CHAPTER 15

Looking at Sounds: Neural Mechanisms in the Primate Brain

Jennifer M. Groh and Dinesh K. Pai

When you hear a salient sound, it is natural to look at it to find out what is happening. Orienting the eyes to look at sounds is essential to our ability to identify and understand the events occurring in our environment. This behavior involves both sensorimotor and multisensory integration: A sound elicits a movement of the visual sense organ, the eye, to bring the source of the sound under visual scrutiny. How are auditory signals converted into oculomotor commands? This chapter describes our recent work concerning the necessary computational steps between sound and eye movement, and how they may be implemented in neural populations in the primate brain.

In principle, the brain must determine the location of the sound, encode that location in a reference frame and format that allows for convergence with visual signals onto a common motor pathway, and create a suitable timevarying signal in the extraocular muscles to move the eyes. In practice, it is not clear exactly how these computations unfold. Several specific hurdles must be overcome. First, auditory and visual signals arise in different reference frames. Binaural and spectral cues provide information about where a sound is located, but only with respect to the head and ears, not the eyes. In contrast, visual information is intrinsically eye centered: The pattern of illumination of the retina depends on the locations of objects in the visual scene with respect to the direction of gaze. These two reference frames vary in their relationship to each other depending on the orbital position of the eyes (Fig. 15.1). This discrepancy in reference frame should be resolved prior to or as part of the convergence of visual and auditory signals onto a common oculomotor pathway.

A second computational hurdle is that visual and auditory signals are not necessarily encoded in the same format. From the retina on, neurons in the early visual pathway have receptive fields that tile the visual scene and produce a "place code" for stimulus location (Fig. 15.2). In contrast, the binaural comput s performed in the auditory pathway do not necessarily produce receptive fields. If they do not, then there may be a discrepancy in the coding format of visual and auditory signals.

Ultimately, either visual or auditory or both signals must undergo a transformation into a reference frame and a coding format that are similar to each other and appropriate for accessing the oculomotor pathway. We will begin by describing the evidence concerning the reference frame of auditory signals as they progress from auditory to multimodal and oculomotor areas before turning to coding format and some computational analyses that shed light on the neural algorithms that may be at play in this process.

REFERENCE FRAME

12:08

The earliest area along the auditory pathway where the reference frame of auditory signals



Figure 15.1 Head-centered receptive fields are fixed in space defined with respect to the head. Eyecentered receptive fields are fixed with respect to the eyes. These reference frames shift with respect to each other when the eyes move with respect to the head. Head- and eye-centered reference frames can therefore be distinguished by evaluating the discharge patterns of individual neurons as a function of head- and eyecentered target location across different fixation positions. "Hybrid" response patterns are defined as those that are not well aligned in either head- or eye-centered coordinates.



Figure 15.2 Place codes for space contain neurons with nonmonotonic (peaked) response functions, such as circumscribed receptive fields, whereas rate codes contain neurons with monotonic location sensitivity.

has been investigated is the inferior colliculus (IC). The IC is part of the ascending auditory pathway, receiving input from the superior olivary complex and projecting to the auditory thalamus (medial geniculate body) (Moore, 1991; Nieuwenhuys, 1984; Oliver, 2000). The IC also projects to an oculomotor structure, the superior colliculus (SC) (for review, see Sparks & Hartwich-Young, 1989b), and thus could play a specific role in the control of eye movements to sound sources.

Originally, it was thought that the IC encodes sound location in a head-centered reference frame (Jay & Sparks, 1987a). We tested this

273

12.08

PRIMATE NEUROETHOLOGY

hypothesis by investigating the responses of IC neurons to sounds as a function of eye position (Fig. 15.3). If IC neurons represent sound location in a head-centered reference frame, then eye position should have no impact on neural responses. In contrast, if the IC uses an eyecentered reference frame, the spatial response functions of IC neurons should shift when the eyes move, and by the same amount that the eyes move (e.g., Fig. 15.1). We found that eye position affects the responses of about 40% of IC neurons (Groh et al., 2001; see also Porter & Groh, 2006, and Zwiers et al., 2004). However, we did not find an eye-centered representation:



Figure 15.3 (a, b) Responses of example inferior colliculus (IC) neuron as a function of head- and eyecentered target location for three different fixation positions. Response functions do not align perfectly in either reference frame. (c) Activity of the same neuron in color as a function of all the eye positions and sound locations that were tested. The triangles indicate the fixation positions corresponding to the data in panels a and b. (d) Population plot showing that the reference frame in the population of IC neurons spanned a continuum from more eye centered to more head centered, with most neurons lying between these two canonical extremes. A correlation coefficient between each neuron's response functions in headvs. eye-centered coordinates was calculated and plotted on this graph; crosses indicate 95% confidence intervals. Neurons were classed as eye > head (red) or head > eye (green) only if the confidence intervals show that the eye-centered correlation was greater than the head-centered correlation or vice versa. See Research Design for details. From Porter, K. K., Metzger, R. R., & Groh, J. M. (2006). Representation of eye position in primate inferior colliculus. *Journal of Neurophysiology*, *95*, 1826–1842. Used with permission.

Page:274

The effect of eye position, while statistically significant, interacted with the auditory response but did not cause systematic shifts related to the change in eye position (Fig. 15.3d). Overall, the representation reflected a hybrid of head- and eye-centered information.

The presence of this hybrid reference frame led us to investigate the reference frame at several later stages of processing: auditory cortex and the intraparietal sulcus. The motivation behind these studies was to see if the hybrid representation in the IC was ultimately converted into a more eye-centered representation at a later stage. Auditory cortex and the intraparietal sulcus are not only situated later in the processing stream but also provide direct input to oculomotor structures, in particular the SC (for review, see Sparks & Hartwich-Young, 1989a).

In core auditory cortex, the representation was similar to that of the IC (Fig. 15.4a) (Werner-Reiss et al., 2003). Approximately one third of individual neurons showed a statistically significant influence of eye position on their responses. Across the population, including all neurons regardless of whether they showed a statistically significant effect of eye position, the spatial sensitivity patterns of the majority of neurons reflected a hybrid of head- and eyecentered information.

The lateral and medial banks of the intraparietal sulcus (lateral intraparietal [LIP] and medial intraparietal [MIP] areas) contain both visual and auditory neurons. It had been assumed that the representation of visual information is generally eye centered, with an eye position gain-modulation affecting the response magnitude but not the location of the receptive fields (e.g., Andersen & Mountcastle, 1983; Andersen & Zipser, 1988; Andersen et al., 1985; Zipser & Andersen, 1988), but this view requires a demonstration that the receptive field location does not change with eye position, and systematic mapping of receptive field locations for each different fixation position¹ had not previously been conducted.

Accordingly, we mapped the visual receptive fields in the LIP and MIP areas at multiple fixation positions (Mullette-Gillman et al., 2005). Our results did not support the interpretation of largely eye-centered representation: We found that visual neurons were nearly as likely to have head-centered as eye-centered receptive fields (Fig. 15.4c). Across the population, the distribution of response patterns spanned a continuum from predominantly eye centered to predominantly head centered, with hybrid reference frames being the most common response pattern (Mullette-Gillman et al., 2005).

In keeping with our results in the IC and auditory cortex, we found that the auditory signals in the parietal cortex reflected a mixture of head- and eye-centered sensitivity (Fig. 15.4b). A quantitative examination of the reference frame of across the IC, auditory cortex, and parietal cortex showed that there was little difference between the auditory signals in these structures. There was a small but statistically significant difference between the visual and auditory reference frame within the parietal cortex, suggesting that even though visual and auditory signals converge onto a common neural population (and, in some cases, onto individual bimodal neurons), there remains a slight discrepancy between how visual and auditory information are encoded.

After parietal cortex, visual and auditory signals pass through the SC prior to reaching the eye muscles. The SC is thought to contain a place code for the eye-centered saccade vector, and the same saccade-related burst neurons are thought to control visual, auditory, and somatosensory saccades (Groh & Sparks, 1992, 1996a,b,c; Jay & Sparks, 1984, 1987a,b; Klier et al., 2001; Meredith & Stein, 1983, 1996; Populin et al., 2004; Robinson, 1972; Schiller & Stryker, 1972; Sparks, 1978; Stein & Meredith, 1993; Stein et al., 1993). However, there are some very puzzling aspects to the current story regarding the SC, which call into question some of these assumptions.

In particular, Jay and Sparks investigated the reference frame of both visual and auditory sensory responses in this structure in primates and reported a discrepancy in reference frame: Visual signals were predominantly eye centered whereas auditory signals were intermediate between head- and eye-centered coordinates (Fig. 15.5) (Jay & Sparks, 1984, 1987a,b). Similar results have been reported in the cat SC

PRIMATE NEUROETHOLOGY



Figure 15.4 Reference frame results for auditory and parietal cortex. In A1 (**a**), lateral intraparietal (LIP), and medial intraparietal (MIP) (**b**,**c**) areas, the observed reference frames span a continuum from head- to eye-centered coordinates for both auditory (b) and (in LIP/MIP) visual signals (c). From Mullette-Gillman, O. A., Cohen, Y. E., & Groh, J. M. (2005). Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *Journal of Neurophysiology*, *94*, 2331–2352; and Werner-Reissm U., & Groh, J. M. (2008). A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies. *Journal of Neuroscience*, *28*(14), 3747–3758. Used with permission.

as well (Hartline et al., 1995; Peck et al., 1995; Populin et al., 2004; Zella et al., 2001). Jay and Sparks did not investigate the alignment between the visual and auditory receptive fields of bimodal neurons, but the implication of their reference frame finding is that these receptive fields cannot maintain perfect alignment across different initial eye positions. Although there have been numerous investigations of the response properties of SC neurons to visual, auditory, and combined modality stimuli, suggesting that visual and auditory receptive fields overlap (e.g., Wallace et al., 1996; for review see Stein & Meredith, 1993), these studies have not addressed the effects of eye position and have generally evaluated the receptive fields in a qualitative fashion. Quantitative data on the locations, shape, and alignment of the receptive fields as a function of eye position at the single neuron and population levels are needed.

If the visual and auditory receptive fields of SC neurons are not aligned, and if these neurons control saccadic eye movements, then one would expect a signature of this misalignment in the accuracy of saccades to sounds across different initial eye positions. Specifically, saccades



Figure 15.5 Results of Jay and Sparks's investigation of visual and auditory reference frame in the superior colliculus. (**a**,**b**) Responses of an auditory neuron as a function of head-centered and eye-centered target position: The receptive field shifts with eye position but not by the exact amount of the change in eye position. (**c**) On average, auditory receptive fields shift about half as much (arrow) as would be needed to maintain a constant position in eye-centered coordinates (dashed line). Visual receptive fields tend to shift the full amount, suggesting a lack of registry between visual and auditory tuning. From Jay, M. F., & Sparks, D. L. (1987a). Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. Journal of Neurophysiology, 57, 35–55. Used with permission.

to a given target location might be more or less affected by initial eye position depending on the modality of the target. Assuming visual signals are the "correct" ones, then saccades to visual targets should compensate completely for initial eye position but saccades to auditory targets should show a characteristic pattern of errors suggestive of a failure to complete a coordinate transformation from head- to eye-centered coordinates.

We looked for such an effect and did not find one (Fig. 15.6) (Metzger et al., 2004; see also Peck et al., 1995, and Populin et al., 2004). Instead, we found that visual and auditory saccades were generally very similar to each other (Fig. 15.6). Both showed only a very modest effect of initial eye position on saccade endpoint, although the auditory saccades were more variable. This suggests that, ultimately, the saccade command does not depend very strongly on whether the target was visual or auditory, implying that visual and auditory signals do end up in a common representation.

It is currently uncertain how this could be accomplished. One potential explanation is that visual and auditory signals might be initially misaligned, at the time of the sensory stimulus, but come into alignment prior to the initiation of the movement. If this is the case, then the visual and auditory *saccade-related bursts*, as opposed to the sensory responses studied by

12.08





Figure 15.6 Eye position did not affect the accuracy of saccades to visual and auditory stimuli: The different colored traces, corresponding to saccades endpoints from three different initial fixation positions, are largely superimposed. From Metzger, R. R., Mullette-Gillman, O. A., Underhill, A. M., Cohen, Y. E., & Groh, J. M. (2004). Auditory saccades from different eye positions in the monkey: Implications for coordinate transformations. *Journal of Neurophysiology*, *92*, 2622–2627. Used with permission.

Jay and Sparks, should be in the same reference frame and in spatial alignment. This possibility could also help account for the better correspondence between auditory tuning and saccade vector that was observed in a recent study of the cat SC (Populin et al., 2004): The response window used in that study included motorrelated activity in addition to sensory activity. More research is needed to resolve this issue.

REPRESENTATIONAL FORMAT

The second potential computational challenge for integrating the visual and auditory codes for space is representational format. As noted previously, visual neurons exhibit receptive fields from the very earliest stages of the visual pathway. These receptive fields arise due to the optics of the eye: Light from a given location in the world passes through the aperture of the pupil and illuminates only a restricted portion of the retina. Each photoreceptor can only "see" out in a particular direction. Receptive fields become more complex as signals progress along the visual pathway, but at base, the code for space remains a code in which the location of a visual stimulus can be inferred from the identity of the neurons that are responding to it. This type of code is referred to as a place code, because neurons are often topographically organized according to their receptive field locations.

278

Page:278

In contrast, in the auditory system, spatial location is inferred by comparing cues such as sound arrival time and level across the two ears. What kind of code is produced as part of this computation cannot be determined from first principles. There would seem to be two possibilities: (1) a place code similar to that for visual information, in which auditory neurons have circumscribed receptive fields that tile the auditory scene; the location of a stimulus could then be inferred from knowing which neurons were responding to that stimulus (e.g., Jeffress, 1948), as is the case for visual information; and (2) a rate code in which neurons respond broadly to a wide range of locations, but with a firing rate that varies with sound locations. The location of a stimulus could be inferred by "reading out" the firing rate of the active neurons rather than the identity of the active neurons.

The key difference between these two types of codes is the shape of the tuning function of individual neurons. Do neurons respond only to a restricted range of locations, with different neurons showing different preferences? Or do individual neurons respond broadly, with the maximum responses occurring at the extremes of the possible range of space (e.g., the axis of the contralateral ear) (Fig. 15.2).

We have conducted several studies to assess the coding format in the primate auditory pathway. We developed a statistical assay based on the success of Gaussian and sigmoidal functions at fitting the responses as a function of sound location. The idea is that Gaussian functions would be substantially better than sigmoids at fitting the response patterns if the neurons had nonmonotonic spatial response functions characteristic of receptive fields and a place code, but that either sigmoids or broad half-Gaussians would be successful at fitting monotonic tuning patterns characteristic of a rate code (Fig. 15.7).

To our knowledge, nothing is known about the coding of spatial location in the primate auditory pathway prior to the level of the IC. The IC itself is known to contain spatially sensitive neurons (Groh et al., 2001, 2003; Zwiers et al., 2004). We evaluated the spatial sensitivity of IC neurons to determine whether they have circumscribed receptive fields tiling the auditory scene. Instead, we found that they showed consistent preferences for locations along the axis of the contralateral ear. This pattern is characteristic of a rate code for sound location (Fig. 15.8a) (Groh et al., 2003).

We found similar results in auditory cortex (Fig. 15.8b) (Werner-Reiss & Groh, 2008). Interestingly, the code was less smooth in auditory cortex than in IC: Individual neurons often had "bumpy" response functions that were broadly tuned for the contralateral ear, but also had other sound locations that they also responded well to. One possible reason for this is that there could be a transformation from rate code to a place code as the auditory signals approach or join with visual signals. If this is the case, then neurons in brain regions such as parietal cortex or the superior colliculus might show circumscribed receptive fields. Quantitative information on the representational format of auditory signals in these structures is currently lacking in primates.

It will be interesting to determine whether auditory signals are ultimately translated into a place code. The chief advantage of this would be to facilitate integration with placecoded visual information. However, other than that, the advantages might be few. Place codes are better than rate codes for encoding many locations simultaneously, but the auditory system may not be able to encode large numbers of sound locations. Perceptually, two very similar simultaneous sound sources tend to be perceived at an intermediate location (summing localization) (for review see Blauert, 1997). Furthermore, if signals are rate coded at one stage and converted into a place code at a later stage, it is not clear that the benefits of place coding would then accrue, as the rate-coding stage would serve as an information processing bottleneck that would prevent subsequent place-coding stages from representing multiple simultaneous stimulus locations.

A possible advantage for retaining auditory spatial information in a rate code is that it may facilitate interactions with eye position





Figure 15.7 Simulations of place and rate codes and how Gaussian and sigmoidal curve fits can be used to distinguish between these representational formats. (a) Simulation of three Gaussian-tuned neurons, showing both Gaussian and sigmoidal curve fits. (b) Simulation of three sigmoidal neurons. The Gaussian and sigmoidal curve fits are so similar as to obscure each other. (c) Population plot of the correlation coefficients of Gaussian and sigmoidal curves for a population of individual neurons whose underlying tuning functions were Gaussian. Gaussian curves were always successful at fitting such response patterns; sigmoidal functions became increasingly successful as the eccentricity (i.e., the absolute value of the azimuthal location) of the Gaussian peak increased. (d) Same as c, but for a population of individual neurons whose underlying tuning functions were sigmoidal. Both Gaussian and sigmoidal curves were successful at fitting such response patterns. From Werner-Reiss, U., & Groh, J. M. (2008). A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies. *Journal of Neuroscience*, *28*(*14*), 3747–3758. Used with permission.

information, which appears to be encoded in a similar format. We found that in the IC, eye position sensitivity is generally monotonic (e.g., Fig. 15.3), consistent with a rate code for eye position. This finding is consistent with studies in the parietal cortex (Andersen et al., 1990), frontal eye fields (Bizzi, 1968), cerebellar flocculus (Noda & Suzuki, 1979), somatosensory cortex (Wang et al., 2007), and premotor circuitry of the oculomotor pathway (Keller, 1974; Luschei & Fuchs, 1972; McCrea & Baker, 1980; Sylvestre & Cullen, 1999a).

MOTOR COMMANDS

12:08

What is the reference frame and representational format of the motor command? The pattern of force needed to move the eyes to look in a particular direction reflects a combination of reference frames and a combination of representational formats. For a movement in a given direction, the amount of force that needs to be applied varies monotonically with the size of the movement, consistent with a rate code. The direction of the movement is controlled by the

280





Figure 15.8 Representational format in inferior colliculus (IC) and core auditory cortex. (a) In the IC, most spatially sensitive neurons responded in a graded, monotonic fashion peaking for sounds along the axis of the contralateral ear, as shown for this example neuron (left panel). Across the population, this pattern is evident in the fact that sigmoidal functions were as good as Gaussians at capturing the response patterns From Groh, J. M., Kelly, K. A., & Underhill, A. M. (2003). A monotonic code for sound azimuth in primate inferior colliculus. *Journal of Cognitive Neuroscience, 15*, 1217–1231. (b) The pattern of results was similar in auditory cortex, although some individual neurons had "bumpy" response functions (data points lie slightly above the line of slope one in the right panel). From Werner-Reiss, U., & Groh, J. M. (2008). A rate code for sound azimuth in monkey auditory cortex: implications for human neuroimaging studies. *Journal of Neuroscience, 28(14)*, 3747–3758. Used with permission.

ratio of activation in different muscle groups, a format that is more akin to a place code: Which muscles are active controls the movement direction.

The reference frame of oculomotor commands is referred to as eye centered by some sources and head centered by others. The confusion stems from both what is meant by motor command—is this term properly reserved only for the extraocular motor neurons or may it be applied to slightly earlier stages such as the SC?—as well as a lack of quantitative investigation into this question. We favor reserving the term "motor command" for the signals carried by the extraocular motor neurons. The reference frame of extraocular motor neurons has not been investigated per se, but their discharge patterns are so well characterized using other means that it is possible to draw some inferences. Specifically, the discharge patterns can be described as a linear differential equation (e.g., Sylvestre & Cullen, 1999b):

$$FR = k_1 + k_2P + k_3V,$$

where FR = instantaneous firing rate, P = eye position, and V = eye velocity. This equation illustrates that the firing pattern depends on both initial and final eye position-that is, fixation position as well as the head-centered location of the target. Thus, the motor command cannot be properly formed if premotor circuitry has access only to target location in a single pure reference frame-some combination of headcentered, eye-centered, and eye position information is needed. This suggests that the use of hybrid reference frames at earlier stages of the audio-oculomotor pathway may reflect the constraints of the motor periphery. Most existing models for how the motor command is formed call for separate representations of head- or evecentered information to be combined with eye position information as the time-varying motor discharge neuron pattern is created (Moschovakis, 1996; Van Gisbergen & Van Opstal, 1989), but it might also be possible to generate this command from an input signal that already has these component signals mixed together.

MODELING

Although it remains unclear exactly what kinds of transformations unfold as visual and auditory signals converge onto the oculomotor pathway, it is certainly evident that some transformations between coding formats and reference frames are needed, and it can be fruitful to explore the neural mechanisms that might underlie such transformations while additional experimental studies are pending. Accordingly, we have worked on several models for transforming signals between different reference frames and between different coding formats. We will begin with the models for transformations of coding format, because how information is encoded impacts which algorithms for coordinate transformations may be most appropriate.

Models for Transformations of Coding Format

We have designed several models that involve transformations of signals from either a place code to a rate code or vice versa. Figure 15.9 illustrates several ways that Gaussian tuning functions can be created from a population of neurons with sigmoidal response functions with varying inflection points (Porter & Groh, 2006). Suppose neurons exhibit sigmoidal tuning functions, with some preferring leftward locations (such as spatial neurons in the right IC) and others preferring leftward locations (e.g., the left IC). Assume further that there is a population of neurons whose inflection points vary across the range of space. Excitatory connections from two neurons with opposite tuning preferences and neighboring inflection points would cause a recipient neuron to be responsive to sound locations between the inflection points of the input neurons (Fig. 15.9a).

A circumscribed receptive field could also be created by combining excitatory and inhibitory inputs from two neurons with sigmoidal tuning functions in the same direction, provided once again that their inflection points are appropriately staggered. Suppose two neurons both prefer

OUP • UNCORRECTED PROOF



Figure 15.9 Models for converting signals between place and rate codes. (a) Sigmoidal units with opposite preferences and different inflection points could converge, producing a receptive field. If sound loudness affected the response patterns of the input neurons as indicated, then the receptive field would be larger for louder sounds but would be centered in the same place. (b) Sigmoidal units preferring the same side but with different inflection points could converge using a combination of excitatory and inhibitory synapses. As in panel a, the receptive field would expand for louder sounds if the two input neurons were affected by sound loudness in different directions. (c) An alternative mechanism involving a cascade of units with different thresholds, with inhibitory interneurons. (d–f) Three models for converting place codes to rate codes, with different ways of normalizing for the level of activity. The first model involves no normalization, the second involves full normalization, and the third involves normalization only for high levels of activity. See text for additional details. Groh, J. M. (2001). Converting neural signals from place codes to rate codes. *Biological Cybernetics*, *85*, 159–165; and Porter, K. K., & Groh, J. M. (2006). The other transformation required for visual-auditory integration: Representational format. *Progress in Brain Research*, *155*, 313–323. Used with permission.

leftward locations, but one has an inflection point at 0 degrees and the other has an inflection point at 10 degrees to the right. A recipient neuron that is inhibited by the first neuron and excited by the second neuron will have a receptive field between 0 and 10 degrees to the right—the region of space where only its excitatory input is active (Fig. 15.9b).

Both of these algorithms involve a certain element of place coding in the input stage: The input neurons must have tuning functions whose *inflection* points show heterogeneity spanning the range of possible spatial locations. Thus, this is not a pure rate code at the input. At present, we have not attempted to determine if the response functions are truly sigmoidal as opposed to some other monotonic function, nor have we assessed whether the inflection points span a range of locations. Thus, it is unclear whether these algorithms are biologically plausible or not.

A third algorithm might apply if the input signals are more linear than sigmoidal and if they therefore lack inflection points, much less variation in inflection points. Figure 15.9c illustrates a local circuit with a cascade of thresholds. The thresholds introduce the necessary nonlinearity into the processing of a linear signal to create the receptive fields. Each output neuron (open circles) has both a threshold and an inhibitory interneuron that is paired with it. The inhibitory interneuron has a slightly higher threshold for activation. Thus, the output neuron is active only when its input exceeds its own threshold but is less than the threshold for its matched inhibitory interneuron. This pattern, when repeated with varying thresholds across the population, can create a range of circumscribed receptive fields across the population (Groh & Sparks, 1992).

We have also developed several models for converting signals from a place code to a rate code (Groh, 2001; Porter & Groh, 2006) (Fig. 15.9d–f). The conversion is accomplished using a graded pattern of synaptic weights. The models differ in whether and how they accomplish normalization for the overall level of activity. The vector summation model

PRIMATE NEUROETHOLOGY

(Fig. 15.9d) simply calculates the weighted sum of activity, with no normalization whatsoever. The problem with such a model is that typically there are many other features that might alter neural activity (e.g., the loudness of a sound or the contrast of a visual stimulus), and without normalization the changes in neural responsiveness associated with these features would affect the read-out. There is some perceptual evidence for this kind of effect: For example, low-contrast visual stimuli appear to move more slowly than high-contrast visual stimuli (e.g., Snowden et al., 1998; Thompson et al., 1996), but more generally such factors appear to be corrected for when determining spatial location.

Accordingly, we developed several additional models for converting place codes to rate codes that include normalization for the overall level of activity. The vector averaging model (Fig. 15.9e) has two read-out pathways, one to calculate the sum of activity weighted by its location in the place code (the numerator channel) and the other to calculate the unweighted sum (the denominator channel). Then, the weighted sum is divided by the unweighted sum, producing a signal corresponding to the average location of activity in the place code.

One problem with this model is that it is not clear how neural circuits might implement the division of one number by another. Inhibitory synapses can exert a divisive-like effect, but more generally the nature of the inhibitory influence will vary with the membrane potential: What seems like division when the membrane potential is near rest (e.g., shunting inhibition) might become more like subtraction when the membrane is more depolarized. The vector averaging model requires that the inhibitory influence of the denominator channel should mimic division for a large range of possible numerator and denominator values.

The third model circumvents this problem by implementing normalization in a different fashion. This model, the summation-withsaturation model, calculates a weighted sum of the activity in the input layer, and then clips off any extra activity above a certain threshold. This is accomplished using a combination of neural

integrators and thresholds. The numerator and denominator channels both integrate their input, weighted by location in the case of the numerator channel. When the denominator channel reaches a certain threshold, it clips off the input to the numerator channel. The activity level of the numerator channel will vary only with the location of the input provided there is sufficient activity to trigger the clipping action of the denominator channel, and otherwise will reflect the weighted sum of the input.

This model successfully mimics the pattern of evoked saccades elicited by microstimulation of the SC. Stimulation above a certain frequency evokes saccades that do not depend on the frequency of stimulation (known as the site-specific amplitude), but below that value the amplitude of the saccade falls off as the frequency or duration of stimulation is reduced. The evoked saccade depends on the total number of stimulation pulses delivered until a saturation point is reached (Stanford et al., 1996).

Models for Coordinate Transformations

We have developed two models for coordinate transformations. At the time these models were designed, little was known about either the representational format or the frame of reference of signals in the auditory pathway, so both models assumed that the input consisted of a head-centered map of auditory space. Since this kind of representation has yet to be found, it is worth updating these models to consider other possible forms of input (as well as output).

The vector subtraction model (Fig. 15.10a) begins by converting head-centered, placecoded auditory signals into a rate code so that rate-coded eye position signals could be subtracted. The resulting eye-centered rate code for sound location was then converted into a place code for eye-centered sound location. This model was essentially constructed from the place-to-rate and rate-to-place component parts.

Since it now appears that sound location may be encoded in a rate code, a simpler version of this model can be constructed (Fig. 15.10b). The input can consist of a rate-coded sound location, from which rate-coded eye position information is subtracted. This produces a rate code for eyecentered sound location, just as in the original version. It may or may not be necessary to then convert these signals into a place code, but if it is necessary, one of the rate-to-place algorithms described previously could still be included.

A second type of model, the dendrite model, was originally proposed with the goal of avoiding rate-coding stages in mind (Groh & Sparks, 1992) (Fig. 15.11). The rationale was that the rate-coding stages would limit the number of sound locations that could be encoded to one. Since it now appears that ratecoding stages do exist in the brain's auditory pathways, the motivation behind this model has been reduced. However, it remains uncertain how the brain handles multiple sound locations, so elements of the dendrite model may yet prove to be of some utility.

CONCLUSIONS

To guide an eye movement to the source of a sound requires a net transformation of auditory information from the initially purely head-centered interaural timing and level cues to a reference frame appropriate for controlling the eye muscles. Our studies as well as others have found evidence for hybrid, but not purely eyecentered, frames of reference at several stages of the audio-oculomotor reference frame. This type of hybrid reference frame may be appropriate for controlling saccades because the motor command requires information about both the initial eye position and the desired amplitude of the saccade.

Eye movements to sounds may also require one or more transformations of auditory signals from one kind of coding format into another. At present, we have only found evidence for rate coding of auditory spatial information. It remains to be seen whether ratecoded auditory spatial information is transformed into a place code, and if so, where and how this transformation occurs. It has long been assumed that this transformation does take place, but it should be noted that it is not necessary to create a place code for auditory





B "Vector Subtraction" model: rate code only version



Figure 15.10 Vector subtraction model. <u>Model</u> for transforming auditory signals from head-centered to eye-centered coordinates. This model assumes place-coded auditory inputs. It now appears possible that the auditory inputs encode sound location in a rate-coded format. In this case, the model could be modified to use the rate-coded auditory input signals. From Groh, J. M., & Sparks, D. L. (1992). Two models for transforming auditory signals from head-centered to eye-centered coordinates. *Biological Cybernetics*, *67*, 291–302. Used with permission.

Dendrite model



Figure 15.11 Dendrite model. A second model for transforming auditory signals from head-centered to eye-centered coordinates. Each dendrite receives input from eye position units and one head-centered auditory unit. Thresholds and synaptic weights are balanced so that the cell body receives net excitation via a given dendrite if an auditory stimulus is present in the head-centered receptive field of the auditory unit that activates that dendrite and the eyes are within a certain range of positions. The range of eye positions that allow excitation to reach the soma varies across each dendrite and is matched to the head-centered receptive field of the auditory input that activates a given dendrite. The combination produces tuning for the eye-centered location of the sound. In short, the dendrites of a given unit sample all the possible head-centered locations that could yield a certain eye-centered location, and the eye position inputs filter out that input unless the eye position is in the appropriate range for that dendrite. From Groh, J. M., & Sparks, D. L. (1992). Two models for transforming auditory signals from head-centered to eye-centered coordinates. *Biological Cybernetics*, *67*, 291–302. Used with permission.

information simply to guide an eye movement. Since the motor output consists of a rate code, any place-coded auditory information would have to be converted back into a rate code to generate a motor command.

Of course, auditory signals do not exist solely to trigger eye movements. The perceptual and behavioral endpoints of auditory processing are many and varied, and natural selection has likely produced auditory informationcoding strategies that serve more than one behavioral and perceptual master. Thus, other constraints may account for the aspects of audio-oculomotor transformations that may appear at face value to be inefficient. Further research on whether the behavioral task affects the type of code employed will therefore be of great interest.

NOTE

1. Several studies did map the receptive fields, but sampled primarily along a dimension orthogonal to the direction in which fixation position varied (Andersen et al., 1985; Batista et al., 1999); the effects of eye position on response patterns observed in these studies were consistent with either the headcentered or eye-centered-with-eye-positiongain hypotheses.

REFERENCES

- Andersen, R., & Zipser, D. (1988). The role of the posterior parietal cortex in coordinate transformations for visual-motor integration. *Canadian Journal of Physiology and Pharmacology*, 66, 488–501.
- Andersen, R. A., Bracewell, R. M., Barash, S., Gnadt, J. W., & Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *Journal of Neuroscience*, 10, 1176–1196.
- Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science*, 230, 456–458.
- Andersen, R. A., & Mountcastle, V. B. (1983). The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *Journal of Neuroscience*, *3*, 532–548.
- Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates [see comments]. *Science*, 285, 257–260.
- Bizzi, E. (1968). Discharge of frontal eye field neurons during saccadic and following eye movements in unanesthetized monkeys. *Experimental Brain Research*, 6, 69–80.
- Blauert, J. (1997). *Spatial hearing*. Cambridge, MA: MIT Press.
- Groh, J. M. (2001). Converting neural signals from place codes to rate codes. *Biological Cybernetics*, 85, 159–165.
- Groh, J. M., Kelly, K. A., & Underhill, A. M. (2003). A monotonic code for sound azimuth in primate inferior colliculus. *Journal of Cognitive Neuroscience*, 15, 1217–1231.
- Groh, J. M., & Sparks, D. L. (1992). Two models for transforming auditory signals from headcentered to eye- centered coordinates. *Biological Cybernetics*, 67, 291–302.

PRIMATE NEUROETHOLOGY

- Groh, J. M., & Sparks, D. L. (1996a). Saccades to somatosensory targets. III. Eye-positiondependent somatosensory activity in primate superior colliculus. *Journal of Neurophysiology*, 75, 439–453.
- Groh, J. M., & Sparks, D. L. (1996b). Saccades to somatosensory targets. II. Motor convergence in primate superior colliculus. *Journal of Neurophysiology*, 75, 428–438.
- Groh, J. M., & Sparks, D. L. (1996c). Saccades to somatosensory targets. I. Behavioral characteristics. *Journal of Neurophysiology*, 75, 412–427.
- Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R., & Inati, S. (2001). Eye position influences auditory responses in primate inferior colliculus. *Neuron*, 29, 509–518.
- Hartline, P. H., Vimal, R. L., King, A. J., Kurylo, D. D., & Northmore, D. P. (1995). Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats. *Experimental Brain Research*, 104, 402–408.
- Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, 309, 345–347.
- Jay, M. F., & Sparks, D. L. (1987a). Sensorimotor integration in the primate superior colliculus.
 II. Coordinates of auditory signals. *Journal of Neurophysiology*, 57, 35–55.
- Jay, M. F., & Sparks, D. L. (1987b). Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *Journal of Neurophysiology*, 57, 22–34.
- Jeffress, L. A. (1948) A place theory of sound localization. *Journal of Comparative Physiology and Psychology*, 41, 35–39.
- Keller, E. L. (1974). Participation of medial pontine reticular formation in eye movement generation in monkey. *Journal of Neurophysiology*, 37, 316–332.
- Klier, E. M., Wang, H., & Crawford, J. D. (2001). The superior colliculus encodes gaze commands in retinal coordinates. *Nature Neuroscience*, *4*, 627–632.
- Luschei, E. S., & Fuchs, A. F. (1972). Activity of brain stem neurons during eye movements of alert monkeys. *Journal of Neurophysiology*, 35, 445–461.
- McCrea, R., & Baker, R. (1980). Evidence for the hypothesis that the prepositus neucleus distributes "efference" copy signals to the brainstem. *Anatomy Research*, 196, 122–123.

OUP • UNCORRECTED PROOF

LOOKING AT SOUNDS: NEURAL MECHANISMS IN THE PRIMATE BRAIN

- Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221, 389–391.
- Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*, 75(5), 1843–1857.
- Metzger, R. R., Mullette-Gillman, O. A., Underhill, A. M., Cohen, Y. E., & Groh, J. M. (2004). Auditory saccades from different eye positions in the monkey: Implications for coordinate transformations. *Journal of Neurophysiology*, 92, 2622–2627.
- Moore, D. R. (1991). Anatomy and physiology of binaural hearing. *Audiology*, 30, 125–134.
- Moschovakis, A. K. (1996). Neural network simulations of the primate oculomotor system.II. Frames of reference. *Brain Research Bulletin*, 40, 337–345.
- Mullette-Gillman, O. A., Cohen, Y. E., & Groh, J. M. (2005). Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *Journal of Neurophysiology*, 94, 2331–2352.
- Nieuwenhuys, R. (1984). Anatomy of the auditory pathways, with emphasis on the brain stem. *Advances in Otorhinolaryngology*, *34*, 25–38.
- Noda, H., & Suzuki, D. A. (1979). Processing of eye movement signals in the flocculus of the monkey. *Journal of Physiology*, 294, 349–364.
- Oliver, D. L. (2000). Ascending efferent projections of the superior olivary complex. *Microscopic Research Technology*, *51*, 355–363.
- Peck, C. K., Baro, J. A., & Warder, S. M. (1995). Effects of eye position on saccadic eye movements and on the neuronal responses to auditory and visual stimuli in cat superior colliculus. *Experimental Brain Research*, 103(2), 227–242.
- Populin, L. C., Tollin, D. J., & Yin, T. C. (2004). Effect of eye position on saccades and neuronal responses to acoustic stimuli in the superior colliculus of the behaving cat. *Journal of Neurophysiology*, 92, 2151–2167.
- Porter, K. K., & Groh, J. M. (2006). The other transformation required for visual-auditory integration: Representational format. *Progress* in Brain Research, 155, 313–323.
- Porter, K. K., Metzger, R. R., & Groh, J. M. (2006). Representation of eye position in primate inferior colliculus. *Journal of Neurophysiology*, 95, 1826–1842.

- Robinson, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research*, 12, 1795–1807.
- Schiller, P. H., & Stryker, M. (1972). Single-unit recording and stimulation in superior colliculus of the alert rhesus monkey. *Journal of Neurophysiology*, 35, 915–924.
- Snowden, R. J., Stimpson, N., & Ruddle, R. A. (1998). Speed perception fogs up as visibility drops [letter]. *Nature*, 392, 450.
- Sparks, D. L. (1978). Functional properties of neurons in the monkey superior colliculus: Coupling of neuronal activity and saccade onset. *Brain Research*, 156, 1–16.
- Sparks, D. L., & Hartwich-Young, R. (1989a). The deep layers of the superior colliculus. In: R. H. Wurtz & M. E. Goldberg (Eds.), *The neurobiology of saccadic eye movements* (pp. 213–255). New York: Elsevier.
- Sparks, D. L., & Hartwich-Young, R. (1989b). The deep layers of the superior colliculus. *Reviews of Oculomotor Research*, 3, 213–255.
- Stanford, T. R., Freedman, E. G., & Sparks, D. L. (1996). Site and parameters of microstimulation: Evidence for independent effects on the properties of saccades evoked from the primate superior colliculus. *Journal of Neurophysiology*, 76, 3360–3381.
- Stein, B. E., & Meredith, M. A. (1993). *The merging* of the senses. Cambridge, MA: MIT Press.
- Stein, B. E., Meredith, M. A., & Wallace, M. T. (1993). The visually responsive neuron and beyond: Multisensory integration in cat and monkey. *Progress in Brain Research*, 95, 79–90.
- Sylvestre, P. A., & Cullen, K. E. (1999a). Quantitative analysis of abducens neuron discharge dynamics during saccadic and slow eye movements. *Journal of Neurophysiology*, 82, 2612–2632.
- Sylvestre, P. A., & Cullen, K. E. (1999b). Quantitative analysis of abducens neuron discharge dynamics during saccadic and slow eye movements. *Journal of Neurophysiology*, *82*, 2612–2632.
- Thompson, P., Stone, L. S., & Swash, S. (1996). Speed estimates from grating patches are not contrastnormalized. *Vision Research*, 36, 667–674.
- Van Gisbergen, J. A., & Van Opstal, A. J. (1989).
 Models of saccadic control. In: R. Wurtz & M. E. Goldberg (Eds.), *The neurobiology of saccadic eye movements*. Amsterdam: Elsevier.

Page:289

12:08

289

- Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, 76, 1246–1266.
- Wang, X., Zhang, M., Cohen, I. S., & Goldberg, M. E. (2007). The proprioceptive representation of eye position in monkey primary somatosensory cortex. *Nature Neuroscience*, 10, 640–646.
- Werner-Reiss, U., & Groh, J. M. (2008). A rate code for sound azimuth in monkey auditory cortex: Implications for human neuroimaging studies. *Journal of Neuroscience.* 28(14), 3747–3758
- Werner-Reiss, U., Kelly, K. A., Trause, A. S., Underhill, A. M., & Groh, J. M. (2003). Eye

PRIMATE NEUROETHOLOGY

position affects activity in primary auditory cortex of primates. *Current Biology*, *13*, 554–562.

- Zella, J. C., Brugge, J. F., & Schnupp, J. W. (2001). Passive eye displacement alters auditory spatial receptive fields of cat superior colliculus neurons. *Nature Neuroscience*, 4, 1167–1169.
- Zipser, D., & Andersen, R. A. (1988). A backpropagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, *331*, 679–684.
- Zwiers, M. P., Versnel, H., & Van Opstal, A. J. (2004). Involvement of monkey inferior colliculus in spatial hearing. *Journal of Neuroscience*, 24, 4145–4156.

12.08