Auditory Saccades From Different Eye Positions in the Monkey: Implications for Coordinate Transformations

Ryan R. Metzger, O’Dhaniel A. Mullette-Gillman, Abigail M. Underhill, Yale E. Cohen, and Jennifer M. Groh

Department of Psychological and Brain Sciences, Center for Cognitive Neuroscience, Dartmouth College.
Hanover, New Hampshire 03755

Submitted 31 March 2004; accepted in final form 25 May 2004

INTRODUCTION

The intermediate and deep layers of the superior colliculus (SC) play an important role in guiding saccadic eye movements to the locations of both auditory and visual stimuli. Experiments involving visually guided saccades, microstimulation, and reversible inactivation have suggested that the SC provides a motor command signal specifying the location of the target with respect to the eyes (Lee et al. 1988; Mays and Sparks 1980; Robinson 1972; Schiller and Stryker 1972). However, it is not clear how auditory information is transformed into a motor command appropriate to move the eyes to look toward the sound-source location. Initially, auditory stimuli are coded in a head-centered frame of reference based on differences in the time of arrival and the level of a sound at the two ears (i.e., binaural cues) and on the frequency- and location-dependent filtering (i.e., monaural spectral cues) of the sound by the head and ears (for review, see Blauert 1997). This information is not sufficient to guide a saccadic eye movement to a sound-source location because the pattern of muscle contraction needed to bring the eyes to the location of a target depends on both the target’s location and the initial position of the eyes in the orbits. Moreover, auditory sensory activity in the SC does not appear to be appropriate either: the SC’s auditory receptive fields shift when the eyes move, but, on average, they shift only about half as far as the eyes move (Jay and Sparks 1984, 1987b). If these signals feed in to the brain stem saccadic pulse-step generator unaltered, then one would expect that saccades to sounds would fail to fully compensate for initial eye position.

Currently little is known about the effects of initial eye position on auditory saccades in monkeys. Although some studies have suggested that monkey auditory saccades are accurate (Jay and Sparks 1990), other studies have disagreed (Grunewald et al. 1999), and the contribution of initial eye position to auditory saccade accuracy has not been studied in quantitative detail (but see Whittington et al. 1981). The question has come to be of particular interest because eye position has been shown to influence localization behavior in humans (Lewald 1997, 1998; Lewald and Ehrenstein 1996a,b) as well as neural activity in many brain regions in monkeys (for review, see Salinas and Thier 2000; see also Groh et al. 2001; Werner-Reiss et al. 2003). The effects of eye position on neural activity in monkeys have been invoked as the underpinnings of the observed behavioral effects in humans. However, the behavior of humans and monkeys can differ: for example, humans show approximately complete compensation for initial eye position in making saccades to tactile stimuli, but monkeys show only partial compensation (Groh and Sparks 1996). Accordingly, we sought to determine whether monkeys compensate for initial eye position when making saccades to sounds.

METHODS

Two adult female rhesus monkeys (monkeys C and X) were subjects for these experiments. All animal procedures were conducted in accordance with the principles of laboratory animal care of the National Institutes of Health (Publication No. 86–23, revised 1985) and were approved by the Institutional Animal Care and Use Committee at Dartmouth. The monkeys underwent a surgery using general anesthesia and aseptic techniques to implant a head post for restraining the head and a scleral eye coil for monitoring eye position (Judge et al. 1980; Robinson 1963).

The experiments were conducted in complete darkness in a single-walled sound-attenuation chamber (IAC) lined with sound-absorbent foam (3-in painted Sonex One) to reduce echoes. The monkeys were seated comfortably in a conventional plastic primate chair (Crist Instruments) with a standard neckplate. The supports for immobilizing the head were located behind the interaural axis. The monkeys faced an array of light-emitting diodes (LEDs) and loudspeakers (Audax...
TWO25V2) located 57 in away. The auditory stimulus was a band-limited white noise (500 Hz to 18 kHz, 10-ms onset ramp, variable duration) at 50 dB SPL (“A” weighting, Bruel and Kjaer, Model No. 2237 integrating sound level meter with Model No. 4137 condenser microphone) measured at the location normally occupied by the monkey’s head. The input signal to each loudspeaker was adjusted to compensate for subtle differences between sound levels from each speaker (± 2 dB SPL). Eye position was sampled at 500 Hz.

Our training method capitalized on the innate tendency of monkeys to orient to unexpected sounds (Gifford et al. 2003). Monkeys were trained first to make visually guided saccades. Before and during this training period, they were never exposed to any sounds from the loudspeakers in the experimental apparatus. Once they were well accustomed to this visual task and to being in the silent experimental apparatus, we began training on a simple auditory saccade task. As reported by Gifford et al. (2003), monkeys oriented readily to the very first presentation of a sound, presumably because it was unexpected and they were startled by it. We were careful to reinforce these first few trials, before the monkeys habituated to the novel sound. In the first few days or weeks of training, and occasionally thereafter, we paired the auditory stimulus with a visual stimulus at the same location. This visual reinforcer came on after sound onset if the animal failed to look to the sound alone. Initially, we allowed the animal to look to this light-sound pairing to receive a reward but gradually we reduced the time allowable to look to the paired light-sound to force the animal to look to the sound alone. With this training regimen, we had little trouble eliciting auditory saccades from monkeys. Our methods contrast with those of Grunewald et al. (1999), who had difficulty training monkeys to perform auditory saccades when they were pretrained not to make auditory saccades, i.e., when they had been previously required to maintain fixation while sounds were presented. Before the data for this study were collected, the two monkeys included here had been performing auditory saccade tasks for the purpose of other unrelated neurophysiological experiments for 2 (monkey X) or 4 yr (monkey C).

During experimental sessions, monkeys performed an “overlap saccade task.” In this task, monkeys fixated an LED (the fixation LED) at one of three randomly presented locations (−12, 0, or 12° horizontally and 17° below the horizontal meridian). After a variable length of time (900–1,300 ms), a target (either visual or auditory) was randomly presented from one of nine locations across the horizontal meridian (−23 to 23° in 5–6° increments). The monkey maintained its gaze at the fixation LED for a variable length of time (600–900 ms) until the LED was turned off, cueing the monkey to saccade to the target to receive a reward. The target stayed on until after this saccade occurred. The sizeable vertical separation between the fixation and saccade targets ensured that a saccade was required for all fixation/target combinations and to avoid overlap with the target acceptance window. We used large target acceptance windows (16–17° in radius) for reasons explained below. With these acceptance windows, 96% of properly initiated trials (i.e., trials in which the monkeys maintained their gaze at the fixation LED until its offset) were rewarded in the experimental sessions. Data included here are from five experimental sessions per monkey with an average of 12 trials per condition (3 initial eye positions × 9 target locations × 2 target modalities) presented in each session. Visual and auditory trials were randomly interleaved during each session.

All properly initiated trials, rewarded and unrewarded, were included for data analysis provided a saccade (with any direction and amplitude) began >100 and <400 ms after the offset of the fixation LED. Saccade onset was defined as the time when eye-movement velocity exceeded 20°/s (velocity had to remain above this threshold for 20 ms to be considered a saccade), and offset was the time when velocity fell <20°/s (for a minimum of 10 ms; Fig. 1). With these inclusion criteria, 88% of all properly initiated trials were included in the final analysis. These generous inclusion criteria minimized experimental bias: limiting our analysis to saccades that landed within target-centered reward windows could potentially introduce a bias in favor of successful compensation for initial eye position. Undoubtedly, some trials in which the monkeys made no attempt to look to the sounds were included. These trials would tend to inflate the observed variability in saccade endpoints but would not introduce a bias either for or against an eye-position effect.

**RESULTS**

The effects of initial eye position on the horizontal accuracy of saccades to visual and auditory targets are shown in Figs. 2 and 3. In Fig. 2, each line represents the mean saccadic endpoints from one initial eye position (left, center, or right). Two important points can be taken from this figure. First, the...
fact that these three lines are largely superimposed suggests that the monkeys compensated appropriately for horizontal changes in initial eye position for both auditory and visual saccades. Second, even though virtually all trials were included in this analysis (see METHODS) and despite the large acceptance windows, the average location of the auditory saccade endpoints was quite close to the true locations of the targets. Figure 3 confirms this basic pattern by showing all of the raw saccade error data for both visual and auditory saccades from each of the three initial eye positions.

How complete was this compensation for initial eye position? The effects of initial eye position did reach statistical significance for both auditory and visual saccades (2-way ANOVA conducted separately for each monkey and target modality, with target location and initial eye position as the factors, P < 0.05). It was also significant using multiple linear regression (Table 1). However, it is known that saccades tend to fall short of an intended target by ~10% of the distance to that target (for review, see Becker 1989). These effects of eye position could therefore reflect the normal hypometria of saccades.

To determine whether the modest but significant effect of eye position that we observed was comparable in size to the predicted 10% undershoot, we calculated the average horizontal difference observed in the endpoints for a 12° horizontal shift in initial eye position. This average endpoint difference was calculated for each target location for the center versus left eye positions and the center versus right eye positions and then averaged together. The average endpoint differences for auditory saccades were 0.6° (monkey C) and 1.6° (monkey X) or 5–13% of the 12° difference in fixation position (Table 1). Shifts for visual endpoints were only slightly smaller: 0.4 and 0.9 or 3–8%, respectively. These effects were comparable to the normal 10% undershoot. We therefore consider the monkeys’ compensation for changes in initial eye position to be virtually complete.

Although initial eye position did not vary vertically, we examined whether the horizontal shifts in initial eye position had an effect on the vertical accuracy of saccades to both the auditory and visual targets located along the horizontal meridian (0° vertically). While vertical localization of the visual targets was quite accurate (mean ± SE vertical endpoint across all initial eye positions for monkey C = 0.84 ± 0.05°, monkey X = 0.04 ± 0.06°), the monkeys differed in their vertical accuracy for auditory targets (mean vertical endpoint across all initial eye positions for monkey C = 7.60 ± 0.09°, monkey X = −2.90 ± 0.12°; Fig. 3). We note that these errors are smaller than the strong upward bias observed for saccades to the remembered locations of visual targets by Barton and Sparks (2001). The average vertical differences (shifts) observed in the endpoints for a 12° rightward shift in initial eye position were −0.08° for both visual and auditory targets for monkey C and −0.17° (visual) and −0.25° (auditory) for monkey X. We therefore consider the monkeys’ vertical saccadic endpoints to be essentially unaffected by horizontal changes in initial eye position.

**DISCUSSION**

We have shown that monkeys can produce eye movements to sounds the endpoints of which are only minimally influenced by initial eye position. Our results are consistent with related
studies of auditory saccades in humans (Frens and van Opstal 1994; Yao and Peck 1997; Zahn et al. 1979; see also Makous and Middlebrooks 1990) and extend these findings to monkeys (see also Whittington et al. 1981 for a brief consideration of this question). The modest influence of eye position that we observed is in some respects similar to that reported by Lewald and Ehrenstein for sound-location discrimination tasks in humans (Lewald 1997, 1998; Lewald and Ehrenstein 1996a,b). However, the eye position effect in our study was similar for visual and auditory saccades, whereas Lewald et al. reported differences between visual and auditory localization judgments. Our findings differ from those of Groh and Sparks, who showed a more sizeable effect of initial eye position on somatosensory saccades in monkeys, although not in humans (Groh and Sparks 1996).

Our findings are surprising in light of the failure of auditory receptive fields in the SC to fully compensate for changes in eye position, unlike visual receptive fields which do show full compensation (Jay and Sparks 1987b). How can the visual and auditory representations in the SC be different and yet produce similar saccades? There are four classes of answers to this question. First, the coordinate transformation of auditory signals from head- to eye-centered coordinates is completed before the SC is “read out,” so that at read-out time the visual and auditory representations are no longer in different coordinates. Second, the representations remain different, but the read-out algorithm is robust to these differences. Third, the representations are read out using different algorithms, and these different algorithms rectify the discrepancies between the visual and auditory representations. Fourth, the SC is not involved in generating auditory saccades after all. We will elaborate on these possibilities in more detail.

The first possibility is that the discrepancy between the visual and auditory reference frames is resolved before the SC is read out. In comparing the visual and auditory representations in the SC, Jay and Sparks (1987b) considered only the activity in a fixed window of time locked to the sensory stimulus well in advance of the movement. It is possible that for auditory saccades, the frame of reference becomes fully eye-centered by the time of the movement. The motor-related portion of the activity would thus be expected to show full compensation for changes in eye position. Jay and Sparks did not analyze the shifting of the motor fields. They did report that many saccade-related burst neurons were active for both visual and auditory saccades (see also Cohen and Andersen 2000; Jay and Sparks 1987a; Linden et al. 1999; Russo and Bruce 1994) for related findings in frontal eye fields and intraparietal cortex) but these data do not fully answer the question because the SC has sufficiently large movement fields that there would be considerable overlap between the visual and auditory movement fields even if they remain in different frames of reference.

The second possibility is that the visual and auditory representations remain different, but the mechanism by which the SC is read out is robust to these differences because it factors in both the site and level of activity. Previous work has shown that both the site of activation in the SC (Lee et al. 1988; Robinson 1972) as well as the level of activation at that site are important for determining the resulting saccade vector (Stanford et al. 1996). Variations in the level of auditory activity with eye position (an eye-position gain field) could therefore serve to complete the compensation for changes in eye position. Indeed, such eye-position gain fields have been demonstrated for visual saccade-related activity, (Van Opstal and Hepp 1995; Van Opstal et al. 1995), but these models are motivated by the SC’s representation of visual saccades and, as we have already mentioned, the visual and auditory representations appear to be different.

The same answer (same saccade vector) could be obtained either through a shift in the site of activity or a change in the level of activity or a combination of both.

Third, it is possible that the SC is simply read out differently for visual and auditory saccades. Either a head-centered signal or an eye-centered signal specifying target location can in principle be successfully combined with information about current eye position to generate the appropriate pattern of force to move and hold the eyes at the appropriate position. Indeed, Robinson’s original model for the saccadic pulse-step generator called for a head-centered signal of target location as the input (Robinson 1975). More recent versions call for the input signal to be an eye-centered signal of target location (e.g., Jurgens et al. 1981; Scudder 1988; for review, see Van Opstal et al. 1995), but these models are motivated by the SC’s representation of visual saccades, and, as we have already mentioned, the visual and auditory representations appear to be different.

J Neurophysiol • VOL 92 • OCTOBER 2004 • www.jn.org

---

### TABLE 1. Results of multiple regression and average endpoint difference analysis

<table>
<thead>
<tr>
<th>Monkey C</th>
<th>Coefficients</th>
<th>Partial Correlation</th>
<th>Regression Parameters</th>
<th>Visual</th>
<th>Auditory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$c_1$</td>
<td>$c_2$</td>
<td>$c_3$</td>
<td>$p_1$</td>
<td>$p_2$</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
<td>0.04</td>
<td>0.00</td>
<td>0.99</td>
<td>0.20</td>
</tr>
<tr>
<td>Monkey X</td>
<td>0.88</td>
<td>0.09</td>
<td>-1.80</td>
<td>0.96</td>
<td>0.20</td>
</tr>
</tbody>
</table>

The multiple regression equation was: final eye position = $c_1$*target position + $c_2$*initial eye position + $c_3$. Although small, all values of $c_2$ differed significantly from 0 ($P < 0.05$), indicating a modest but significant influence of initial eye position on the final endpoint. For monkey C but not monkey X, the 95% confidence intervals for the visual and auditory values of $c_2$ overlapped with one another. All values of $c_1$ differed significantly from 0 ($P < 0.05$). All values of $c_3$ except monkey C’s visual $c_3$ differed significantly from 0 ($P < 0.05$). The partial correlations of target position and initial eye position with final eye position are given in the columns labeled $p_1$ and $p_2$, respectively. See RESULTS for additional details.
Precedence for the idea that the SC might provide a saccade command signal that is not limited to strictly eye-centered information comes from work by van Opstal and colleagues (Van Opstal and Hepp 1995; Van Opstal et al. 1995). As noted in the preceding text, Van Opstal et al. have demonstrated that changes in initial eye position do affect the magnitude of the saccade in the preceding text, Van Opstal et al. have demonstrated that (Van Opstal and Hepp 1995; Van Opstal et al. 1995). As noted in this dual-read-out view with the results of microstimulation can in principle be read out successfully (Groh et al. 2001; Pouget and Sejnowski 1997; Van Opstal and Hepp 1995; Van Opstal et al. 1995), behavioral confirmation of this in non-human primates has been lacking. Our results provide support for this view.

The main sequence, a tool for studying sensory information into a novel frame of reference (e.g., Cohen and Andersen 2000), in other instances, the coordinate transformation seems to remain incomplete. Although there has been ample modeling work to suggest that these representations appear to reflect a full transformation of sensory information into a novel frame of reference. Not only is it of interest to study the SC’s representation of auditory space in greater detail, but also the assumption that the premotor circuitry is accessed in the same way for auditory and visual saccades needs to be tested. Recording the activity of brain stem excitatory burst neurons (EBNs) or other saccade-related neurons in the brain stem reticular formation and/or cerebellum during both visual and auditory saccades would shed important light on this question.

Since the seminal work of Jay and Sparks in the SC, many other regions of the primate brain have been found to combine information about eye position with their sensory responses to either visual or auditory stimuli (Andersen and Zipser 1988; Andersen and Mountcastle 1983; Andersen et al. 1985; Boussaoud and Bremmer 1999; Bremmer et al. 1997a,b, 1998, 1999; Cohen and Andersen 2000; Duhame et al. 1997; Galletti et al. 1993; Groh et al. 2001; Guo and Li 1997; Jay and Sparks 1984 1987b; Mullette-Gillman et al. 2003; Russo and Bruce 1994; Squatrito and Maioli 1996; Strianne et al. 1996; Trotter and Celebrini 1999; Werner-Reiss et al. 2003; Weyand and Malpeli 1993). In many of these neurons, the resulting activity conveys a blend of information about the external sensory stimulus and the eye-position signal. Although some of these representations appear to reflect a full transformation of sensory information into a novel frame of reference (e.g., Cohen and Andersen 2000), in other instances, the coordinate transformation seems to remain incomplete. Although there has been ample modeling work to suggest that these representations can in principle be read out successfully (Groh et al. 2001; Pouget and Sejnowski 1997; Van Opstal and Hepp 1995; Van Opstal et al. 1995), behavioral confirmation of this in non-human primates has been lacking. Our results provide support for this view.

**Acknowledgments**

We thank A. S. Clark, K. A. Kelly, S. G. Lisberger, D. L. Sparks, and U. Werner-Reiss for helpful comments.

**Grants**

This research was supported by the Alfred P. Sloan Foundation (J. M. Groh), McKnight Endowment Fund for Neuroscience (J. M. Groh), Whitehall Foundation (Y. E. Cohen and J. M. Groh), John Merck Scholars Program (J. M. Groh), Office of Naval Research Young Investigator Program (J. M. Groh), EJLB Foundation (J. M. Groh), National Institutes of Health Grant NS-17778-19 (Y. E. Cohen and J. M. Groh), NIH B/START and Shannon Awards (Y. E. Cohen), NIH Grant DC-05292 (R. R. Metzger), and The Nelson A. Rockefeller Center at Dartmouth (J. M. Groh).

**References**


