

Eye and Head Movements to Auditory Targets*

D. A. Whittington, M.-C. Hepp-Reymond¹, and W. Flood

Department of Psychology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Summary. Three adult female monkeys were trained to direct their gaze toward auditory targets. When the animals were free to move their heads about the vertical axis, this was accomplished with short-latency, coordinated eye-head movements reminiscent of responses to visual targets. The similarity of response to auditory and visual targets suggests a common motor program elicited by stimuli of different modalities. Since these modalities do not share the same reference system, this implies a remapping between the two reference systems.

Key words: Eye movements – Saccade – Coordinate remapping – Auditory localization – Eye-head coordination

It is a general observation that a variety of animals (monkeys and humans as well) will orient visually to an auditory target. That is, they will make coordinated eye and head movements so that their gaze is directed towards a sound. In theory, such responses are inherently different from visually elicited movements because the auditory stimulus is detected in a head-centered reference system as opposed to the retinocentric frame of visual targets. This observation raises several questions regarding the mechanisms involved in generating saccades and in the transformation, or remapping, from one coordinate

system to another in this type of sensorimotor interaction.

Two hypotheses have been proposed for the generation of saccades. The first holds that the size and direction of visually-elicited saccades is completely specified in terms of retinal error as defined by the directed distance from the center of the fovea to target image on the retina. The retinal error and corresponding saccade can be computed without considering either afferent or refferent information about eye position in the orbit. The second hypothesis describes saccades in terms of the position of the eye in the orbit at the end of a saccade. This description is more complex because the eye's final position is equal to the sum of the retinal error and the initial position of the eye. The advantage of simplicity in specifying saccade size in terms of retinal error is lost when non-visual targets are considered. The location of an auditory target, for instance, is detected in a head-centered system and, therefore, does not correspond directly to retinal error. The saccade required to fixate a sound source is equal to detected auditory position minus the eye's initial position in the orbit. It follows that, unless one proposes that the specification of saccades changes with the sensory modality of the target, whichever specification (retinal error or final position) is used, one target modality or the other must be remapped to conform to that specification. Work by Hallett and Lightstone (1976) tends to support the final position argument by demonstrating that in some instances accurate saccades can be elicited to a visual target even though the retinal error produced by that target is inappropriate. Nevertheless, the issue is not resolved, and information on saccadic responses to non-visual stimuli may help.

The problem of coordinate remapping becomes even more pronounced if the head moves during the response to a stimulus. When the head moves, the

* This research has been supported by the National Institutes of Health Research Grant NS 09343, awarded by the National Institute of Neurological Diseases and Stroke, National Aeronautics and Space Administration Grant NGR 22-009-798, and the National Institute of General Medical Science, NIGMS 5 T01-GM01064

¹ Present address: Institut für Hirnforschung der Universität Zürich, August-Forel-Strasse 1, CH-8029 Zurich, Switzerland
Offprint requests to: Dr. D. A. Whittington (address see above)

saccade size or final position must be modified to account for the movement of the head. Bizzi et al. (1971) have delineated the mechanisms for foveating peripheral visual targets involving coordinated eye and head movement. They showed that the eye moves first towards the target, followed by the head. The movement of the head modulates the size of the saccade via the vestibular system so that the eye is not carried past the target (Morasso et al. 1973), and the continuing rotation of the head is now compensated by a smooth eye movement equal in magnitude but opposite in direction to the head movement. Whether a similar strategy of coordinated eye-head movements is elicited by auditory targets has not been previously investigated.

A number of studies have demonstrated that humans, monkeys, and cats can accurately judge the location of sounds with reference to their heads and bodies. In humans, auditory localization has been extensively studied by Stevens and Newman (1936), Wallach (1939), Mills (1958) and Thurlow et al. (1967), either by measuring the smallest detectable angular separation between two sound sources ("minimum audible angle") (Mills 1958), or by matching the position of one sound to another. The role of head movements has also been considered but mainly with emphasis on the localization improvement gained by the movement (Thurlow and Runge 1967). Latency as well as accuracy of eye movements during localization have recently been investigated in humans with the head fixed (Zahn et al. 1978).

In monkeys (Heffner and Masterton 1975) and in cats (Thompson and Masterton 1978), orienting behavior toward auditory targets has been studied; specifically, the deficits caused by lesions of auditory structures and pathways. The first study involved monkeys walking toward the sound source, while in the second study, latency and accuracy of head orientation movements to unexpected noise were reported.

In view of the lack of information about the strategy of foveation of auditory targets, and about the attendant questions of coordinate remapping, the following experiments were undertaken in awake monkeys.

Methods

Pretraining and Surgery

Three adult female monkeys (*Macaca mulatta*) were trained to discriminate between horizontal and vertical thin black lines (3 min in width) superimposed on a one degree luminous spot. The monkeys received a reward of water for bar-pressing only during the presentation of the vertical hairline, while inappropriate bar-

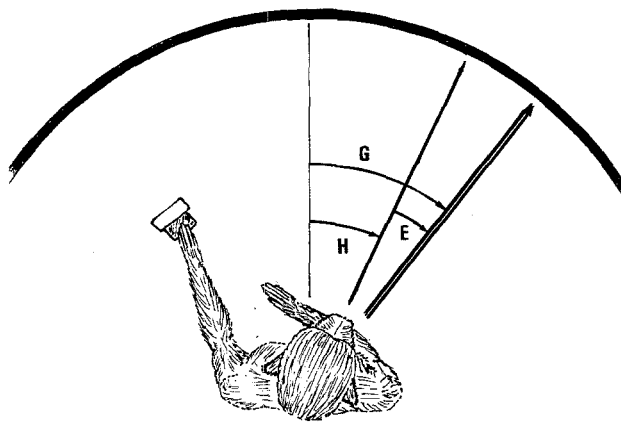


Fig. 1. Plan view of monkey seated in front of target-carrying arc, showing relation between eye position (E), head position (H), and gaze position (G). These positions are expressed as angular deviations from a reference line. For head and gaze, the zero reference is a line between the center of the monkey's head and the center of the arc; for the eyes, the reference is the mid-sagittal plane of the head

pressing delayed the next target sequence. The target sequence consisted of a luminous spot of randomized (0.4–1.5 s) duration followed by superimposition of a horizontal or vertical hairline and appeared randomly at different horizontal positions on a perimeter arc 60 cm from the animal (Bizzi et al. 1971; Morasso et al. 1973).

After the monkeys became proficient at this task, they were anesthetized with Nembutal, and silver-silver chloride electrodes were implanted in their orbital bones adjacent to the outer canthus of each eye for recording extraocular potentials. Stainless steel screws were implanted in their skulls for attachment to the head movement recording apparatus.

Experimental Procedure

After recovery from surgery, the monkeys were trained to direct their visual axis toward auditory targets. These targets consisted of 15 pps click trains or of a single click presented through one of nine 5 cm diameter loudspeakers. The intensities of the clicks in the various loudspeakers were equalized by means of resistors in series with each speaker to a level of 57 db SPL as measured with a General Radio SPL meter (Model 1551-C) at the location of the monkey's head. The nine loudspeakers were positioned at 10 deg intervals on the perimeter arc, directly above the target lights, level with the monkey's eyes. To minimize sound reflections which would complicate the task, draping was installed behind and beside the primate chair and sound absorbant foam fitted to the chair and the arc. The monkeys sat in the primate chair with their heads attached, by way of the skull screws, to a head holder which limited head movement to rotation about the vertical axis. Head movements were monitored from a potentiometer mechanically coupled to the head holder, while eye movements were obtained from the implanted silver-silver chloride electrodes. The eye movement signal was electronically linearized and summed with the head movement signal to produce a signal called "gaze" which denotes the position of the visual axis in reference to the target (Fig. 1).

Early stages of training involved pairing the auditory targets with the visual ones and then, later, making the presentation of the visual target dependent upon the monkey looking towards the

auditory target. In the final stage of training the visual targets were no longer used and the arc was draped with a coarse weave cloth which made it impossible to see the loudspeakers and the lights. A water reward was given only when the monkey responded to an auditory target by bringing its gaze to within $\pm 5^\circ$ of the sound source and holding it there for 500 ms. No attempt was made to force the monkeys to make the most accurate or rapid movements possible. Rather, we wanted to elicit, as nearly as possible, a natural response to an auditory stimulus of interest.

Data Collection and Analysis

For the data collection, series of auditory targets were presented at randomized positions along the arc, and the elicited eye and head movements recorded. Data were collected when the monkeys were free to move their heads as well as when the heads were restrained and with both single click and click-train targets. Calibration of eye position for these recordings was performed by presenting a sequence of target lights at reference positions at the beginning and end of the session. For the sake of comparison, eye and head movements triggered by visual targets were also stored in the head fixed and head free conditions for two monkeys. The data were recorded simultaneously on a paper tape oscillograph (Honeywell Visicorder 906) and on an FM magnetic tape recorder.

Results

Strategy of Eye-head Movements

Figure 2A shows a typical response to an auditory target when the head is free. The onset of the target in the periphery triggers a saccade to the target followed by a head movement in the same direction. During the head movement, a compensatory, counter rotation of the eye serves to keep the gaze on target. Comparison with Fig. 2B, which shows the same monkey's response to a visual target, reveals the striking similarity of both motor strategies. The data in Table 1 support the similarity, showing that for both auditory and visual targets, on the average, the eyes move slightly before the head. As Bizzi et al. (1971) has shown for visually elicited targets, this corresponds to a head before eye EMG strategy which, because of the large mass of the head, results in the detectable onset of head movement being delayed to roughly coincide with the onset of eye movement.

Latency of Eye and Head Movements

From the analysis of our data the only feature which differentiates auditory from visually-triggered movements is response latency. These differences are shown in Table 1, which gives the mean movement latencies for visual and auditory targets. On the average, the monkeys' eye and head movements

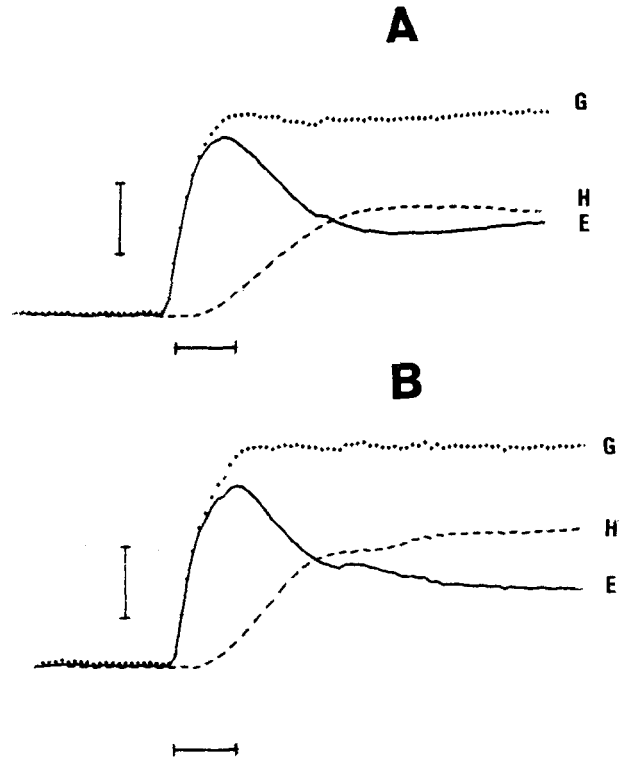


Fig. 2A, B. Eye (E), head (H), and gaze (G) position traces. A is a response to an auditory target; B is a response to a visual target. The initial position of the eye and head have been adjusted to coincide to make the pattern of movement clearer. Calibrations: vertical = 20° , horizontal = 100 ms

have significantly shorter latencies when elicited by auditory rather than by visual stimuli. The mean differences in latency are 56 ms for the eye movements and 59 ms for the head movements. The eye movement latency was also investigated in the situation where the monkey's head was restrained, so that orientation to the auditory (or visual) targets was accomplished by eye movements only. This experiment confirmed the shorter latency to auditory targets with an average difference in latency of 65 ms (Table 2), and it also allowed the investigation of the effect of initial position of the eye on the accuracy of the eye movements to auditory targets.

Effect of Eye Position on Saccadic Accuracy

In the condition when the monkeys' heads were restrained, the animals responded to the presentation of auditory targets by making saccadic eye movements toward the targets. Since the position of the eyes at the start of each movement was not constrained, the collected data consist of saccades from a random distribution of initial eye positions to each of

Table 1. Latencies for eye and head movements for visual and auditory targets with the head free. For the two animals which have both visual and auditory responses (b-76 and 6-7) the differences in latencies between auditory and visual stimuli are all significant ($P < 0.01$). For the two animals with both click and click train responses (b-74 and 6-7) none of the latencies are significantly different ($P < 0.01$) for those two stimuli. N = number of movements, s.d. = standard deviation

monkey no.	single click		click train		visual target		
	eye	head	eye	head	eye	head	
b-76	no data	no data	173	177	231	236	Latency (ms)
			30	30	30	30	n
b-74	109	122	126	126			lat.
	30	30	50	50	no data	no data	n
6-7	26	20	35	37			s.d.
	143	146	138	142	192	200	lat.
b-74	50	50	50	50	35	35	n
	39	43	49	49	36	32	s.d.

Table 2. Latencies for eye movements for visual and auditory targets with the head fixed. The differences in eye and head latencies between auditory and visual targets is significant ($P < 0.01$) for all entries. The differences for the two types of auditory targets are not significant ($P < 0.01$) for the two monkeys (b-74 and 6-7) for which data exist. N = number of movements, s.d. = standard deviation

monkey no.	single click	click train	visual target	
b-76	no data	190	232	latency (ms)
		30	30	n
b-74	138	114	216	lat.
	30	50	50	n
6-7	58	42	46	s.d.
	139	143	194	lat.
b-74	50	40	50	n
	35	42	35	s.d.

the nine targets. By plotting the error of the final position of each saccade versus its starting position (Fig. 3), one obtains a measure of the influence of initial position on accuracy. Figure 3 shows that while there is considerable scatter (the monkeys were not required to be highly accurate), the error is constant over the range of initial positions.

Role of Types of Auditory Stimulus

In an effort to determine the role of the stimulus in movements to auditory targets, each of the experiments was repeated with two different auditory

targets; (1) a 15 Hz click train continuing throughout the movement and (2) a single click. Apparently auditory feedback does not influence the motor strategy since both stimuli produce indistinguishable movements. Tables 1 and 2 bear this out with respect to latencies which do not show statistically significant differences and, within the requirements of the task, there was also no difference in accuracy between the two types of auditory targets. In order to gauge the appropriateness of click stimuli, natural sounds rather than the highly artificial clicks and click trains, were also presented through the speakers. These produced the same short-latency stereotyped movements as the clicks and click trains.

Finally, in order to be sure that the monkeys were moving to auditory targets, as opposed to making trained movements to slight variations in tonal quality of the different sound sources, the positions of the sources were altered, and the experiments repeated. This produced accurate movements to the appropriate targets, eliminating tonal variation as a possible source of error.

Discussion

Our results show that the motor strategy used by Rhesus monkeys to fixate auditory targets is virtually identical to the strategy used to fixate visual targets. They suggest that there exists a motor program for directing the gaze toward targets of interest, and that this same program is triggered by stimuli of various modalities. Further, since the two types of sensory stimuli used in the experiments do not share a

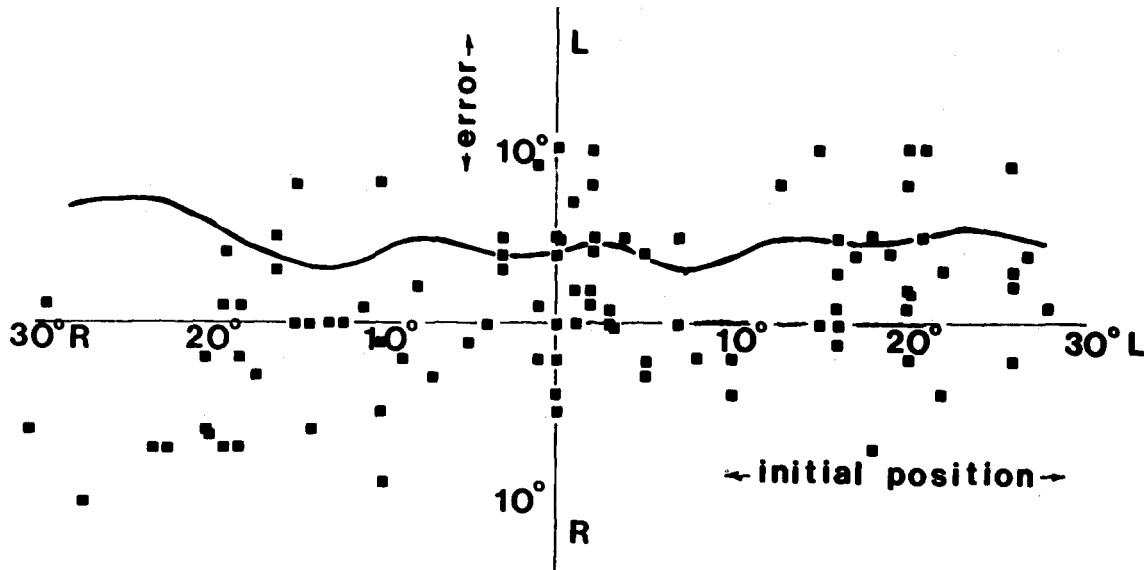


Fig. 3. Plot of error of saccades to auditory targets versus initial position of the eye at the start of the saccades. Line is plot of the root mean square of the errors, showing that the error is independent of initial position

common reference system, the results imply a remapping from one coordinate system into another.

This last point is somewhat easier to see if consideration is restricted to the head fixed experiment, where the motor response was simply a saccadic eye movement. For those responses in which the eye is pointed straight ahead at the onset of the auditory stimulus, the size of an accurate saccade is just equal to the displacement from straight-ahead of the sound source. For all other initial eye positions, however, the size of an accurate saccade is equal to the displacement of the auditory target minus the initial position of the eye. If one subscribes to the idea that the motor program for saccades specifies the size and direction of the saccade, the fact that Fig. 3 displays points clustered about the horizontal axis, rather than showing error directly proportional to the initial eye position, forces one to conclude that the auditory input is remapped, by subtraction of the eye's initial position, into a saccadic motor command. If one chooses instead to argue that saccades are specified in terms of their final position, then our results can be explained by noting that saccade final position and the auditory target position are both head-centered and do not need to be remapped. However, now saccades to visual targets must be remapped. Here initial eye position information is needed to convert the retinal error of a visual target into the appropriate final position command. So, however, saccades are specified by the motor program, our results demonstrate that initial eye position must be used to remap from retinocentric to

head-centered coordinates or vice versa. Given this, do our results favor either final position or size specification of saccades?

One possible clue is the difference in response latencies between visual and auditory targets. It is tempting to argue that the longer latency of the response to visual stimuli is the result of the remapping process. Such an assumption cannot be justified, since, depending on the intensity of the stimulus and adaptation, a delay of approximately 20–100 ms occurs between the time light strikes the retina and the onset of activity in the retinal ganglion cells (Gouras 1967). There is no similar delay in the auditory system, where activity in primary auditory afferents follows an auditory stimulus by less than even 2 ms (Rupert et al. 1963). These differences in the rapidity of the transduction mechanism preclude making any deductions about the relation of latencies to remapping in our experimental arrangement.

So far, the discussion has centered on eye movements, but what part does the head play? As is the case with coordinated movements to visual targets, in movements to auditory targets the head plays a secondary role. The motor strategy is to get the eye on target first, and then to subtract out the influence of any head movement through the use of the vestibulo-ocular reflex (VOR). The head does, of course, move toward the auditory targets, but, because of the VOR, this movement does not affect the accuracy of the gaze. The contribution of the head could not, however, be disregarded in the case of say, pointing with the arm. To indicate a sound

source with the arm, the position of the head with respect to the torso must be known. Therefore, in pointing to a sound, there must be another remapping between a body centered reference frame and a head centered one.

Finally, the lack of difference in response between punctate and continuous auditory stimuli shows that the eye-head movement strategy is not a searching movement controlled by auditory feedback, but rather a triggered movement program.

Zahn et al. (1978) investigated the ability of human subjects to make accurate saccadic eye movements to auditory targets with head fixed. According to these investigations, the eye movements elicited by auditory stimuli are less accurate and had longer latencies than those made to visual stimuli. While it is conceivable that monkeys and humans may vary on this point, it may also be that we are looking at different types of behavior. In Zahn et al.'s experiment the instruction to the subjects was to be as accurate as possible, whereas in our experiment the targets had to be acquired only within $\pm 5^\circ$. Therefore, the monkeys may have been behaving in a reflexive manner while human subjects may have been doing something more measured, possibly involving higher neural control mechanisms. In a study on cats (Thompson and Masterton 1978), head orienting reflex to an unexpected tone burst of white noise occurred with latencies varying between 20–80 ms. This is remarkably shorter than the latencies of 122–177 ms measured for head movements in our monkeys (Table 2). These discrepancies stress the importance of the behavioral conditions which might involve different pathways, and the possibility of inter-species variation.

However, for all the studies on eye and head movements to auditory stimuli, two main questions remain still open: where does the coordinate remapping occur and how is it done? Whereas the transformation mechanisms remain for the moment unsolved, there are some clues concerning the putative structures involved. A likely candidate, as pointed out by Zahn et al. (1978), is the superior colliculus where some neurons respond to moving visual and auditory stimuli with almost overlapping receptive fields (Gordon 1973). Results from lesion experiments also stress the importance of the inferior colliculus in auditory localization (Strominger and Oesterreich 1970), and more recently, Thompson and Masterton (1978) conclude that initiation and direction of reflexive head movements towards a sound source is mediated by pathways ascending to the vicinity of the inferior colliculus, whereas their accuracy involves pathways bypassing the inferior colliculus. Most probably the superior olivary com-

plex, which gets information from both ears (Brugge and Geisler 1978), is also implicated in the auditory localization, but the evidence for its role in auditory-visual interaction is still very indirect (Harrison and Irving 1966).

It would be interesting to know if an auditory map exists in one of the structures mentioned previously, and if the neuronal activity there is modulated by the initial eye position or if incorporation of this information occurs only in the immediate premotor structures to eye and head movements – as, for instance, the paramedian pontine and the medullary reticular formation.

References

- Bizzi E, Kalil RE, Tagliasco V (1971) Eye-head coordination in monkeys: Evidence for centrally patterned organization. *Science* 173: 452–454
- Brugge JF, Geisler CD (1978) Auditory mechanisms of the lower brainstem. *Ann Rev Neurosci* 1: 363–394
- Gordon B (1973) Receptive fields in deep layers of cat superior colliculus. *J Neurophysiol* 36: 157–178
- Gouras P (1967) The effects of light-adaptation of rod and cone receptive field organization of monkey ganglion cells. *J Physiol (Lond)* 192: 747–760
- Hallett P, Lightstone A (1976) Saccadic eye movement to flashed targets. *Vision Res* 16: 107–114
- Harrison JM, Irving R (1966) Visual and non-visual auditory systems in mammals. *Science* 154: 738–743
- Heffner H, Masterton B (1975) Contribution of auditory cortex to sound localization in the monkey. *J Neurophysiol* 38: 1340–1358
- Mills AW (1958) On the minimum audible angle. *J Acoust Soc Am* 30: 237–246
- Morasso P, Bizzi E, Dichgans J (1973) Adjustment of saccade characteristics during head movements. *Exp Brain Res* 16: 492–500
- Rupert A, Moushegian G, Galambos R (1963) Unit responses to sound from the auditory nerve of the cat. *J Neurophysiol* 26: 449–456
- Stevens SS, Newman EB (1936) The localization of actual sources of sound. *Am J Psychol* 48: 297–306
- Strominger N, Oesterreich R (1970) Localization of sound after section of the brachium of the inferior colliculus. *J Comp Neurol* 138: 1–18
- Thompson GC, Masterton RB (1978) Brain stem auditory pathways involved in reflexive head orientation to sound. *J Neurophysiol* 41: 1183–1202
- Thurlow WR, Mangels JW, Runge PS (1967) Head movements during sound localization. *J Acoust Soc Am* 42: 489–493
- Thurlow WR, Runge PS (1967) Effect of induced head movements on localization of direction of sounds. *J Acoust Soc Am* 42: 480–488
- Wallach H (1939) On sound localization. *J Acoust Soc Am* 10: 270–274
- Zahn JR, Abel LA, Dell'Osso LF (1978) Audio-ocular response characteristics. *Sens Processes* 2: 32–37