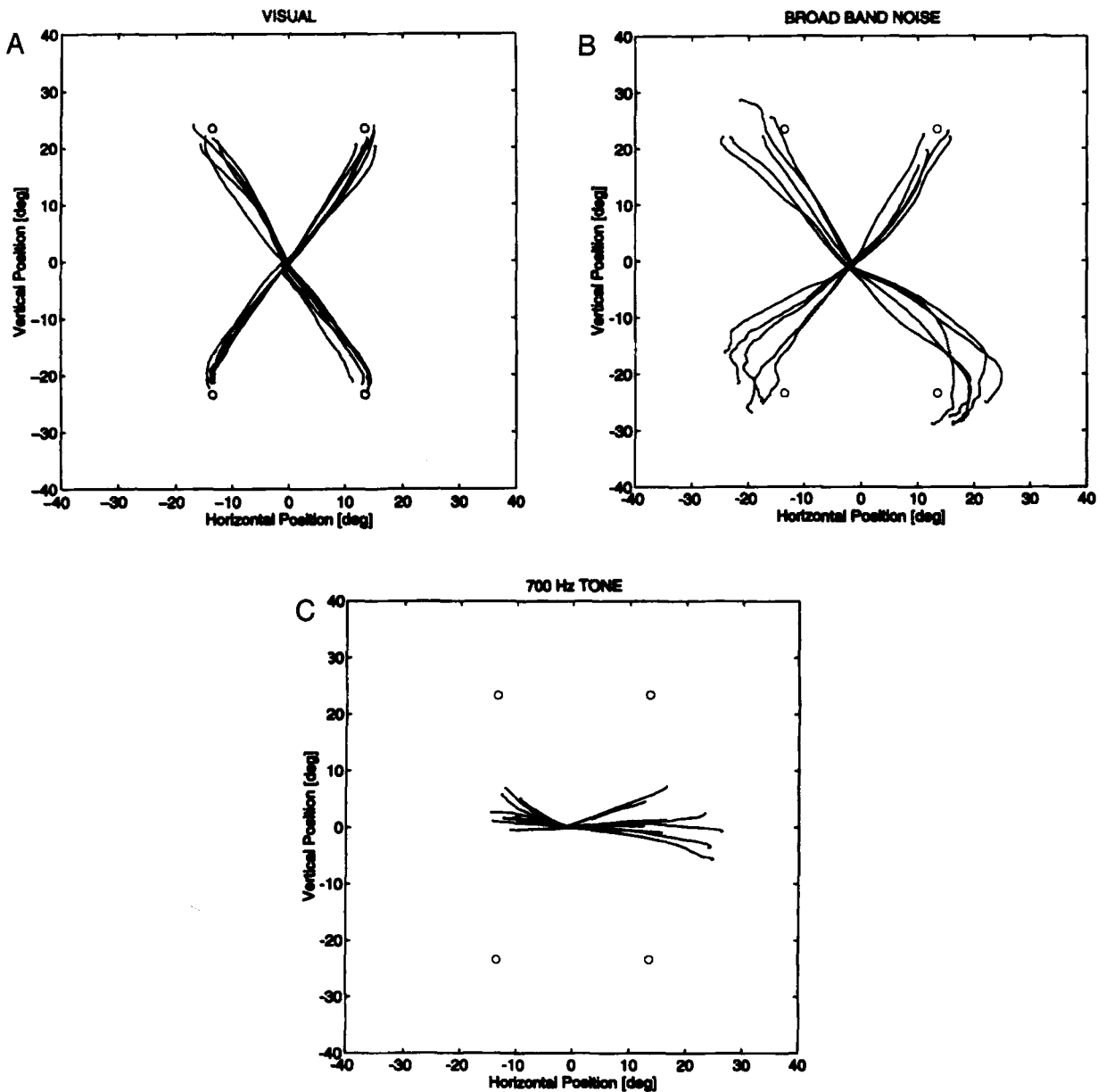


Figure 4. Auditory evoked saccade traces: Data of Subject M.F. Trajectories of primary saccadic eye movements from a visual fixation spot at (0,0) toward each of the four target positions that were presented in Experiment 1. (A) Saccades to visual targets. (B) Saccades to broad-band noise targets in darkness. Scatter in the end points of these saccades is larger than visually elicited saccades. Nevertheless, their accuracy is quite high (see text for details). (C) Saccades to 700-Hz tone targets. Accuracy of the horizontal component in these responses is comparable to the saccades evoked by broad-band noise, whereas the vertical component is not related to the elevation of the target.

Most notably, the horizontal component of the auditory-evoked saccades had the same high accuracy for both types of sound stimuli. The rank correlations between target azimuth and the azimuth of the end points of the primary saccades of the subject presented in Figures 4B and 4C are .97 (noise) and .92 (tone), respectively. Elevation was determined accurately for the broad-band noise sound ($r = .94$), whereas the vertical component was virtually constant and was not related to sound source elevation for the 700-Hz stimulus ($r = .07$). This

phenomenon has been described in an earlier paper (Frens & Van Opstal, 1994). Thus, although the physical positions of the acoustic stimuli in Figures 4B and 4C are identical, the subject's saccadic end points depended highly on the spectral content of the sounds.

In Figure 5, it is shown what happens if, instead of the well-localizable noise stimulus, the poorly localizable (at least in its vertical component) 700-Hz stimulus is used in the same experiment as is shown in Figure 3A. Note that, instead of the four different levels for saccade

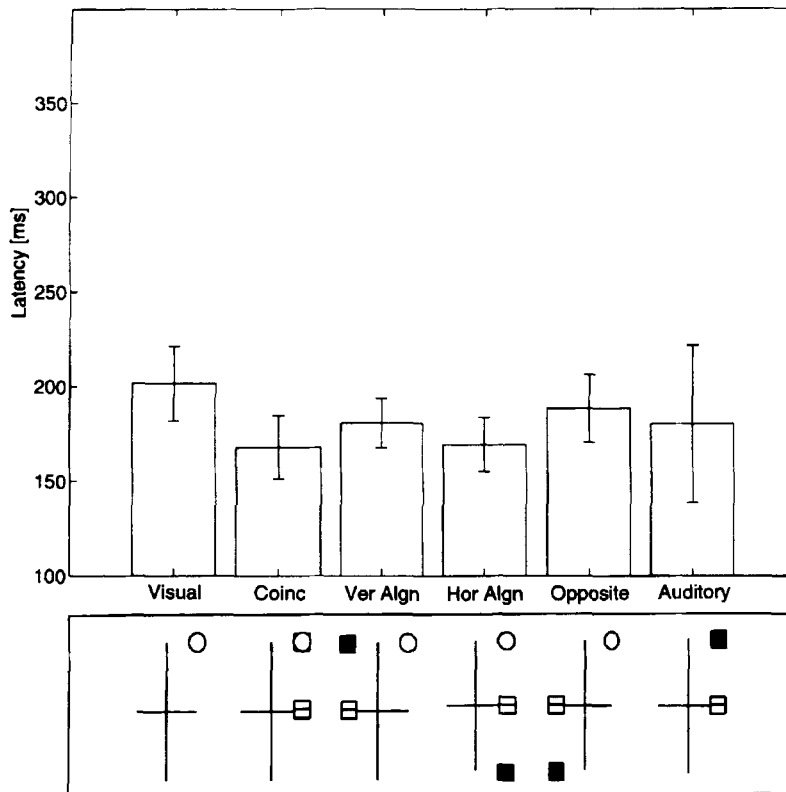


Figure 5. Experiment 1: Latencies (Subject M.F.). Same format as Figure 3 for high-intensity visual targets in combination with the auditory 700-Hz stimulus. In the bottom diagram, the perceived auditory stimulus positions, as deduced from the data in Figure 4C, are represented as shaded squares.

latencies in the audiovisual conditions of Figure 3A, now only two can be discerned. Reaction time reduction was equally effective for the auditory stimuli that had the same horizontal component as the visual target (coincident and horizontally aligned configurations). Similarly, the two audiovisual stimulus configurations, in which the horizontal components differed by the same amount (vertically aligned and opposite configurations), showed the same reduction of saccadic latency. Under these experimental conditions, in which only the horizontal component of the sound source could be localized, the distance between the *perceived* stimulus positions seemed to be the parameter that determined the spatial component of auditory-visual interactions. As a result, both the correlation and the slope are decreased with respect to the high-intensity condition (see Table 1). Rank correlation with ΔR_h (the horizontal component of the interstimulus distance), however, proved to be highly significant in all three subjects: J.O., .42; M.F., .40; and J.G., .37.

Discussion

The results of Experiment 1 show that saccadic eye movements made to a visual target that is spatially coincident with an auditory nontarget have shorter latencies than saccades that are directed toward the same visual

target without an auditory costimulus. When the visual and auditory stimuli were increasingly separated in space, this reduction of reaction time diminished in a gradual way (Figure 3A).

The finding that no differences were found between the latency distributions of the four visual targets shows that the effects cannot be due to the presentation of visual or auditory targets in different locations in space. Only their relative configuration with respect to the fovea (coincident, opposite, horizontally aligned, or vertically aligned) determined the saccadic latency.

Using two different (somewhat arbitrary) sets of visual and auditory intensities, we found no change in the strength of this spatial effect. The only objective of the low-intensity experiment was to increase the latency difference between unimodal auditory and visual saccades. Though this study is too limited to infer the effect of stimulus intensity over the full range, it does show that the effect observed is a rather robust one. Furthermore, the fact that the two unimodal latency distributions differed much more in the low-intensity condition than in the high-intensity condition but that the resulting facilitation was equal for both conditions argues against the possibility that the observed latency reduction was due to a statistical facilitation (Raab, 1962; see General Discussion section).

Under conditions where only the azimuth of the sound source could be accurately determined (tone stimuli), the horizontal separation between the stimuli, not their physical distance, determined the spatial component of the interaction. This suggests that it is the *perceived* position of the targets that rules the observed effect. Note that the site at which these interactions occur must have access to both the binaural cues that encode sound azimuth and the spectral cues that are used for elevation detection. This follows from the fact that, if these elevation cues are present in the sound signal, they also play a role in determining the amount of facilitation.

We have shown that, under low-stimulus-intensity conditions, the trajectory of visually evoked saccades can also be influenced by the presence of an auditory nontarget, provided that it is presented close to the visual stimulus (Figure 2). Under these circumstances, the saccade initially starts in a direction between the auditory and the visual stimulus.

In short, these findings indicate that there exists a spatial component in the auditory-visual interactions that subserve human oculomotor behavior. Experiment 2 was designed to make a more quantitative estimate of the observed phenomena.

EXPERIMENT 2

Inspired by data such as those seen in Figure 3, we suspected that a relevant parameter for quantifying the spatial effect on saccade latencies by auditory broad-band noise stimuli would be the perceived vectorial distance ΔR (see Method section) between the auditory and visual stimuli. However, since we only tested a limited set of distances (0° , 27° , 39.2° , and 54°), we felt that we had to explore this point further. Therefore, we performed an additional set of experiments in which the visual targets were presented throughout the upper half of the oculomotor range, in combination with one of two optimally localizable broad-band noise stimuli, thus creating a large set of perceived interstimulus distances.

Method

Subjects

Six male volunteers, ranging from 21 to 36 years of age, served as subjects. Two of these subjects (J.O. and M.F.) also participated in Experiment 1.

Experimental Setup and Protocol

All experimental equipment was the same as described in Experiment 1. Visual targets were chosen in random order from 15 locations in the upper hemisphere of the oculomotor range (see Figure 1B). Target positions were at $R \in [7^\circ, 14^\circ, 21^\circ]$ and $\phi \in [0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ]$. Synchronous auditory nontargets were at $(R, \phi) = (27^\circ, 45^\circ)$ and $(27^\circ, 135^\circ)$. Auditory stimuli always consisted of broad-band noise, and stimulus intensities were high (see Method section of Experiment 1 for further details).

Data Analysis

As a quantitative measure of the latency change in a response, due to the presentation of an auditory stimulus, we defined ΔL for each primary saccade as

$$\Delta L = L - \bar{L}_v, \quad (2)$$

in which L is the latency of the saccade (msec), and \bar{L}_v is the mean latency of a subject's responses under visual-only conditions toward that same target position. In this way, we could compare the latency effects between different targets and subjects. Perceived interstimulus distance, ΔR , was computed as the vectorial distance between the means of the unimodally evoked first-saccade end points.

Results

In Figure 6A, saccadic latency differences ΔL are shown as a function of stimulus separation ΔR (data from one representative session with Subject J.O.). As a first approximation, the relation between ΔR and ΔL was fitted by a straight line,

$$\Delta L = \Delta L_{\max} + a \cdot \Delta R, \quad (3)$$

where ΔL_{\max} is the maximum effect (in milliseconds) at spatial stimulus alignment, and a is the slope of the relation (in milliseconds/degree). This particular experiment yielded an offset $\Delta L_{\max} = -27.0$ msec and a slope $a = 0.7$ msec/deg. Spearman's rank order correlation coefficient between ΔR and ΔL ($r = .38$) was significant ($p < .01$).

Although significant rank order correlations were obtained for all subjects, sometimes an experiment yielded nonsignificant results. Thus, it was possible that, during one experiment, the presentation of the speaker at one position caused a significant effect, whereas the presentation at the other position did not. Of the obtained correlations, 50% proved to be not significant. However, the lines that were fitted through the individual data sets always had a positive slope ($a > 0$). The mean rank order correlation coefficient was $.3 \pm .2$; the mean fit parameters were $\mu_{\Delta L_{\max}} = -33 \pm 15$ msec and $\mu_a = 0.4 \pm 0.4$ msec/deg.

Since Equation 3 is an expression between relative variables, ΔL and ΔR , it is possible to directly compare the data of all subjects. We therefore pooled the data of all subjects and determined the optimal fit parameters of Equation 3 on this much larger database (Figure 6B). From the pooled data ($N = 2,130$), we obtained a highly significant ($p \leq 10^{-6}$) rank order correlation coefficient between ΔR and ΔL ($r = .25$), which was of the same order of magnitude as the mean value obtained in individual experiments (see above).

Also the offset ΔL_{\max} of -36.2 msec and slope $a = 0.55$ msec/deg are in good agreement with values obtained in the individual experiments (see above). Thus, it seems reasonable to assume that the failure to reach significance in some of the separate experiments was due to the fact that the observed effect on saccadic latency was relatively small and was masked by the large intrinsic scatter in saccade latency.

The stability of the correlation and the fit parameters of the pooled data was tested, using the bootstrap procedure.¹ In 100 bootstrap trials, the mean values of r ,

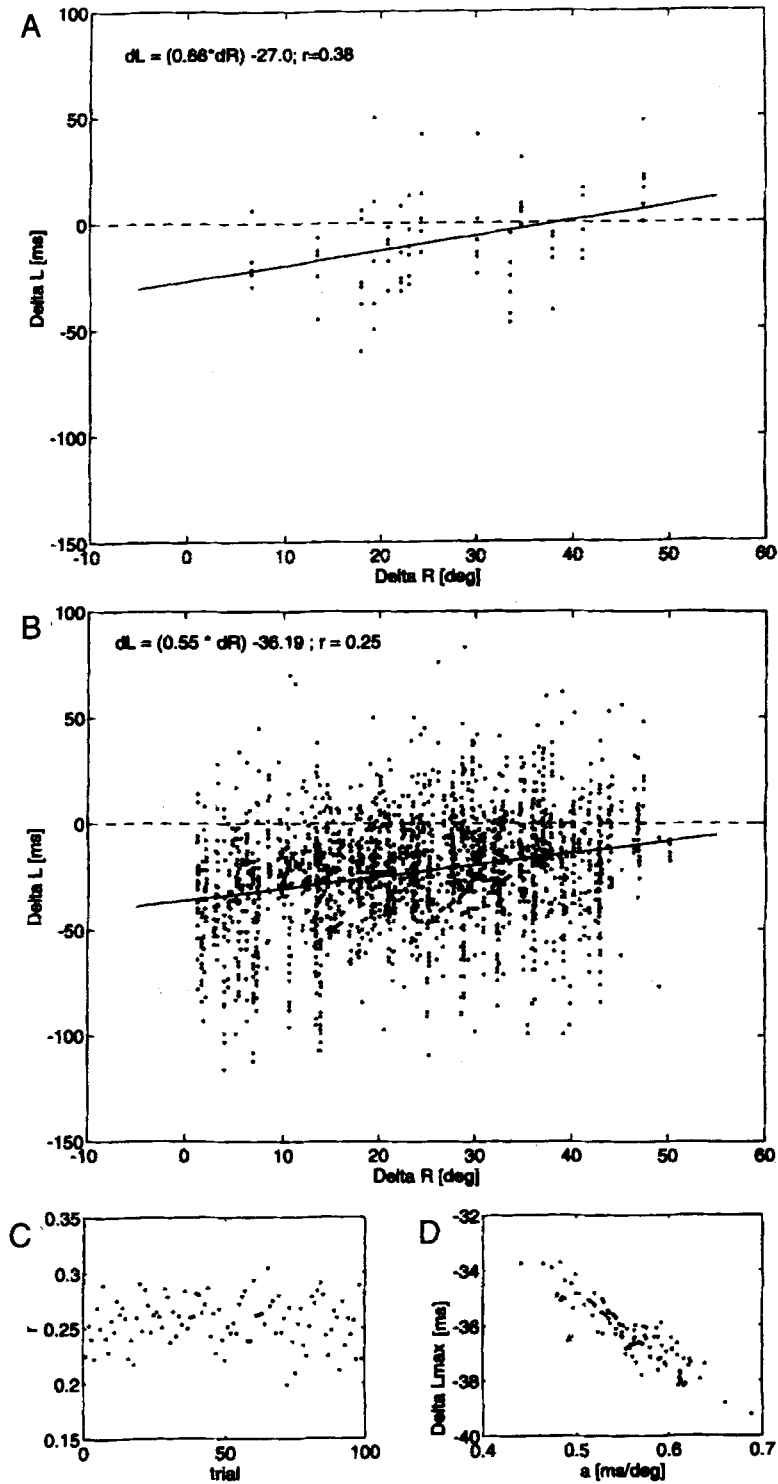


Figure 6. Experiment 2: Latencies. (A) The relation between ΔR and ΔL (see text) for a typical experimental session with Subject J.O. Data are shown for all visual targets in combination with an auditory stimulus at $(R, \phi) = (27^\circ, 45^\circ)$. Spearman's rank order correlation coefficient and the parameters of a linear fit (Equation 3; indicated by the solid line) are given in the figure ($N = 75$). (B) Same plot for the pooled data of all 6 subjects ($N = 2,130$). (C) Spearman's rank order correlation of data in 100 subsequent bootstrap trials. All correlations obtained were highly significant. Note that the $p = .01$ significance threshold is at $r = .06$. (D) Values of a (slope) and ΔL_{max} (offset) of linear fits through data in the same 100 bootstrap trials. Note that the parameter values are in the same range as obtained in the individual experiment of panel A.

ΔL_{\max} , and the slope a were identical (within 1%) to the values of the original data set. The standard deviation of the rank order correlation coefficient was $\sigma_r = .02$; the standard deviations of the fitted slope and offset were $\sigma_a = 0.05$ msec/deg and $\sigma_L = 1.2$ msec, respectively (see also Figures 6C and 6D). Thus, all three parameters were significantly different from 0 ($p \leq 10^{-6}$).

Discussion

The results of Experiment 2 confirm and extend the findings of Experiment 1, showing that there is a significant correlation between the auditory-visual stimulus distance, ΔR , and the effect of auditory costimulation on visually evoked saccadic latency, ΔL . Despite the large intrinsic scatter in saccadic reaction times, we were able to estimate the maximum effect, ΔL_{\max} , and its spatial sensitivity, a , in a robust way, due to the large data set at $\Delta L_{\max} = -36 \pm 1$ msec and $a = 0.55 \pm 0.05$ msec/deg, respectively.

Note that we found only latency reduction, since the vast majority of the data points corresponded to a negative ΔL . Even at large values of ΔR , no increase of latencies (as compared with purely visual saccades) was obtained. This point will be further elaborated upon in the General Discussion section.

EXPERIMENT 3

Experiment 3 was designed to study the temporal factor in the effect of an auditory stimulus on the latencies of saccadic eye movements toward visual targets. In contrast to the previous experiments, the onsets of the visual and auditory targets were not necessarily synchronous.

Method

Subjects

The subjects were the same as those in Experiment 1.

Experimental Setup and Protocol

All experimental conditions were identical to those of the previous experiments. In this experiment, we restricted the number of spatial target configurations to four. Only two possible visual and auditory target positions were used, at 27° to the left and to the right of the central fixation spot. In this experiment, temporal and spatial disparities between the two stimuli were introduced. The spatial configurations can be characterized as coincident and opposite, respectively. The auditory stimulus could start with a temporal onset difference of $\Delta T = -50, 0, 50,$ or 100 msec with respect to the visual target onset (see Figure 1C). Stimulus durations were always 500 msec.

Results

All 3 subjects showed similar behavior in this experiment. The relative timing of the visual and the auditory stimulus appeared to have a clear effect on the saccadic latencies. Asynchronous target onsets resulted in a gradual change of saccade latency, provided that the visual and auditory stimuli were spatially coincident (see Figure 7). Remarkably, in the condition in which the visual and the

auditory stimulus were antimeretrically positioned, the timing of the auditory target had no significant effect.

Latency reduction appeared to be optimal in our experiments if the auditory stimulus *preceded* the visual target by 50 msec ($\Delta L = -65$ msec). If auditory stimulation started 50 msec after visual target onset (but well before the onset of the saccades), the interaction was strongly reduced ($\Delta L = -25$ msec). Note that the condition $\Delta T = 0$ msec corresponds to our previous experiments. Again, we obtained a latency reduction of about 50 msec in this latter condition.

Discussion

The results of Experiment 3 show that also the temporal relation of visual and auditory stimuli affects the latency distribution of saccadic eye movements. The shortest latencies were observed not when the visual stimulus and the auditory stimulus were presented synchronously but when the auditory stimulus preceded the visual stimulus.

Due to the different speeds of light and sound, synchronously generated visual and auditory stimuli will arrive at different moments at the receptor organs. In our setup, this timing difference was about 2.5 msec. In natural conditions, these differences can be much larger. However, in all these conditions, the visual information will always precede the auditory signal. The different processing times at the sensory organs (visual, ± 60 msec; auditory, ± 20 msec) have been thought to play a role in compensating for these timing differences in order to promote auditory-visual coincidence detection at a multimodal level (Meredith, Nemitz, & Stein, 1987). However, our results show that the auditory stimulus has to precede the visual target. Therefore, the observed phenomenon is hard to explain in terms of a "designed" coincidence detection at the neuronal level.

The curve we measured for the coincident configuration resembles the data reported by Ross and Ross (1981) for the effect of the onset of a spatially fixed auditory target that served as a warning signal for a visually guided saccade. Similar effects have been found when a visual, rather than an auditory, warning cue was used to speed up choice reaction times (Bertelson & Tisse, 1969).

Surprisingly, we found no effect on the latency distributions if the auditory stimulus was presented opposite to the visual stimulus. Therefore, our results cannot be described as a general warning effect, where the auditory stimulus announces the presentation of a visual target.

In the General Discussion section, a model is presented that incorporates our spatial and temporal interaction results.

GENERAL DISCUSSION

Influence on Trajectories and Kinematics

It has previously been shown that when two visual targets are presented simultaneously, a subject often re-

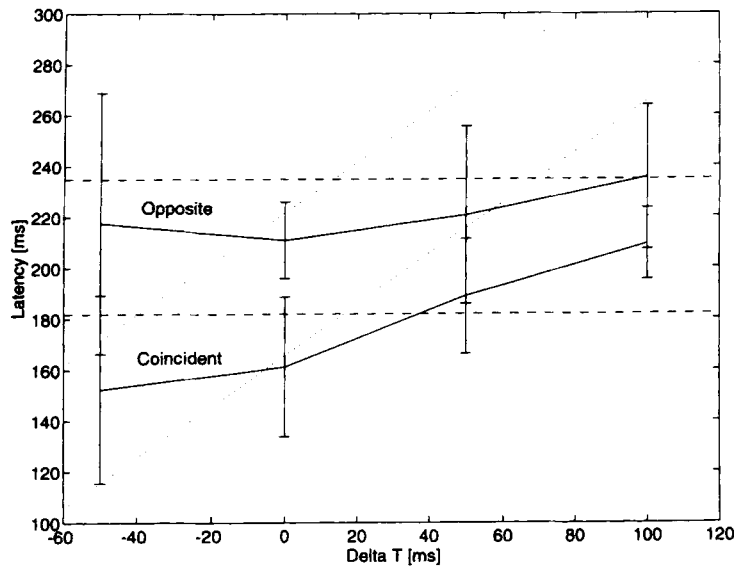


Figure 7. Experiment 3: Temporal factors in multimodal interaction on saccadic latency (Subject M.F.). Saccadic latency as a function of the temporal disparity between the onset of the visual target and the auditory stimulus (ΔT). Note that a negative value of ΔT means that the auditory target precedes the visual target (Figure 1C). Solid lines indicate the latency distribution (mean \pm SD) for the coincident and the opposite target configurations. The broken lines show the expected latency distributions (bottom line, mean $-$ SD; top line, mean $+$ SD), if the saccades were purely visually triggered (dashed lines) or purely auditorily triggered (dotted lines). Note that all latencies are calculated with respect to the onset of the visual target.

sponds with a saccade toward the average position of these two targets, provided that they are separated by an angle that does not exceed a certain maximum. This phenomenon is known as *target averaging* or the *global effect* (Findlay, 1982; He & Kowler, 1989; Ottes, Van Gisbergen, & Eggermont, 1984).

When the two visual targets are not of equal size or matched intensity, the saccade is directed toward a weighted-average position in which target size or intensity serves as a weighting factor (Findlay, 1982). It is difficult to compare the relative intensities of the visual and auditory targets that were used in our experiments. However, it should be noted that, in the high-intensity condition of Experiment 1, no averaging was observed, whereas in the same spatial configuration, averaging of initial saccade direction did occur when low-intensity stimuli were used in 2 of 3 subjects. This effect vanished if the spatial separation between the visual and the auditory stimulus was too large. Apparently, the relative intensity of the visual and auditory stimuli does play a significant role in the occurrence of averaging responses to audiovisual stimulation as well.

A similar phenomenon was reported by Lueck, Crawford, Savage, and Kennard (1990), who showed in a different task that the amplitude distribution of auditory-evoked saccades is influenced by the presence of a nearby (but not by a distant) visual stimulus in the horizontal plane.

Our findings indicate that the presence of an auditory stimulus can also influence the trajectory of saccades that are directed toward a visual target when these stimuli are separated in direction by an angle of about 60° , under conditions in which the visual signal is of poor perceptual quality. A strong competition between the auditory and visual modalities is then observed.

When the trajectory of an audiovisual saccade is not altered with respect to the purely visually driven saccades, the saccade velocity profile and its duration remain unchanged, even under low-intensity conditions. Therefore, saccades toward spatially coincident as well as widely separated multimodal targets appear to be normal main-sequence saccades (see Bahill, Clark, & Stark, 1975). This suggests that the observed interaction between the visual and the auditory system takes place at a stage where both the metrics and the latency (see below) of a saccade are specified but which is not involved in its kinematic properties.

Facilitation of Reaction Times

Our results can be tentatively explained in terms of a spatially and temporally dependent neuronal interaction. However, other factors, which have earlier been proposed in the literature, should first be considered.

For instance, Zahn, Abel, and Dell'Oso (1978) and Perrott et al. (1990) reported a reduction of saccadic reaction time to coincident audiovisual stimuli with re-

spect to purely visual targets. However, Raab (1962) argued that multimodal facilitation is not necessarily due to a neural interaction between the visual and auditory pathways. On the basis of theoretical considerations, it was shown that if the latency distributions of both modalities overlap to some degree, the subject may respond when triggered by the modality that happens to be available first in a certain trial. Thus, a *statistical facilitation*, rather than a neural interaction, may explain the shortening of average reaction times in this type of experiment. This phenomenon may indeed play a dominant role in the results obtained by Zahn et al. and Perrott et al., since their subjects made saccades toward spatially coincident audiovisual stimuli only and were not instructed to use one specific modality.

Gielen, Schmidt, and Van den Heuvel (1983) showed that, in a manual reaction time task, statistical facilitation could not account for the differences observed between latencies to unimodal and multimodal stimuli. The latency change due to presenting multimodal targets appeared to be stronger than could be expected on purely statistical grounds (see also Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994, for more recent results).

A more direct way to rule out statistical facilitation was used in this study. By making the position of the auditory nontarget completely independent of the position of the visual target, and therefore making the auditory cues irrelevant for the correct performance of the task, the subjects could not benefit from responding to the onset of the auditory stimulus. The fact that the strength of the spatial facilitation in the present study did not depend on the unimodal latency distributions (Figure 3) strongly suggests that statistical facilitation cannot explain our data (however, see also below).

Spatial Factors in Facilitation of Reaction Times

As was mentioned in the introduction, the spatial alignment of a visual and an auditory stimulus has earlier been reported to contribute to a reduction of reaction times. Lee et al. (1991) reported a decrease of 20 msec in saccadic reaction times when a tone was presented spatially coincident with a visual target, relative to purely visual reaction times. They found that this effect was not present when the tone was presented in the field contralateral to the visual target.

We obtained a mean auditory facilitation of spatially coincident stimuli of about 35 msec (see Figure 6B). That the effect appears to be stronger in our study may be due either to the localizability of the sound source (we used white noise as opposed to the tones used by Lee et al.) or to other experimental factors, such as the relative intensities of the visual and auditory stimuli. With respect to the latter point, it is important to note that we used two combinations of relatively arbitrary intensities (either high or low) and did not find a large difference in the strength of the effect, even though the difference between the latencies of saccades toward unimodal stimuli (visual only and auditory only) strongly increased (see

Figure 3). Furthermore, we showed that the reduction of reaction times steadily decreased with increasing stimulus distance, revealing a two-dimensional spatial interaction. Saccadic latencies increased with about 0.5 msec per degree stimulus separation.

We have previously shown (Frens & Van Opstal, 1994) that, unlike with broad-band noise stimuli, the elevation of tonal acoustical stimuli is not reflected in the output of the audio-oculomotor system. In contrast, the horizontal component (the azimuth) is still as accurate as it is under auditory noise conditions. Presumably, this relates to the fact that, in the process of auditory localization, azimuth and elevation of a sound source are determined on the basis of different cues in the sound signals at the eardrums. Typically, the azimuth of a sound source is derived from binaural cues, such as interaural timing and intensity differences. Sound elevation, however, is based on spectral filtering by the pinnae, which constitutes a monaural cue (see Blauert, 1983, for review). Since, in the case of a single tone, the spectral filtering is not uniquely related to elevation of a tonal sound source, this spatial parameter cannot be extracted unambiguously.

The results of the experiments using tones, described in this paper (Figure 3C), show that these features of localizability are also found in the interaction between auditory and visual stimulation. Under these circumstances, the horizontal component of the sound source position determines the strength of the interaction, whereas the actual vertical component does not play a role. This agrees with the perceived elevation of the tone stimuli, which is constant (Figure 4C).

Although the presentation of an auditory nontarget could decrease the saccadic reaction times considerably (in the order of 20%–30%), the spatial dependence could be weak and often resulted in nonsignificant relations. The main reason for this seems to be that the scatter in saccadic reaction times is rather large by nature. Therefore, a weak effect is likely to disappear in the background noise. Pooling all normalized data, however, resulted in an effect that was highly significant and consistent with the results of the first experiment (cf. Figures 3A and 6B). The values of the rank order correlation coefficient and the linear fit parameters were comparable with the means that were found in the separate data sets.

Temporal Factors in Facilitation of Reaction Times

We also found that the relative timing of the visual and auditory stimuli influenced the strength of the latency reduction, provided that the stimuli were in spatial proximity. In contrast to the electrophysiological findings of Meredith et al. (1987), we found that an optimal facilitation arose if the auditory stimulus was presented slightly *before* or synchronously with the visual target.

Strikingly, in many of our subjects, the unimodal auditory latencies were on average *shorter* than the uni-

modal visual latencies. Thus, the observed optimal relative timing seems to play no role in aligning the two unimodal latency distributions.

As discussed above (see Discussion section of Experiment 3), the facilitation also cannot be explained by an aspecific warning signal (see Ross & Ross, 1981), since temporal facilitation takes place only when the stimuli are spatially coincident. In what follows, we will present a neurophysiologically inspired model that deals with this paradox.

Neurophysiological Interpretation

Metrics

Electrophysiological recordings have shown that, in the frontal eye fields (FEF; Bruce & Goldberg, 1985) and in the deep layers of the superior colliculus (DLSC; McIlwain, 1982; Sparks, Holland, & Guthrie, 1976), saccadic eye movements are encoded by a large population of cells in a two-dimensional motor map.

If two targets are sufficiently close, the populations that would encode the individual target positions and the corresponding saccades might merge into a single region of activity, with its peak activity at the region that encodes the average saccade vector. Indeed, simultaneous electrical stimulation at two sites in the DLSC produces weighted-averaging saccades (Robinson, 1972). Recordings in the DLSC during averaging saccades show that their coordinates are already represented at this stage as a single peak of activity (Glimcher & Sparks, 1993; Van Opstal & Van Gisbergen, 1990).

It has been shown that, in the DLSC (Jay & Sparks, 1987) and in FEF (Russo & Bruce, 1994), eye movements toward auditory targets are represented in an oculocentric motor map, which is thus in spatial register with the visual modality. The averaging responses between visual and auditory targets that we found may therefore be caused by the same mechanism that is thought to be responsible for the averaging responses toward two visual targets (see also Lueck et al., 1990).

Latency

In anaesthetized cats, it has been shown that, when near-threshold visual and auditory stimuli were presented at the same location and with a characteristic time interval, multimodal cells in the DLSC responded vigorously with a firing rate that could be more than 10 times the optimal unimodal stimulus response (Meredith & Stein, 1986b). Spatial and temporal disparities, however, caused a *gradual* decrease of this nonlinear bimodal interaction and could even be reversed into an inhibitory interaction (Meredith et al., 1987; Meredith & Stein, 1986a). These interactions have been demonstrated to depend on a specific cortical input. Multimodal neurons from the feline anterior ectosylvian sulcus (area AES) project in spatial register to the DLSC (Wallace, Meredith, & Stein, 1993). Reversible inactivation of AES completely disrupts the multisensory interactions in the DLSC (Wilkinson, Meredith, & Stein, 1992).

The multimodal cells in the superior colliculus of the cat respond optimally when a certain characteristic delay is introduced between the onset of the visual and the auditory stimulus (Meredith et al., 1987). Typically, the optimal auditory stimulus must be presented about 0–100 msec *after* the visual stimulus.

Our experimental results show striking similarities to those obtained electrophysiologically in the DLSC of the cat. If one substitutes *latency* with *firing rate*, the same rules of multisensory integration seem to apply for human saccadic eye movements and activity in the superior colliculus. How could the firing rate of collicular neurons and saccade latency relate?

Current ideas suggest that fixation neurons in the rostral part of the DLSC excite “omnipause” neurons (OPNs) in the brainstem, which are generally thought to act as an inhibitory gate on the saccade burst generator (BG; e.g., Munoz & Wurtz, 1993). In this way, the rostral pole of the DLSC may control active fixation by inhibiting saccade generation. Therefore, in order to elicit a saccade, this rostral fixation zone first has to be silenced. Enhancement of the firing rates of presaccadic neurons in the colliculus may facilitate the crossing of a certain threshold needed to overcome the fixation-related activity and may therefore reduce the saccadic latency (see Figure 8).

The acoustic stimulus may function as a “warning signal” if it is assumed that the auditory modality, due to a faster access to the brainstem, also exerts a nonspecific inhibitory effect on the OPNs. As a result, the DLSC–

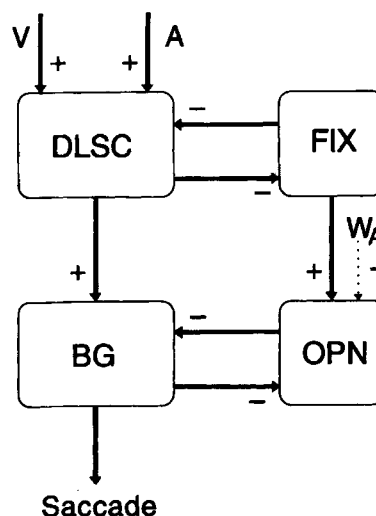


Figure 8. Proposed mechanism for saccade facilitation (based on Munoz & Wurtz, 1993). Visual (V) and auditory (A) signals project to saccade-related burst neurons in the DLSC. Bimodal interactions may either facilitate or detain the crossing of a threshold, necessary to silence the fixation neurons (FIX). These neurons act through the omnipause neurons (OPNs) as an inhibitory gate on the generation of a saccade by the saccadic burst generator (BG). The auditory stimulus also has an aspecific inhibitory effect on the OPNs (W_A), which may serve as a warning signal through disinhibition of the BG.

BG pathway may cross the saccade initiation threshold sooner. This is especially the case when the target and nontarget are spatially coincident, because, in this case, local excitatory interactions in the motor map may further boost the collicular presaccadic activity. If, however, the targets are spatially unaligned, the benefit caused by the warning signal may diminish or even be completely canceled as a result of lateral inhibition within the map. In short, such a scheme may explain (1) the observed facilitation of saccades to spatially coincident stimuli, (2) the lack of a latency increase of saccades to spatially unaligned stimuli, (3) the lack of a temporal effect on movements to spatially unaligned stimuli, and (4) why coincidence detection alone does not underlie our results.

A similar process is currently thought to explain the frequent occurrence of short-latency (express) saccades in the "gap paradigm" (Fischer & Weber, 1993). In this case, the early offset of the fixation point may exert the same "aspecific" warning effect by OPN-BG disinhibition, allowing spatially selective collicular activity to overcome the threshold at an earlier moment.

Electrophysiological recordings in the DLSC of awake animals are needed in order to reveal the link between multimodally driven behavior and the underlying neural processes.

REFERENCES

- BAHILL, A. T., CLARK, M. R., & STARK, L. (1975). The main sequence: A tool for studying human eye movements. *Mathematical Biosciences*, **24**, 191-204.
- BERTELSON, P., & TISSEYRE, F. (1969). The time course of preparation: Confirmatory results with visual and auditory warning signals. *Acta Psychologica*, **30**, 145-154.
- BLAUERT, J. (1983). *Spatial hearing: The psychophysics of human sound localization*. Cambridge, MA: MIT Press.
- BRUCE, C. J., & GOLDBERG, M. E. (1985). Primate frontal eye fields: I. Single neurons discharging before saccades. *Journal of Neurophysiology*, **53**, 603-635.
- COLLEWIJN, H., VAN DER MARK, F., & JANSEN, T. J. (1975). Precise recording of human eye movements. *Vision Research*, **15**, 447-450.
- FINDLAY, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, **22**, 1033-1045.
- FISCHER, B., & WEBER, H. (1993). Express saccades and visual attention. *Behavioral & Brain Sciences*, **16**, 553-610.
- FRENS, M. A., & VAN OPSTAL, A. J. (1994). Auditory-evoked saccades in two dimensions: Dynamical characteristics, influence of eye position, and sound source spectrum. In J. Delgado-García, E. Godaux, & P. P. Vidal (Eds.), *Neural mechanisms underlying gaze control* (pp. 329-339). Oxford: Pergamon Press.
- GIELEN, C. C. A. M., SCHMIDT, R. A., & VAN DEN HEUVEL, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics*, **34**, 161-168.
- GLIMCHER, P. W., & SPARKS, D. L. (1993). Representation of averaging saccades in the superior colliculus of the monkey. *Experimental Brain Research*, **95**, 429-435.
- HE, P. Y., & KOWLER, E. (1989). The role of location probability in the programming of saccades: Implications for "center-of-gravity" tendencies. *Vision Research*, **29**, 1165-1181.
- HUGHES, H. C., REUTER-LORENZ, P. A., NOZAWA, G., & FENDRICH, R. (1994). Visual-auditory interactions in sensorimotor processing: Saccades versus manual responses. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 131-153.
- IRVINE, D. R. F. (1986). The auditory brainstem. In D. Ottoson (Ed.), *Progress in sensory physiology* (Vol. 7). Berlin: Springer-Verlag.
- JAY, M. F., & SPARKS, D. L. (1987). Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *Journal of Neurophysiology*, **57**, 22-34.
- LEE, C., CHUNG, S., KIM, J., & PARK, J. (1991). Auditory facilitation of visually guided saccades. *Society for Neuroscience Abstracts*, **17**, 862.
- LUECK, C. J., CRAWFORD, T. J., SAVAGE, C. J., & KENNARD, C. (1990). Auditory-visual interaction in the generation of saccades in man. *Experimental Brain Research*, **82**, 149-157.
- MANLY, B. F. J. (1991). *Randomization and Monte Carlo methods in biology*. London: Chapman & Hall.
- MCILWAIN, J. T. (1982). Lateral spread of neural excitation during microstimulation in the intermediate gray layer of cat's superior colliculus. *Journal of Neurophysiology*, **47**, 167-178.
- MEREDITH, M. A., NEMITZ, J. W., & STEIN, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons: I. Temporal factors. *Journal of Neuroscience*, **10**, 3215-3229.
- MEREDITH, M. A., & STEIN, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, **365**, 350-354.
- MEREDITH, M. A., & STEIN, B. E. (1986b). Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, **56**, 640-662.
- MUNOZ, D. P., & WURTZ, R. B. (1993). Fixation cells in monkey superior colliculus: II. Reversible activation and deactivation. *Journal of Neurophysiology*, **70**, 576-589.
- OTTES, F. P., VAN GISBERGEN, J. A. M., & EGGERMONT, J. J. (1984). Metrics of saccade responses to visual double stimuli: Two different modes. *Vision Research*, **24**, 1169-1179.
- PERROTT, D. R., SABERI, K., BROWN, K., & STRYBEL, T. Z. (1990). Auditory psychomotor coordination and visual search performance. *Perception & Psychophysics*, **48**, 214-226.
- PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A., & VETTERING, W. T. (1992). *Numerical recipes in C* (2nd ed.). Cambridge: Cambridge University Press.
- RAAB, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, **24**, 574-590.
- ROBINSON, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research*, **12**, 1795-1808.
- ROSS, S. M., & ROSS, L. E. (1981). Saccade latency and warning signals: Effect of auditory and visual stimulus onset and offset. *Perception & Psychophysics*, **29**, 429-437.
- RUSO, G. S., & BRUCE, C. J. (1994). Frontal eye field activity preceding aurally guided saccades. *Journal of Neurophysiology*, **71**, 1250-1253.
- SIMON, J. R., & CRAFT, J. L. (1970). Effects of an irrelevant auditory stimulus on visual choice reaction time. *Journal of Experimental Psychology*, **86**, 272-274.
- SPARKS, D. L., HOLLAND, R., & GUTHRIE, B. L. (1976). Size and distribution of movement fields in the monkey superior colliculus. *Brain Research*, **113**, 21-34.
- STEIN, B. E., HUNNEYCUTT, W. S., & MEREDITH, M. A. (1988). Neurons and behaviour: The same rules of multisensory integration apply. *Brain Research*, **448**, 355-358.
- STEIN, B. E., & MEREDITH, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- VAN OPSTAL, A. J., FRENS, M. A., & VAN DER WILLIGEN, R. F. (1993). A spatial component for auditory-visual interactions in human saccadic eye movements. *European Journal of Neuroscience*, **6** (Suppl.), 273.
- VAN OPSTAL, A. J., & VAN GISBERGEN, J. A. M. (1990). Role of monkey superior colliculus in saccade averaging. *Experimental Brain Research*, **79**, 143-149.
- WALLACE, M. T., MEREDITH, M. A., & STEIN, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *Journal of Neurophysiology*, **69**, 1797-1809.
- WILKINSON, L. K., MEREDITH, M. A., & STEIN, B. E. (1992). Cortical deactivation disrupts multisensory integration. *Society for Neuroscience Abstracts*, **18**, 1031.

ZAHN, J. R., ABEL, L. A., & DELL'OSSO, L. F. (1978). Audio-ocular response characteristics. *Sensory Processes*, **2**, 32-37.

NOTE

1. The bootstrap method is a statistical procedure to estimate the confidence limits of the fit parameters for a data set when the underlying probability distribution of the errors is unknown (see, e.g., Press et al., 1992; Manly, 1991). In short, from the measured set of N data points, a new hypothetical data set is generated by randomly drawing,

with uniform probability, N points from the original data. In general, the overlap between the new and the original data set will be approximately 63%. The fit procedure is applied on the new set as if it were the real data set, resulting in different fit parameters. Repeating this procedure P times results in a distribution of P values for each parameter, from which the confidence limits can straightforwardly be determined.

(Manuscript received June 13, 1994;
revision accepted for publication January 27, 1995.)