Authors: Junchao Hu, Petra Vetter (2023)
Title: How the eyes respond to sounds
Journal: Annals of the New York Academy of Sciences
doi: 10.1111/nyas.15093
Hu, J., & Vetter, P. (2023). How the eyes respond to sounds. Ann NY Acad Sci., 1–19. https://doi.org/10.1111/nyas.15093

CONTENTS

How the eyes respond to sounds

Junchao Hu [a, *]
Petra Vetter [a]

[a] Visual and Cognitive Neuroscience Lab, Department of Psychology, University of Fribourg, Fribourg, Switzerland

Corresponding author at: Department of Psychology, University of Fribourg, Fribourg, Switzerland.
E-mail address: junchao.hu@unifr.ch (J. Hu)

Graphical Abstract

We suggest Fig. 4 to be included here.

Auditory evoked eye responses have been studied much less than eye responses evoked by visual input. We review the evidence on how different sounds drive different eye responses and which cortical and subcortical brain areas are involved. Overall, auditory evoked eye responses differ markedly from those visually evoked. The auditory system interacts in an intricate way with the oculomotor system, but the underlying computational and neural mechanisms are still unclear.

Abstract

Eye movements have been extensively studied with respect to visual stimulation. However, we live in a multisensory world, and how the eyes are driven by other senses has been explored much less. Here, we review the evidence on how audition can trigger and drive different eye responses and which cortical and subcortical neural correlates are involved. We provide an overview on how different types of sounds, from simple tones and noise bursts to spatially localized sounds and complex linguistic stimuli, influence saccades, microsaccades, smooth pursuit, pupil dilation and eye blinks. The reviewed evidence reveals how the auditory system interacts with the ocular-motor system, both behaviorally and neurally, and how this differs from visually driven eye responses. Some evidence points to multisensory interaction, and potential multisensory integration, but the underlying computational and neural mechanisms are still unclear. While there are marked differences in how the eyes respond to auditory compared to visual stimuli, many aspects of auditory-evoked eye responses remain underexplored, and we summarize the key open questions for future research.

Keywords

Eye movements, saccades, audition, pupil response, eye blinks, superior colliculus, frontal eye fields.

Introduction

The study of eye movements has a long tradition in vision science, and the majority of research has focused on eye responses to different features in the visual world. The characteristics of different eye responses such as saccades, microsaccades, smooth pursuit, pupil size or eye blinks, as well as their neural mechanisms, have been studied in depth, both in animals and humans, with respect to visual stimulation, e.g. with suddenly appearing or moving visual objects, attentionally or emotionally salient visual features, text for reading etc. (for reviews, see e.g., Refs 1-7). However, how the eyes respond to non-visual stimulation has been studied much less. We live in a multisensory environment and given that our brain constantly receives information from multiple senses, our eyes as well as the underlying neural circuits are very likely to be influenced by information from other senses. Here, we focus on audition, and review the empirical evidence on how different types of sounds can elicit different types of eye responses. We provide an overview on how saccades, microsaccades, smooth pursuit, pupil size and eye blinks respond to auditory stimulation in comparison to visual stimulation. We discuss how different sound features, from simple tones or noise bursts, spatially localized sounds to complex linguistic stimuli, influence different eye responses. We first review how auditory stimulation drives saccadic behavior, for example, how it modulates accuracy and latency of saccades in darkness, how it affects visual saccades, and how it can facilitate and inhibit both saccades and microsaccades. We then discuss the effects of different sound and speech stimuli on smooth pursuit, pupil responses and eye blinks. Eye responses to audiovisual stimuli are covered, too, and their potential as to how much they could indicate multisensory interaction or even integration. In addition, we review the evidence of the neural correlates, both cortical and subcortical, and as such the neural mechanisms, that could mediate the influence of audition on eye responses. Thus, we provide an overview on how the auditory system interacts with the oculomotor system to sketch potential explanations of how sounds can drive the eyes.

1. Auditory-evoked saccades

1.1 The mechanisms of sound localization

Saccades are spatially oriented eye movements, therefore we briefly explain first how the brain localizes sounds in space. Sound localization partly determines how saccades are directed towards spatial auditory targets. The brain infers sound location from spatial acoustic cues that arise from the physical interactions of sounds with the head and external ears. Acoustic cues on which the brain relies for source localization (Fig. 2**A** left) are different for the horizontal plane (sound azimuth) and the vertical plane (sound elevation). Front/back localization is similar to vertical localization, and depth/distance of sounds is derived from intensity/volume differences. Sound azimuth is determined by comparing the differences between acoustic signals arriving at each ear (binaural differences). Depending on sound spectrum, the determinant cue is either interaural time difference at low frequencies, or interaural intensity difference at high frequencies. This is known as the Duplex Theory of sound localization.[8] Sound elevation, on the other hand, is derived from spectral-shape cues, also referred to as "monaural spectral cues", because acoustic changes on the vertical plane carry no binaural difference. These cues arise from the location-dependent filtering of the sound source's spectrum (Fig. 2**A** right) by the pinna geometry of each ear.[9,10] Initially, independent neural pathways in the brainstem process the different cues of sound azimuth and elevation (for review, see Refs 11, 12). The midbrain inferior colliculus (IC) is a major convergence center for these pathways and in Section 5.1 we review its role in encoding two-dimensional (2D) sound location. Due to this early separation of sound signals in the horizontal and vertical dimension, auditory localization is initially confined to a head-centered frame of reference. A saccadic eye movement, however, is executed based on an eye-centered organization of the oculomotor system. We will discuss the transition of reference frames between auditory and oculomotor systems in more detail below.

Both vision and hearing provide spatial information, but hearing also provides information about events outside the field of view.[13] Saccade studies, however, mostly employed a fixed head position, thereby limiting the spatial range of saccadic orienting responses to the field of view. One study displayed sounds in the rear space and found that eye positions biased behavioral sound localization,[14] but otherwise, very little is known about how the eyes respond to sounds in rear space.

1.2. Paradigms used to study auditory-evoked saccades

Saccades are rapid shifts in the direction of gaze that guide the fovea to information of interest in the environment. There are many types of saccadic eye movements (Fig. 1). Exogenously driven sensorimotor saccades, often termed prosaccades, are reflexive saccades to unexpected appearances of novel stimuli. In contrast, endogenously driven volitional saccades encompass saccades directed opposite to the location of a target, known as antisaccades, and saccades directed towards the remembered location of a target, referred to as memory-guided saccades. Compared to visual saccades, research on auditory saccades has largely been limited to stimulus-driven prosaccades. Note that while the term 'reflexive' is commonly used for prosaccades, auditory saccades may rarely be truly 'reflexive' in nature. For example, if a sound is poorly localized, saccade initiation may partly rely on volitional control processes in search for the sound rather than being purely stimulus-driven and 'reflexive' (see Ref 15 for a related discussion of somatosensory saccades and a broader conceptualisation of saccades). The distinction is particularly necessary when considering saccadic latency. For example, response latency is longer for saccades relying on internal endogenous processes to specify saccadic direction and amplitude.[16] Thus, interpretations of auditory prosaccade findings should consider possible influences from internal factors.

A number of paradigms developed to test visually guided saccades have also been used for studying auditory guided saccades. These paradigms (Fig. 1) modulate the time delay between the initial fixation target and the target to which the saccade is directed to.

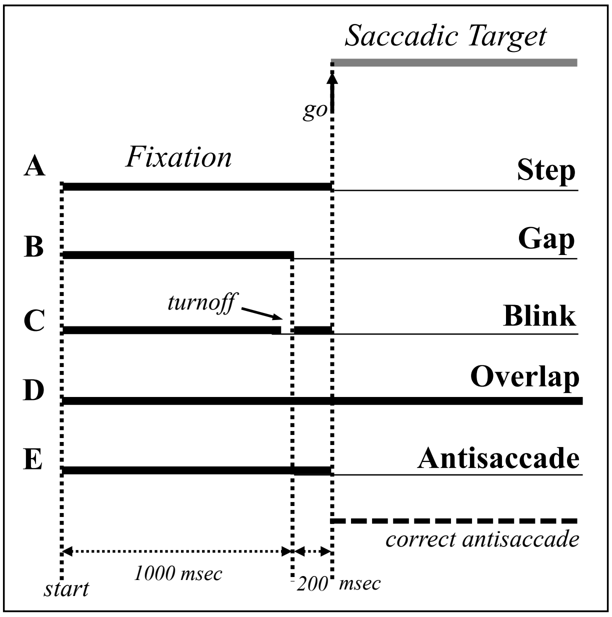

Fig. 1. Schema of stimulus paradigms used for testing saccadic behavior. Example timescales are shown at the bottom. **A**. Step or no-gap paradigm, the most common one, in which the fixation target is turned off at the same time as the saccade target is turned on. **B**. Gap paradigm, in which the fixation target is switched off (typically 200 ms) before the saccade target is switched on. **C**. Blink paradigm, in which the fixation target is turned off very briefly and then back on before the saccade target appears. **D**. Overlap paradigm, in which the fixation remains after target appearance. **E**. Antisaccade task, in which the participant is required to look in the opposite direction to where the target appears.

In the following, we will discuss the main characteristics of auditory-evoked saccades, such as their accuracy and latency, under these various paradigms in healthy participants, and how they differ from visually evoked saccades (see Table 1 for an overview of the key studies on auditory saccades).

1.3 Auditory saccade accuracy

Auditory-evoked prosaccades are slightly less accurate than visually evoked saccades when executed in darkness on the horizontal plane[17-19] and in darkness in 2D space.[20,21] This is consistent with the notion that the neural representation of acoustic space is less elaborated and less precise than the representation of visual space. For example, optimal vernier acuity of visual discrimination, i.e., the finest threshold of visually distinguishing two spatially separated lines, is much smaller (0.0008 degree)[22] than the minimum audible angle (MAA) of sound localization, i.e., the smallest angular separation at which two sounds are perceived as coming from different locations (about 1 degree).[23] The further away targets are to the eyes (that is, saccade amplitude is big), the faster auditory saccades on the horizontal plane are initiated, but the less accurate they are compared to visual saccades.[24] There is also an asymmetry with respect to cardinal direction: saccades towards auditory targets on the vertical plane are less accurate than those on the horizontal plane (Table 1 and Fig. 3A, Refs 20, 21). Another phenomenon about auditory saccade accuracy is the occurrence of multiple saccades, i.e., when a primary saccade is followed by corrective secondary saccades. While the response to visual targets is mostly a single and highly accurate saccade, the response to an auditory target usually requires multiple saccades,[20,25] with their likelihood of occurrence increasing with target eccentricity,[20] probably due to poorer sound localization accuracy. Only a few studies investigated acoustically cued antisaccades, showing reduced accuracy for auditory than for visual antisaccades.[26] Note that antisaccades are typically initiated to a blank location opposite the target, as such accuracy is poorer in both modalities due to uncertainty of the exact target location.

Despite controlled task requirements and stimulus setups, it is difficult to make direct comparisons between auditory and visual saccades. Saccade generation is determined by the combined representation of a target's bottom-up saliency and its top-down behavioral relevance to the observer.[27] These representations likely differ between modalities and few studies have directly compared auditory and visual saccades under the same stimulus conditions (Ref 28; also see Sections 1.7 and 1.8 on how stimulus relevance modulates auditory saccades).

1.4. Auditory saccade latency and eccentricity effect

Due to differences in the receptor systems, sound signals reach the cortex faster than visual signals, along with a faster conduction time to motor cortex. Hence behaviorally, simply detecting an auditory target is on average faster than detecting a visual target.[29,30] Despite faster sound detection, the brain is slower to initiate an auditory saccadic response compared to a visual saccadic response to targets close to the eyes (that is, saccades with small amplitudes), and with a lower peak velocity for a given saccade size.[17,20,21,25] This difference is explained by the processing delay between determining the spatial position of auditory cues and transforming them into oculomotor commands for accurate saccade initiation (see also Section 1.5 below). A related robust finding is the *eccentricity effect* of auditory saccades being initiated faster the further the target is in the periphery.[17,19-21,24,25,31]This eccentricity effect is influenced by the initial eye position.[17,19-21,31,32] For example, when keeping the head fixed at a straight-ahead position but varying initial fixation locations along the horizontal plane, then saccades are initiated faster the further away the target sound is from the starting fixation location.[17,20,21,31] For visually guided saccades, the eccentricity effect is the opposite, i.e., visually guided saccades are faster the closer the target is to the initial eye position. The visual eccentricity effect is also modulated by stimulus intensity, i.e., it is present only for dim light targets, but not bright ones.[33] For auditory guided saccades, no modulation by sound intensity has been observed.[24]

It is important to note that early studies varied target positions only in one dimension, the horizontal plane (i.e., target azimuth). Related reports suggested that auditory saccades are considerably slower in latency than visual saccades on the horizontal plane.[17-19,25,34] A few other studies (Table 1, Refs 20, 21, 31) looked at auditory saccades in the full two-dimensional space (i.e. towards targets with both a horizontal and a vertical component) and found that auditory saccades are faster than visual saccades for more peripheral targets in 2D space. As such, studies measuring auditory saccades only on the horizontal dimension may have underestimated saccade latency and peak velocity, and peripheral versus non-peripheral target positions may be the stronger modulating factor when comparing auditory and visual saccades. In terms of saccade accuracy, auditory saccades towards broad-band noise are accurate in both dimensions, although saccade accuracy in elevation is slightly lower than on the azimuth. For single frequency tones (Fig. 2D), saccades are driven mainly by the azimuth component of tones, whereas elevation localisation of tones is rather poor (see also Section 1.7 and Fig. 3A).

Table 1 Summary of auditory-evoked saccade studies in humans.

Saccade Study	Paradigm	Fixation	Auditory Stimulus	Auditory Target Range
Zahn et al., 1978 (Ref 17)	Overlap	LED center	Narrow band noise bursts (0.75, 1.5, 3.0, 6.0 kHz) Duration 400-600 ms	Horizontal: ± 20 deg
Zahn et al., 1979 (Ref 18)	Overlap	LED, 15 & 5 deg left, center, 15 & 5 deg right	Narrow band noise burst (1.5 kHz) Duration 1000 ms	Horizontal: ± 20 deg
Zambarbieri et al., 1982 (Ref 25)	Step	LED center	Square wave (7 Hz)	Horizontal: ± 30 deg
Jay and Sparks, 1990 (Ref 20)	Step	LED, 24 deg left, center, 24 deg right	Broad band noise burst (0.02 - 20 kHz)	Horizontal: ± 24 deg Vertical: ± 20 deg
Lueck et al., 1990 (Ref 34)	Step	LED, 15 deg left, center, 15 deg right	Sinewave (3.4 kHz) Duration 200 ms	Horizontal: ± 15 deg
Frens and Van Opstal, 1994, 1995 (Refs 21, 31)	Step	LED center	Broad band noise burst (0.15 – 20 kHz), tones (0.5, 1.0, 2.0, 5.0, 10 kHz) Default duration 500 ms	Horizontal: ± 40 deg Vertical: ± 40 deg
Yao and Peck, 1997 (Ref 19)	Step	LED, 10 deg left, center, 10 deg right	Broad band noise burst (0.2 - 20 kHz)	Horizontal: ± 30 deg
Shafiq et al., 1998 (Ref 44)	Step, Gap, Blink, Overlap	500-Hz tone	Broad band noise Duration 1200 ms	Horizontal: ± 15 deg
Taylor et al., 1999 (Ref 45)	Step, Gap, Overlap	LED center Broad band noise center	Tone Duration max. 1200 ms	Horizontal: ± 10 deg
Gabriel et al., 2010 (Ref 24)	Step	LED center	Broad band noise burst (0.1 - 1.2 kHz), narrow band noises (0.25, 0.7, 1.5, 5.0, 9.0 kHz) Duration 500 ms	Horizontal: ± 36 deg

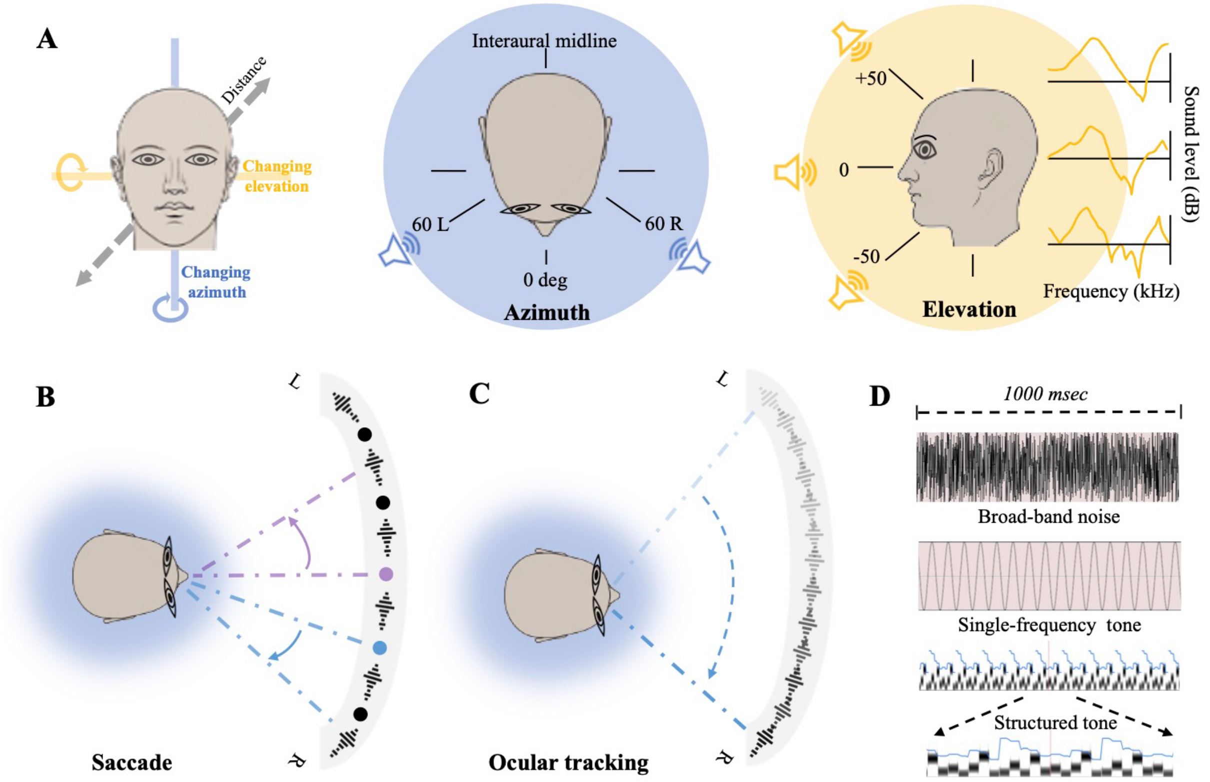

Fig. 2. Human auditory space and ocular responses to sounds in darkness. **A.** Auditory localization in space is organized in three dimensions: azimuth (horizontal, blue), elevation (vertical, orange), and distance (depth, gray). Human binocular range is around 120 degrees on the horizontal plane (Azimuth, L, left; R, right) and around 100 degrees on the vertical plane (Elevation). With fixed head position, saccades to sounds are restricted in this range. Sound azimuth causes interaural time differences, and elevated sound provides spectral cues. Right panel shows examples of head-related transfer functions. The same sound is presented to the observer at three elevations: at level to the ears, above, and below the ears. Note how the head/pinnae geometry shapes the perceived sound spectrum, that is, it affects the intensities of component frequencies of the sound, as detected by the eardrum. **B.** Auditory saccades from different starting fixations. Dots are fixation points, spaced by auditory targets (illustrated as sound waves). Two examples are shown: a leftward saccade from a central fixation point (purple) and a rightward saccade from an eccentric fixation point (blue). **C.** Ocular tracking of moving sounds. A sound source (in darkness) moves from left to the right. Auditory target pursuit is contaminated by saccades, here indicated by dotted arrows, as opposed to smooth pursuit of visual targets. **D.** Examples of sounds used to elicit ocular response. Top rows are broad-band noise

and a single-frequency tone, both typical in saccade studies (Table 1). Bottom rows show pitch and spectrum of a structured, regular tone from Zhao et al.[90] Figure produced with Praat (Version 6.2.14).[164]

1.5. What underlies the auditory eccentricity effect?

What leads to the eccentricity effect of auditory saccades and the modulation thereof? First, it was argued that sound localization ability decreases for central targets because when the source of a sound is placed near the midline of the subject's head, both phase and amplitude differences at the level of the subject's ears are very small. Hence it is more difficult to estimate the absolute target location[25] and it takes more time to determine where to look when the sound is straight ahead. This is intuitive for when direction of the head and eyes are aligned and probably could explain the eccentricity effect in auditory saccades. Yet, it was shown that saccade latency depends on the relative position of the sound with respect to the eyes rather than the head. The pattern here is that the closer a sound is to the starting fixation point when head and eyes are not aligned, the longer it takes to direct the saccade (see Fig. 3**C** for illustration). Note the opposite is found for visual saccades which are faster the closer the target is to the initial fixation point.[17,18,20,25] As mentioned earlier, compared to visual targets, sounds are detected much faster, but this advantage is lost during auditory saccades, especially for sounds close to the eyes. This longer latency is thought to reflect a higher degree of uncertainty of determining the auditory target location, resulting in additional time being required to generate a saccade towards sounds.

To untangle this phenomenon, the neural processing underlying auditory saccade generation and execution (Fig. 4) should be considered. Unlike visual signals, which are predominantly represented in the brain with an eye-centered, retinotopic reference frame, auditory signals (although originally head-centered) arrive in the brain with hybrid coordinates, intermediate between head- and eye-centered coordinates. This has been demonstrated in primates for neurons in the core auditory cortices,[35] parietal cortex,[36] superior colliculus (**SC**),[37-39] and inferior colliculus (**IC**).[40] It shows that auditory neuronal representations are substantially modulated by concurrent eye position relative to the head/ears and this modulation of reference frame occurs as early as in the IC along the auditory pathway (Section 5.1). It has also been suggested that the auditory reference frame in the SC changes between sound detection and saccade execution,[39] such that in the sensory period, the auditory reference frame is encoded as hybrid, and in the motor period, auditory saccades have an eye-centered encoding. Moreover, with the change in the angle of gaze, the auditory receptive field of SC neurons shifts correspondingly so that it maintains a constant spatial relationship to the eyes, hence minimizing the misalignment to the visual reference frame.[39,41] Such shift likely allows for accurate saccades to auditory targets regardless of initial eye positions.[18,42] Overall, signals passing the acoustic-auditory system are heavily influenced by the position of the eyes and there is a mixed interplay of reference frames used by the brain to generate accurate oculomotor commands. The fact that physical change in the visual axis can reorient a neuron's auditory receptive field (so that it now samples information from a different area of space) may also explain the dependence of the auditory eccentricity effect on eye position.

These effects of eye position on auditory saccades have been studied only on the horizontal plane, but whether the same effects can be found for the vertical plane is still unclear.

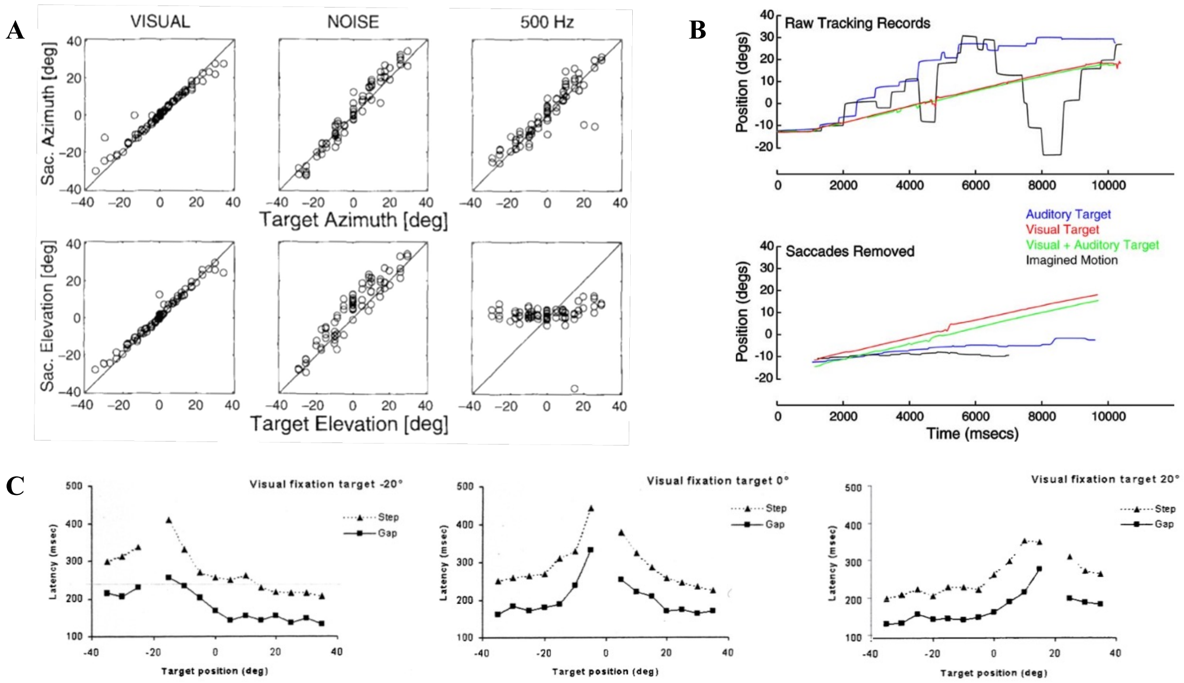

Fig. 3. Example measurements of auditory saccades in response to sound. **A.** End-point azimuth (top) and elevation (bottom) of primary saccade are shown as a function of target azimuth (horizontal plane) and elevation (vertical plane). Left column shows visual saccades, center column saccades to 500 ms broad-band noise, right column saccades to a 500-Hz tone. Saccades to broad-band noise are less accurate on the vertical plane and the eyes fail to respond to the elevation component of a single frequency tone. Figure from Frens and Van Opstal.[21] Springer Nature. **B.** Ocular tracking of different moving targets. Top row shows eye position over a 10 s span. Ocular tracking of sound (blue) is not continuous as the typical smooth pursuit of visual targets (red). Bottom row shows relative changes of eye position when saccades are removed. Figure from Boucher et al.[75] Elsevier. **C.** Mean latency of auditory saccades as a function of target sound position in the step and gap paradigms, with three different starting fixations. Note that the eccentricity effect depends on initial eye position. Figure from Zambarbieri[32] Elsevier.

1.6. Gap effect and express saccades

Another classic setting to test saccades is the gap paradigm (Fig. 1). The associated gap effect occurs when a visual fixation point is extinguished shortly (typically 200 ms) before the appearance of a peripheral visual target. Then a subsequent visual saccade to that target is initiated faster compared to when there is no temporal gap between fixation point offset and target onset (step paradigm) and when the fixation stays on after target onset (overlap paradigm).[43] With auditory stimuli, there are mixed findings for the reduction of saccadic latency (when comparing step and gap paradigms). One study found a smaller gap effect for auditory than for visual targets, that is, the gap effect is less strong for auditory targets,[44] whereas Taylor et al.[45] found a larger gap effect in the auditory modality. A second phenomenon under visual gap conditions, although less reliable, is the occurrence of express saccades, which are reflex-like saccades of extremely short latency (80-120 ms),[46] usually followed by saccades of normal latency approximately 50 ms later. Express saccades for auditory targets have not been observed so far.

For visual saccades, it has often been suggested that the gap effect and express saccades are due to disengagement of covert visual spatial attention following fixation offset.[47] Engaged attention, for instance during fixation, acts as an inhibitor of the saccade system, while disengaged attention (e.g., when the fixation point disappears) permits saccade preparatory processes which facilitate saccade responses. That is, when there is a transient gap during fixation, disengaged attention gets the saccadic system ready for response, which reduces the time needed for saccade execution and manifests as the gap effect and sometimes as express saccade. However, using visual fixation and visual saccade targets, Klein et al.[48] showed that attention to the visual fixation point did not play a role in the visual gap effect, and that the governing factor was the release of ocular fixation when the fixation point was removed. To disengage attentional effects from fixation release in auditory saccades, Shafiq et al.[44] included an auditory “blink” task[47] under auditory gap conditions in which the fixation tone was turned off 20 ms before target presentation (blink paradigm, Fig. 1). The brief turnoff created an easily discernible click sound that served as a warning signal to disengage attention from current auditory fixation before subsequently occurring auditory targets. In line with Klein et al.,[48] Shafiq et al.[44] found that the acoustic warning signal did not decrease saccade latency compared to the step, no-warning condition. Therefore, the gap effect of auditory saccades, as the gap effect of visual saccades, cannot be explained by attentional modulation, and may rather be a result of ocular release of fixation. For example, there is direct evidence for a visual fixation zone in primate SC that neurons in its rostral pole (rSC) discharge during active visual fixation and pause during saccadic execution.[49,50] Chemical inhibition of these fixation cells leads to a reduced ability to fixate and increases unrestrained visual express saccades (< 100 ms). Early fixation release, as in the gap paradigm, may remove the gating process of the fixation system, thereby facilitating saccade initiation (i.e., leading to shorter saccadic latency). These visual fixation neurons may not be purely unimodal, as there is clear reduction of saccadic latency in the auditory gap paradigm after fixation tone release.[44] However, direct evidence of these mechanisms for auditory saccades is still lacking.

Overall, saccades to auditory targets are more variable and less accurate than saccades to visual targets at the same eccentricities. However, all the so far mentioned studies on auditory saccades were done in darkness. Thus, humans are able to direct their gaze towards sound targets in complete darkness with correct, if not precise, eye movements. This implies that the positional information derived by the auditory system is transformed into valid oculomotor commands. From this point of view, differences in saccadic latency and endpoint variability are likely due to the time and error costs during coordinates transformation between the auditory and the visuomotor system.

1.7. Sound features modulating auditory saccades

Typically, the sounds used in the above mentioned studies included broad-band noise burst,[19-21] narrow-band noise burst,[17,18,24] low frequency pulses,[25] single frequency tones[21] and sine wave tones.[34] Notably, the sounds varied not only in their frequency, but also in duration (Table 1). These acoustic features were later found to be important determinants for the accompanying saccades, particularly on the vertical plane. For example, Frens and Van Opstal[21] used a set of tone stimuli with various single frequencies and showed that, compared to broad-band noise, participants' saccades failed to respond to the elevation component of tones, i.e., saccade elevation was independent of the tones' vertical location (Fig. 3**A**). Thus, sufficient spectral cues (present in broad-band noise but absent in single frequency tones) are necessary for accurate sound elevation localization and for accurate vertical saccade initiation. A similar effect was found for the duration of broad-band noise bursts, that is, noise-evoked saccade elevation accuracy decreases as noise duration decreases as the brain is unable to extract spectral cues from short presentations of noise.[21] On the horizontal plane, Gabriel et al.[24] found that saccades were initiated more quickly to narrow-band noise bursts centered at midrange frequencies than at very low and very high frequencies. Interestingly, the observed mid-range frequency preference is in line with an acoustic fovea,[51] with the highest auditory acuity occurring for frequencies falling within the range of human speech. Thus, the human auditory saccade system seems to be more sensitive to speech-like sound frequencies. All in all, this evidence shows that sound by itself is not sufficient to evoke an accurate saccade in the horizontal or vertical dimension, the ease with which sound location is extracted from different sound features also modulates auditory saccades.

A lot of research on auditory saccades used simple auditory stimuli such as noise and sine wave tones, with limited variations in saliency except for sound intensity and frequency. These sounds are often very short in duration and contain no meaning due to the nature of the tasks and the fact that e.g., auditory saccades are mostly driven by abrupt, short sounds. Nevertheless, these studies are important as they reveal how oculomotor dynamics are driven by sounds at the very fundamental level. In the natural environment, however, oculomotor behavior usually operates in an acoustically rich context. Recently, more studies have used more complex sounds with linguistic content, and, for example, language comprehension has been shown to modulate the general orienting response (as well as modulating pupil dilation and eye blinks, see Sections 3 and 4 below). For example, Huette et al.[52] manipulated the grammatical aspect of speech and compared eye movements to sentences in simple past tense (e.g., “walked”), which emphasizes the end-state of action, versus past progressive tense (e.g., “was walking”) that emphasizes the ongoing nature of actions. Fixations were shorter and saccades more dispersed for the progressive condition, potentially reflecting the representation of more dynamic events. This shows that high-level cognition such as linguistic and grammatical interpretation can guide eye orienting behavior.

1.8. Inhibition and facilitation of visual saccades by sounds

In natural behavior, gaze orienting to objects of interest is often affected by other competing stimuli in close spatial or temporal distance to the target that are either relevant or irrelevant (distractor), in the same (unimodal) or different modality (crossmodal) to the target. Presentation of sounds affects visual saccadic responses in different ways. For example, during exploration of a visual scene, irrelevant tones prolong fixations and inhibit visual saccades (*distractor effect*).[53,54] Moreover, saccades and blinks are more strongly inhibited prior to predictable and probabilistic sounds compared to unpredictable sounds,[55] indicating that eye movement freezing could be a biomarker of auditory temporal expectations. On the other hand, facilitation of saccades due to sounds is also observed when a visual target and an acoustic non-target (i.e., distractor) are presented synchronously.[56,57] In this case, facilitation occurs such that the saccadic reaction to multimodal targets is faster than any of the responses to unimodal stimuli. This speed-up in reaction time is often referred to as *redundant target effect*. For example, Diederich and Colonius[57] found that saccadic reaction time to a visual target was reduced in the presence of an auditory non-target, and the effects were greatly enlarged when two non-targets (auditory & tactile) were simultaneously presented.

Originally, the redundant/facilitated saccadic response with multimodal stimulation has been explained with a race model such that the measured reaction time is the winning racer's time of parallel sensory processes (racers) triggered by multimodal signals (e.g., visual & auditory). In particular, the idea postulates context independence, that is, individual racers proceed at the same speed in the unimodal and multimodal conditions, and the winner's advantage is purely statistical.[58] However, Raab's *context-independent race model* cannot explain that behavioral responses to a multimodal target stimulus are often found to be faster than the responses to a unimodal stimulus.[59] An alternative account, therefore, is that the redundant target effect of saccades is likely due to coactivation, that is, the redundant response is influenced by both racers.[57] Some coactivation models also assume "integrated processing of the combined stimuli",[60] i.e., sensory convergence. Taken together, irrelevant sounds can both inhibit and facilitate saccades to visual targets, and in the case of facilitation, the underlying multisensory interaction mechanisms are still debated.

1.9. Inhibition of microsaccades by sounds

Microsaccades, or fixational saccades, are rapid, miniature saccades occurring involuntarily during attempted fixation, with a frequency of 1-2 per second.[61] There was a debate whether microsaccades had any specific functional role for visual perception, until Engbert and Kliegl[62] demonstrated that microsaccades are optimized motor acts with rich spatio-temporal dynamics to enhance visual perception during fixation. Similar to saccadic inhibition induced by sounds, microsaccades are strongly inhibited after sound presentation,[63,64] with their direction being biased towards the location of auditory stimuli (*orientation effect*).[63] Further, microsaccadic inhibition (MSI) is typically followed by a rebound of enhanced microsaccade rates.[63-65] Furthermore, attended deviant tones in an auditory sequence induce prolonged MSI,[66-68] regardless of stimulus-specific properties like pitch.[69] Moreover, when passively attended deviant tones are frequent and predictable, MSI occurred faster as a function of the pitch difference between the deviants and the standard tones.[69] Using more natural sounds (e.g., car alarm, sneeze, camera shutter, kiss, etc.), Zhao et al.[70] found that MSI was systematically modulated by auditory salience (as rated subjectively) such that more salient sounds evoked earlier and larger MSI. Therefore, MSI may index an orienting mechanism that is sensitive to not just the onset, but also the perceived salience of sounds.[69,70] Compared to visual- or auditory-alone conditions, stronger MSI was observed in audiovisual conditions when participants made saccades to audiovisual targets, possibly reflecting multisensory integration affecting ocular control.[71] These multimodal effects on MSI point to a potentially important role of the superior colliculus in both multisensory integration[41] and microsaccade generation (Ref 64; for review, see Ref 3; see also Section 5 on neural correlates).

2. Smooth pursuit

Aside from saccades, another important ocular orienting response is smooth pursuit eye movements (SPEMs). Smooth pursuit eye movements cannot be performed at will but require the percept of a moving *visual* stimulus.[72] Smooth pursuit allows the observer to follow moving visual objects and maintain retinal stability of images on the fovea. Compared to the rapid, abrupt nature of saccadic response, smooth pursuit of a moving target is slow and continuous. Whereas the saccadic system operates on position signals, sustained smooth pursuit eye movements require velocity signals from the target to track the relative retinal motion.

Unlike the release of saccade, which is under voluntary control, smooth pursuit depends heavily on a visible moving target. Motion detectors in the visual system provide the inputs to engage the smooth pursuit control system. However, aside from visual motion, kinesthetic and cutaneous motion also provides signals sufficient to trigger smooth pursuit,[73] although in these cases, smooth pursuit is contaminated with saccades.[74] Here, the motion information is likely provided by proprioceptive and somatosensory afferents. Moving auditory targets, to the contrary, are unable to generate smooth pursuit responses. Boucher et al.[75] and Berryhill et al.[76] compared smooth pursuit responses for moving auditory and visual stimuli, and found that sounds cannot trigger smooth pursuit, and that the tracking of continuous auditory motion is indistinguishable from when participants perform pursuit to imagined motion in darkness, both in the horizontal, vertical and depth planes. While ocular tracking of visual motion is smooth and continuous, ocular tracking of auditory motion is primarily carried out by a sequence of saccades (Fig. 3**B**).[74]

Although sounds cannot trigger smooth pursuit, they still affect pursuit of visual targets. For example, sounds can help maintain high pursuit velocities when a moving visual target disappears. Madelain and Krauzlis[77] used a static tone as a positive reinforcer contingent on smooth eye movements. They found that feedback from sounds helped participants to maintain high pursuit velocities even after the disappearance of the visual target. This indicates that the motor output of the pursuit system can be influenced by learned contingencies from auditory signals.

In general, smooth pursuit cannot reach large velocities in a short period of time and cannot track very fast targets (see Ref 2, for a review). For visual targets with a predictable motion trajectory, the oculomotor system uses prediction to anticipate the future target path for smooth pursuit. However, this mechanism cannot manage unpredictable or very fast-moving targets. Consequently, pursuit and saccades are combined for tracking of unpredictable targets to avoid large position error and the eye lagging behind the target. In this situation, a saccade that helps realigning the eye when smooth pursuit falls behind is called a catch-up saccade. In contrast with saccades to stationary targets accounting only for positional error, catch-up saccades necessitate the retinal velocity error (motion of the eye relative to the target) to make a predictive, and accurate “jump”.[78]

Based on the auditory distractor effect in visual saccades (see Section 1.8), Kerzel et al.[79] investigated how catch-up saccades were affected by abrupt acoustic onsets. They showed that a loud irrelevant click sound was able to inhibit catch-up saccades but only weakly inhibited pursuit, indicating a privileged access of acoustic input to the saccadic system. Thus, inhibition of saccadic eye movements may dissociate from inhibition of smooth pursuit. While the effect of a click (i.e., auditory distractor) on smooth pursuit was much smaller than the effect of a flash (i.e., visual distractor), catch-up saccades were inhibited to the same degree by clicks and flashes. The inhibition of catch-up saccades is consistent with the acoustic distractor effect on saccades during visual exploration.[53,54] And the (relative) resistance of smooth pursuit gain to acoustic distraction shows that visual signals are the main input to the pursuit system.

3. Pupillometry

So far, we have covered two important aspects of gaze orienting behavior, namely saccades and smooth pursuit. However, gaze orienting is accompanied by a coordinated interplay of multiple eye responses, including non-spatial pupil responses that change the pupil size to admit more or less light into the eyes. Pupil response is in part reflexive to visual input, for example, it responds to changes in brightness and focal distance, allowing sufficient light to arrive at the retina. In addition, pupil changes are often driven by cognitive processes and considered as an integral part of orienting[80] and a key indicator for arousal.[5,81] These cognitively driven pupillary modulations are also seen in pupil responses to sounds, but to what extent pupil responses are driven by arousal, cognitive processes or both, is still unclear.

In monkeys and humans, acoustic stimuli evoke transient pupillary dilation, and dilation responses are modulated by sound intensity,[82,83] and subjective salience of the sound.[83] The pupil also dilates more to deviant than standard tones.[69] Saccades and pupil responses often act together and correlate during orienting behavior. For instance, faster saccadic responses to targets are accompanied by larger pupil dilations, and this effect is present across sensory modality of the saccade target (visual, auditory, audiovisual).[71] Note while saccades are often fast (< 250 ms), pupil responses are slow signals peaking around 1 s after stimulus onset. The relationship between saccadic latency and magnitude of pupil dilation may be mediated by an interplay of arousal, attention and cognitive factors. Faster saccadic responses are associated with a state of increased arousal and attention which in turn are reflected in larger pupil dilations. Compared to visual stimuli, auditory stimuli elicit faster transient pupil dilations, although not as fast as dilations to bimodal audiovisual stimuli.[71] Whether pupillary responses can be a sign of multisensory integration is much under debate. In humans, one study[84] suggests pupil size changes to multisensory signals to be super-additive and thus a sign of multisensory integration whereas another[85] concludes pupil size changes to be simply additive, i.e., a linear summation of the unisensory pupil size changes. As most studies differ in stimulus intensity (e.g., supra-threshold or near-threshold), stimulus intensity contrast (e.g., high visual & low auditory signal, or high visual & high auditory signal), and sound type (e.g., noise or tones), it remains unclear to what extent pupil size changes indicate multisensory integration.

In naturalistic sounds, one often finds regularities such as rhythms in music and speech. Our brain is highly adapted to learn and predict such frequently reoccurring signals such that expected signals require comparatively less computational demands.[86,87] Recent studies explored how pupil diameter may reflect modulation of sensory and cognitive processing by structured versus random auditory streams. For example, Milne et al.[88] asked participants to listen to rapid tone-pips[89] containing either regular or randomly changing patterns, while the base rate of tones was kept the same. They observed a smaller sustained (tonic) pupil diameter for regular than for random auditory sequences. Another study[90] found stronger transient (phasic) pupil dilations when changes in the auditory sequence pattern violate structural regularity (Fig. 2**D**) compared to when a predictable structure emerged. Interestingly, when the auditory pattern changes were rendered behaviorally relevant through active monitoring tasks, marked differences in the dynamics of tonic pupil activity emerged. At first, when the task was entirely passive, there were no differences between sustained pupil responses to random and regular parts of the sound sequence. However, when participants monitored and detected transitions between sound patterns, there were diverging tonic pupil dilations to the random and regular sound structures themselves (as in Ref 88), even in the absence of a transition. Together, these studies showed that inherent regularity and statistics embedded in sound sequences evoke distinct pupil responses, and that these responses are task modulated.

Pupillometry has also been increasingly used as a measure for the brain's evaluation of meaningful sounds or linguistic stimuli. For example, changes in pupil size provide an objective tool to assess listening effort under various acoustic challenges during speech comprehension. When listening to speech, a larger pupil is observed for syntactically more complex sentences compared to simple sentences,[91,92] and acoustically more degraded versus less degraded speech.[93-95] Further, the semantic demands of speech also increase pupil dilation. When listening to everyday naturalistic sentences containing homophones (ambiguous words with more than one meaning), pupil size is larger and peaks later compared to matched sentences without homophones.[95] While there is microsaccadic inhibition (MSI) following sentence onset, neither sentence clarity nor semantic ambiguity affects microsaccade rate, suggesting that changes in pupil dilation are probably not related to microsaccades. This evidence demonstrates that pupil response is modulated also by the cognitive demands of speech comprehension.

4. Blinks

Spontaneous eye blinks are far more frequent than necessary for simply maintaining the tear film of the eye.[96,97] Besides environmental conditions, endogenous factors such as attention strongly affect the temporal patterns of eye blinks. When exclusively acoustic stimuli are presented, spontaneous blinking is momentarily suppressed,[98-100] and the suppression is stronger as listening difficulty increases.[100] Moreover, cognitive task requirements in either the visual or auditory modality (e.g., counting the number of visual or auditory targets) result in an increased post-stimulus blink frequency compared to passive viewing or listening conditions, respectively.[98] This pattern closely resembles the stereotypical inhibition-release-baseline response of microsaccades[64], suggesting that blinks and microsaccades might rely on a common mechanism for oculomotor inhibition that acts independently of stimulus modality.[101] Recently, Huber et al.[102] found that blink patterns were modulated by the predictability of presented tones and the requirement for a motor response. Specifically, when a manual reaction is required (compared to when no action is required), there was increased blink suppression for tones with high predictability, followed by an excess of blinking after tone offset. The rebound of eye blinks was not merely a secondary reaction to manual response because it was observed prior to keypress, too. Another study[99] asked participants to respond or hold response (go/no-go task) according to the information in spoken sentences. They found that blink suppression was released significantly earlier in no-go trials (information absent, response not needed) than in go trials. This suggests that eye blinks are not merely triggered by task sound offset, instead, they are meaningfully paced in time for the evaluation of linguistic input. Together, these studies showed that spontaneous blinking is modulated by anticipation and the cognitive evaluation of sound input.

Other studies looked at how eye blinks track spoken sentences. Jin et al.[103] demonstrated that ocular muscle activity controlling eye blinks (as measured with electro-oculography, EOG) synchronizes to high-level linguistic structures, in both eyes open and eyes closed conditions. In particular, when participants listen to a sequence of simple four-word sentences with open eyes, blinks and EOG responses are synchronized to the temporal frequency of words and sentences, both with and without visual stimulus. The same pattern was observed in the EOG when the eyes were closed. When the word order was presented randomly, blinks and EOG responses were uncoordinated. This demonstrates that ocular muscle activity can track the linguistic structure of natural speech. Additionally, the temporal dynamics of vertical EOG power displayed an inhibition-release-baseline pattern: EOG power decreased after the speech sequence onset, was suppressed during speech perception and sharply increased after its offset. Such dynamic change of EOG occurred even with closed eyes. As the vertical EOG likely reflects eyelid movements, this finding mirrors the blinking and MSI patterns in earlier studies.[98,101]

5. Neural correlates

In previous sections, we have reviewed how different sounds guide ocular responses. Here we briefly describe the neural mechanisms (Fig. 4) proposed to control the generation of eye movements to sounds. Typically, in response to visual targets, the main cortical ocular-motor brain areas include the frontal eye field (FEF) and the supplementary eye field (SEF) in the frontal lobe; several associative, attentional, and motor areas in the posterior parietal cortex (PPC), such as the parietal eye field (PEF); and the cingulate eye field in the anterior cingulate cortex (for a comprehensive review, see Ref 104). The most important subcortical structure involved in saccade generation is the superior colliculus (SC). Modulation of eye movements is likely to be implemented by the dorsolateral prefrontal cortex (DLPFC). The middle occipito-temporal visual area (hMT+/V5) is essential for visual motion detection, and also involved in auditory motion.[105-107]

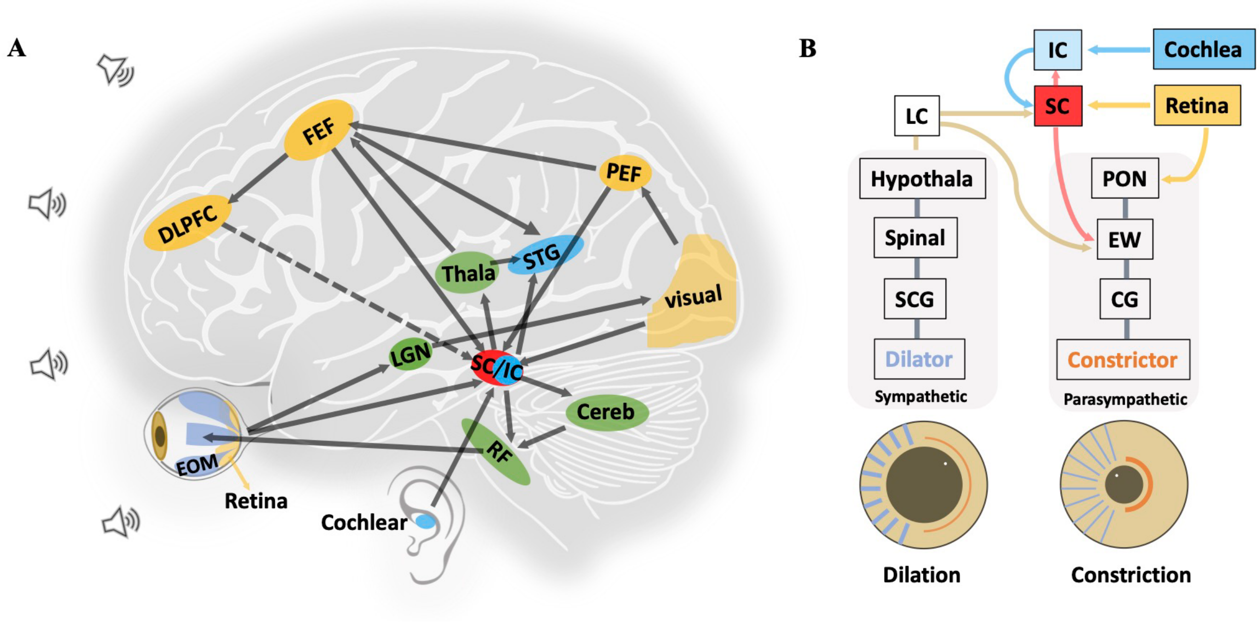

Fig. 4. Brain areas controlling saccades and pupil response. **A.** Simplified neural circuitry underlying saccade generation. DLPFC, dorsal lateral prefrontal cortex; FEF, frontal eye fields; PEF, parietal eye fields (homologue to the lateral intraparietal area, LIP, in monkey); STG, superior temporal gyrus; visual, visual cortex; Thal, Thalamus; LGN, lateral geniculate nucleus; SC, superior colliculus; IC, inferior colliculus; RF reticular formation in the brainstem; Cereb, cerebellum; EOM, extraocular muscles. Figure modified, from Martinez-Conde et al.[3] Springer Nature. **B**. Simplified neural pathways of auditorily or visually induced pupil response. Sympathetic and parasympathetic innervations lead to pupillary dilation and constriction, respectively. Dilation is controlled by the dilator pupillae (blue) in the periphery of the iris, and constriction by the circular sphincter pupillae (orange). IC, inferior colliculus; LC, locus coeruleus; Hypothala, hypothalamus; SGC, superior cervical ganglion; PON, pretectal olivary nucleus; EW, Edinger-Westphal nucleus of third nerve; CG, cilliary ganglion. Figure modified from Wang and Munoz[147] Elsevier, and Mathôt[5] Ubiquity Press.

5.1. Neural correlates of auditory saccades

For accurate shifts of gaze, the superior colliculus is a vital brain structure because it receives convergent inputs from almost all the cortical and subcortical structures involved in the generation of saccades (Fig. 4**A**).[108] Evidence from animal electrophysiology shows that neurons in the deep layers of the SC respond to inputs from a range of sensory modalities, and their visual, auditory, and somatosensory receptive fields are in close topographical alignment. Further, motor neurons in the intermediate and deeper layers of the SC register the initiation of saccades of particular amplitudes and to particular spatial locations.[109,110] For visual spatial signals, the SC sensory and motor maps are closely aligned, allowing a highly accurate visuomotor coordination for orienting,[111] as seen behaviorally in the fast and accurate visual saccadic response. Auditory space representation, however, is larger and less determined than visual space representation (see Section 1.3). Compared to the visual modality, the auditory space map in the deeper layers of the SC emerges developmentally later and over a much longer period of postnatal life.[112] Interestingly, during overt orienting behavior, when the gaze angle changes, the auditory receptive field of SC neurons shifts correspondingly such that the alignment of the visual and auditory maps can be maintained.[41] Thus, visual input may play an instructive role in the development[112] and maintenance[41] of appropriately tuned auditory-evoked eye responses. And it could to some extent explain the behavioral characteristics of auditory saccades in darkness, such that their low accuracy may be due to larger receptive fields in the SC (compared to the much finer visual receptive fields), and that the saccade eccentricity effect is modulated by initial eye fixation position.[20]

A remarkable property of the SC is that its multisensory neurons demonstrate unusual integrative properties (e.g., super-additivity) to multisensory stimuli, probably contributing to the distinct saccadic response facilitation in the redundant target effect (see Section 1.8, Refs 57, 60). According to Stein and Meredith[41], stimulus intensity and the relative spatial and temporal presentation of audiovisual signals influence the timing and size of collicular neural responses (as derived from cat SC electrophysiological data). However, the link between collicular and behavioral/ocular responses to multimodal signals is not straightforward, because multisensory neurons vary in their response properties, and the response of a single multisensory neuron can vary from sub-additive to additive and super-additive depending on the stimulus strength.[113] Thus, inferences from animal SC evidence to human behavioral studies should be drawn with caution.

Together with the frontal eye fields, the SC projects directly to the brainstem reticular formation to provide the necessary input to the saccadic premotor circuit for saccade initiation. The SC is necessary because a collicular lesion leads to a lack of orienting reflex and a lack of responses to novel visual and auditory stimuli.[114] On the other hand, deactivation of cortical visual or auditory afferents to SC neurons does not affect collicular responses to unimodal stimuli, although it seems to render SC unable to integrate visual and auditory signals.[115]

Inferior colliculus (IC) is an important subcortical node early in the auditory pathway for the encoding of 2D sound locations (for review, see Ref 116). Not only is it necessary for relaying ascending and descending auditory information (for review, see Ref 117), IC also receives inputs from non-auditory sources such as vision and somatosensation (for review, see Ref 118). Importantly, IC has reciprocal connectivity to SC, the main site of multisensory saccade generation.[119] IC projects to SC as shown in cats[120,121] and ferrets.[122] Direct neurophysiological evidence from behaving primates demonstrated IC being sensitive to absolute sound level, to sound-source azimuth and elevation.[123] Given the collicular connections, it is likely that SC incorporates sound-source inputs from IC, at least partly, for precise oculomotor control such as auditory saccades. IC responses to sounds also seem to vary with eye positions non-linearly,[40] suggesting a hybrid reference frame (see Section 1.5) for sound localisation at the level of IC.

For auditorily guided saccades, another important cortical ocular-motor area are the frontal eye fields (FEF). FEF motor neurons are active preceding saccades to visual targets as well as to auditory targets made in the dark.[124] The crucial factor for a FEF discharge before auditorily guided saccades is the location of an auditory target relative to the current direction of gaze.[124] Therefore, sound location cues, even though mainly head-centered, appear to already conform to the visual coordinate system at the FEF stage (see Section 1.5, for the discussion of reference frame at the level of the SC). The FEF projects directly to the pons at the brainstem[125,126] for saccade execution. It also sends afferents to the SC[127] such that the medial subdivision of FEF targets those collicular neurons exhibiting visual and auditory sensory responses.[128] Unlike a collicular lesion, a FEF lesion does not abolish saccades, it very mildly affects most reflexive visual saccades but might delay eye movements that involve a voluntary component (e.g., fixation disengagement, for a review see Table 4 of Ref 129). It is thus hypothesized that FEF plays a context-dependent modulatory role over the cortical and subcortical structures involved in different reflexive visual saccades. It is less clear how the FEF modulates auditory saccades, as studies thus far examined mainly the visual modality, but FEF of normally sighted individuals is responsive to auditory stimuli in covert shifts of spatial attention (without eye movements).[130] In primates, the rostral part of FEF has extensive projections to the superior temporal gyrus (STG)[131] which is part of auditory cortex. Combining intracranial recording and eye tracking, a recent study[132] in humans found that FEF shares bidirectional functional connectivity with the STG during the saccade-fixation cycle. Specifically, during naturalistic viewing of movie clips with and without viewing, bottom-up influences from the STG to FEF is strongest after fixation onset whereas top-down interactions from FEF to STG peak in later phase of fixation and during ensuing saccades. It is possible that auditory cortex is involved in an attentional network that enhances oculomotor responses to visual inputs,[132] or that auditory stimulation drive FEF and eye movements directly. It remains to be investigated how the interactions between FEF and STG are modulated purely by sounds without visual stimulation.

Eye movements, especially reflexive orientation eye movements such as saccades, have to be inhibited at moments when they are unwanted. Dorsolateral prefrontal cortex (DLPFC) is thought to be specifically involved in saccadic inhibition. The DLPFC has direct connections with the main cortical ocular motor areas (for review, see Ref

1). Under the visual antisaccade paradigm in humans, a DLPFC lesion, but not a FEF lesion, leads to increased misdirected saccades, whereas the latency of correct antisaccades increases after FEF lesion, but not after DLPFC lesion.[133] This suggests that inhibition of unwanted reflexive saccades is under the specific control of the DLPFC, as later confirmed in both primate[134] and human fMRI studies.[135] And this control is probably exerted directly through connections from the DLPFC to the SC,[136,137] but not via the FEF. Again, like for FEF, studies on the modulation of DLPFC looked at visual saccades alone, and it is unclear to what extent DLPFC inhibits auditorily guided saccades. Behavioral findings suggest that DLPFC may be involved in the oculomotor inhibition mechanism of auditory temporal expectations, for instance the saccadic inhibition prior to sounds with predictable structure.[55] Another source of inhibition on saccade neurons could be the substantia nigra pars reticulata (SNpr)[138] by exerting tonic inhibition to the SC to avoid unwanted eye movements.[108]

5.2. Neural correlates of pupillary response to sounds

Pupil size is controlled by the balanced activity between the sympathetic constriction and parasympathetic dilation pathways (Fig. 4**B**; for a detailed review, see Ref 5). For many aspects of cognitive pupillary responses, the locus coeruleus-norepinephrine system (LC-NE) is regularly involved,[139] particularly when related to arousal and mental effort.[140] Locus coeruleus (LC) is the major source of noradrenergic innervation in the brain and it mediates cognition through arousal.[141] There is strong electrophysiological evidence for functional relationships between LC activation (i.e., arousal) and pupil change.[142] A human fMRI study[143] also demonstrated that pupil diameter covaries with activations in LC. Noradrenaline release not only accompanies pupil dilation, but its innervation from LC also affects processing in the auditory cortex.[144,145] The detailed cause-effect links are difficult to establish,[146] but it is possible that sound-specific features, such as saliency and regularity assessed in higher order cortical networks, could lead to changes in arousal which in turn affects pupil size.[70]

Another structure mediating the pupil pathways is the SC.[71,83,147] Having direct and indirect connections to the parasympathetic pupil circuit and receiving descending cognitive signals,[146] the SC likely inhibits pupil constriction and drives the pupil dilation that typically accompanies orienting responses, for example, to salient auditory events.[147] Characteristically, the integrative role of the SC is sometimes apparent in pupillary responses to multisensory stimuli. For example, audiovisual stimuli evoke larger, super-additive, pupil response magnitudes, compared to visual or auditory alone stimuli (see Refs 82, 84; but see Ref 85, for an additive account of pupil changes to audiovisual stimuli). Overall, projections from the SC to the pupil circuitry (Fig. 4**B**) may provide sound-relevant inputs to coordinate orienting-related pupil responses.

There is now accumulating evidence of cortical modulations on pupillary changes. For example, pupil dilation is modulated by subthreshold electrical microsimulation of the FEF in monkeys,[148] but it is unclear whether these effects are mediated via projections from FEF to SC or from FEF to the pretectal olivary nucleus (PON) in the sympathetic pupil pathway. Similarly, the auditory cortex may also play a role on sound-induced pupillary responses. Cortical processing of sound-specific features could lead to heightened arousal which may affect pupil dilation.[70] Yet the extent of pupil changes result from cortical auditory feedback is unclear.

5.3. Neural correlates of smooth pursuit

For moving visual targets, the smooth pursuit system is subserved by a cortical network involving the pursuit subregion of FEF[149] and the middle occipito-temporal area MT+/V5 that responds preferentially to visual motion.[150] These regions, among a few others (for a review, see Ref 104), project to the pontine nuclei that relays the information to ocular motor neurons in the cerebellum. Of these, hMT+/V5[151] is seen as the main visual motion processor that feeds the network with motion signals. In situations where there are unpredictable, or very fast moving, visual targets, the pursuit system collaborates with the saccadic system to avoid large position error and the eye lagging behind the target (see Ref 2, for an overview). For the auditory modality, human imaging studies using PET or fMRI have identified cortical areas with enhanced activity to sounds containing simulated sound-source motion. The human planum temporale (PT), for example, engages preferentially in the processing of moving sounds,[152-155] and its neural coding for the direction of sounds follows an axis-of-motion organization.[106] The middle temporal visual area (hMT+/V5) also represents auditory motion information to some extent,[156] with the anterior portion of hMT+/V5 showing motion selectivity for both vision and audition in sighted and blind people (Refs 105-107; but see Ref 155, for similar results only in the blind, but not in the sighted). Compared to visual stimuli, however, the ocular tracking of moving sounds is not continuous and composed only of saccades and no smooth pursuit (see Section 2), suggesting that moving sounds may not be able to activate the ocular pursuit pathway. Indeed, as noted by Middlebrooks,[157] it is difficult to distinguish cortical responses to continuous auditory motion (velocity computation) from cortical responses to changes in sound-source location (position computation), the latter potentially mistaken as evidence for sound motion sensitivity. Unlike visual motion detection which is supported by direction-selective neurons in early visual cortex,[151] processing of continuous sound motion is not as well defined on the neural level (for review, see Ref 158). There is a lack of evidence that auditory motion velocity signals are extracted (if at all) and fed to the smooth pursuit pathway. Rather, ocular tracking of moving sound sources appears to be a sequence of saccades (position pathway) that move the eyes in discrete steps along the motion path.[159,160]

5.4. Neural correlates of blinks

Spontaneous blinking is suppressed at the presentation of sounds and increases after sound offset. Neural explanations for this are lacking. It is possible that eye blink suppression shares some oculomotor inhibition mechanisms with microsaccades (the inhibition-release-rebound response pattern).[101] In visual fixation tasks during voluntary blink inhibition, fMRI-EOG evidence shows the involvement of bilateral parahippocampal cortex, visual cortex, and precentral gyrus.[161] It is unclear, however, how these areas might be involved in sound-induced *involuntary* blink inhibition. When environmental sounds are occasionally played in the background during a visual reading task,[162] the orbicularis oculi (the muscle controlling eyelid closure) shows a short-lasting inhibition of EMG (electromyography) activity followed by a gradual increase in EMG activity during auditory stimulation. Muscular motor activity again confirmed the inhibition-increase pattern. However, the neural correlates for these effects remain to be explored.

Summary

We summarize the key aspects of how sounds modulate eye responses, particularly in comparison to visual modulation of eye responses, and the potential contribution of the involved brain structures.

Auditory evoked saccades are less accurate than visual saccades[17-21] due to the lower spatial resolution of auditory space compared to visual space. Detecting sounds is initially faster than detecting visual stimuli, but saccade initiation is slower due to the transformation of auditory spatial cues into oculomotor commands in visual space. Auditory saccades are faster the further away they are from the initial eye position,[17,19-21,31,32] whereas visual saccades are faster the closer they are to the initial eye position (eccentricity effect). Some of this eccentricity effect is explained by the transformation of auditory cues from a head-centered reference frame into an eye-centered reference for saccade execution, likely mediated by the superior colliculus. The gap effect (faster saccades when the fixation dot is extinguished before target appearance) is present also for auditory saccades but there are no auditory express saccades. In both modalities, the gap effect cannot be explained by disengagement of attention but rather reflects ocular release of fixation.[44,48] Different sound features affect auditory evoked saccades, such that vertical auditory saccades necessitate spectral cues (present in noise but not in single tones) of sufficient duration. Horizontal auditory saccades are initiated quickest with sound frequencies similar to human speech,[24] and saccades can even be modulated by high-level grammatical aspects of speech.[52] Sounds also affect visual saccades: irrelevant sounds can inhibit visual saccades (distractor effect), and more so when sounds are predictable, but can also facilitate saccades to visual targets when presented synchronously (redundant target effect). The underlying multisensory interaction mechanisms of either effect are still debated.

Microsaccades are strongly inhibited by sounds, and this inhibition is modulated by attention[66-68] and sound saliency.[69,70] Multimodal effects on microsaccade inhibition point towards a potential involvement of the superior colliculus due to its role in both multisensory integration and microsaccade generation.

Smooth pursuit cannot be triggered by sounds alone; the oculomotor system rather reacts with a sequence of saccades to moving sounds. However, a static sound can help maintain smooth pursuit when the visual target disappears,[77] and catch-up saccades during fast visual tracking are inhibited by sounds (similar to the distractor effect).

Pupil size is modulated by sound intensity, saliency and regularity, by auditory task demands and by the semantic and syntactic processing demands of speech. Pupil responses to audiovisual stimuli are stronger than to unimodal stimuli,[71] but whether this indicates multisensory integration is still under debate.

Eye blinks are inhibited by sounds, similar to microsaccades, and this inhibition is modulated by sound predictability and cognitive evaluation of sounds. Eye blinks also synchronize to the high-level linguistic structure of spoken sentences.[103]

In terms of neural correlates, the superior colliculus (SC) is a critical structure in the generation of auditory saccades. Auditory receptive field of SC neurons are larger than visual receptive fields and are shifted during gaze orientation such that visual and auditory spatial SC maps align. The SC's multisensory neurons could contribute to saccade facilitation during the redundant target effect, as well as to microsaccade inhibition and pupil size changes, but the link between neural response properties measured in animals and oculomotor behavior in humans is still unclear. Another important midbrain nucleus is the inferior colliculus (IC). IC is implicated in coding sound azimuth and elevation,[123] receiving and sending inputs to and from SC. IC neurons are modulated by eye position, suggesting a role in integrating retinal and auditory space, and importantly, supporting SC for sound orienting. The frontal eye fields (FEF) also play a critical role in auditory saccades,[124] and head-centered auditory spatial cues are already aligned to eye-centered visual space at the FEF processing stage. FEF seems to play a context-dependent modulatory role in visual saccades, but evidence for its role in auditory saccades is still sparse. Dorsolateral prefrontal cortex (DLPFC) is critical for saccadic inhibition and the control of antisaccades, but its role in auditory saccades is underexplored.

Pupil responses, particularly related to arousal and mental effort, involve the locus coeruleus-norepinephrine system (LC-NE), but its role in auditory-evoked pupil response is again unclear. Also, the SC plays a modulatory role in pupil response to salient auditory stimuli (along with parallel orienting eye responses), and its multisensory integration properties may be linked to additive or super-additive pupil responses to multisensory stimuli.

The neural correlates of auditorily induced eye blink inhibition are still unknown.

Conclusions

This review provided evidence that sounds drive and modulate eye responses and often differently to the way visual stimuli drive eye responses. Sounds can speed up, slow down or inhibit saccades depending on spatial location, sound features and presence of co-occurring visual stimulation. Sounds also inhibit microsaccades and eye blinks, and this inhibition is modulated by attentional and cognitive factors. Pupil size is modulated by both low-level sensory features and high-level cognitive demands of auditory stimulation. Sounds alone cannot induce smooth pursuit but can modulate visually induced smooth pursuit. In response to audiovisual stimuli, saccades, microsaccades and pupil response show signatures of multisensory interaction, and potential multisensory integration, but the underlying computational and neural mechanisms are still unclear and require further investigation. Also, the involvement of the key neural correlates of visual eye responses in auditory-evoked eye responses are underexplored in many aspects. We summarize the key open questions for future research on auditory-evoked eye responses in **Box 1**. All in all, we can say that the auditory system interacts with the ocular-motor system in an intricate way, from the transformation of low spatial resolution, head-centered auditory cues to high-resolution, eye-centered oculomotor commands to the interaction between high-level cognitive content in speech with basic eye responses. These interactions between the auditory and the eye response systems are complex and likely to involve an interplay between several brain circuits, but the exact underlying neural mechanisms remain underexplored and call for further research.

Box 1

Open questions for future research on auditory evoked eye responses:

1. How do the eyes respond to sounds in rear space?
2. What are the characteristics and neural mechanisms of antisaccade response to sounds?
3. Is there an eccentricity effect of auditory saccades on the vertical plane?
4. How can the redundant target effect (i.e., the speeding up of visual saccades by an irrelevant sound) be explained?
5. Do pupil responses to audiovisual stimuli indicate multisensory integration?
6. Does the superior colliculus mediate multisensory effects of eye responses?
7. Are the FEF and/or STG involved in planning and execution of saccades to sounds?
8. What is the role of the DLPFC in sound induced saccadic inhibition?
9. How is the LC-NE system involved in eye responses to auditory stimuli?
10. What are the neural correlates for auditorily induced eye blink inhibition?
11. Does hearing sounds in the absence of visual stimulation modulate retinal processing, similar to eye movements modulating the eardrums (c.f., Ref 163)?

Acknowledgements

This work was supported by a PRIMA grant from the Swiss National Science Foundation to PV [PRIMA grant PR00P1_185918/1]. The authors wish to thank two anonymous reviewers for their helpful comments on this manuscript.

Author Contributions

JH performed the literature research, wrote the first draft and created the figures. PV edited and wrote further drafts and acquired the funding. JH and PV edited and wrote the final manuscript.

Conflicts of Interests

None.

Figure captions

Fig. 1. Schema of stimulus paradigms used for testing saccadic behavior.

Fig. 2. Human auditory space and ocular responses to sounds in darkness.

Fig. 3. Example measurements of auditory saccades in response to sound.

Fig. 4. Brain areas controlling saccade and pupil response.

Table

Table 1. Summary of auditory-evoked saccade studies in humans.

References

1. Pierrot-Deseilligny, C., Müri, R. M., Nyffeler, T., & Milea, D. (2005). The Role of the Human Dorsolateral Prefrontal Cortex in Ocular Motor Behavior. *Annals of the New York Academy of Sciences*, *1039*(1), 239–251. https://doi.org/10.1196/annals.1325.023
2. Orban De Xivry, J., & Lefèvre, P. (2007). Saccades and pursuit: Two outcomes of a single sensorimotor process. *The Journal of Physiology*, *584*(1), 11–23. https://doi.org/10.1113/jphysiol.2007.139881
3. Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, *14*(2), 83–96. https://doi.org/10.1038/nrn3405
4. Einhäuser, W. (2017). The Pupil as Marker of Cognitive Processes. In Q. Zhao (Ed.), *Computational and Cognitive Neuroscience of Vision* (pp. 141–169). Springer Singapore. https://doi.org/10.1007/978-981-10-0213-7_7
5. Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, *1*(1), 16. https://doi.org/10.5334/joc.18

6. Binda, P., & Morrone, M. C. (2018). Vision During Saccadic Eye Movements. *Annual Review of Vision Science*, *4*(1), 193–213. https://doi.org/10.1146/annurev-vision-091517-034317
7. Mahanama, B., Jayawardana, Y., Rengarajan, S., Jayawardena, G., Chukoskie, L., Snider, J., & Jayarathna, S. (2022). Eye Movement and Pupil Measures: A Review. *Frontiers in Computer Science*, *3*, 733531. https://doi.org/10.3389/fcomp.2021.733531
8. Rayleigh, Lord. (1907). XII. *On our perception of sound direction. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *13*(74), 214–232. https://doi.org/10.1080/14786440709463595
9. Batteau, D. W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *168*(1011), 158–180. https://doi.org/10.1098/rspb.1967.0058
10. Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *The Journal of the Acoustical Society of America*, *75*(4), 1195–1200. https://doi.org/10.1121/1.390770
11. Yin, T. C. T. (2002). Neural Mechanisms of Encoding Binaural Localization Cues in the Auditory Brainstem. In D. Oertel, R. R. Fay, & A. N. Popper (Eds.), *Integrative Functions in the Mammalian Auditory Pathway* (Vol. 15, pp. 99–159). Springer New York. https://doi.org/10.1007/978-1-4757-3654-0_4
12. Young, E. D., & Davis, K. A. (2002). Circuitry and Function of the Dorsal Cochlear Nucleus. In D. Oertel, R. R. Fay, & A. N. Popper (Eds.), *Integrative Functions in the Mammalian Auditory Pathway* (Vol. 15, pp. 160–206). Springer New York. https://doi.org/10.1007/978-1-4757-3654-0_5
13. Spence, C., Lee, J., & Van Der Stoep, N. (2020). Responding to sounds from unseen locations: Crossmodal attentional orienting in response to sounds presented from the rear. *European Journal of Neuroscience*, *51*(5), 1137–1150. https://doi.org/10.1111/ejn.13733

14. Lewald, J., & Ehrenstein, W. H. (2001). Effect of gaze direction on sound localization in rear space. *Neuroscience Research*, *39*(2), 253–257. https://doi.org/10.1016/S0168-0102(00)00210-8
15. Amlôt, R., & Walker, R. (2006). Are somatosensory saccades voluntary or reflexive? *Experimental Brain Research*, *168*(4), 557–565. https://doi.org/10.1007/s00221-005-0116-9
16. Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1039–1054. https://doi.org/10.1037/0096-1523.28.5.1039
17. Zahn, J. R., Abel, L. A., & Dell'Osso, L. F. (1978). Audio-ocular response characteristics. *Sensory Processes*, *2*(1), 32–37.
18. Zahn, J. R., Abel, L. A., Dell 'Osso, L. F., & Daroff, R. B. (1979). The audioocular response: Intersensory delay. *Sensory Processes*, *3*(1), 60–65.
19. Yao, L., & Peck, C. K. (1997). Saccadic eye movements to visual and auditory targets: *Experimental Brain Research*, *115*(1), 25–34. https://doi.org/10.1007/PL00005682
20. Jay, M. F., & Sparks, D. L. (1990). Localization of auditory and visual targets for the initiation of saccadic eye movements. In *Comparative perception* (Vol. 1, Ser. Basic mechanisms, pp. 351–374). essay, John Wiley & Sons.
21. Frens, M. A., & Van Opstal, A. J. (1995). A quantitative study of auditory-evoked saccadic eye movements in two dimensions. *Experimental Brain Research*, *107*(1), 103–117. https://doi.org/10.1007/BF00228022
22. Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R., Klatzky, R. L., & Merfeld, D. M. (2021). *Sensation & perception* (Sixth edition). Sinauer Associates, Oxford University Press.
23. Mills, A. W. (1958). On the Minimum Audible Angle. *The Journal of the Acoustical Society of America*, *30*(4), 237–246. https://doi.org/10.1121/1.1909553

24. Gabriel, D. N., Munoz, D. P., & Boehnke, S. E. (2010). The eccentricity effect for auditory saccadic reaction times is independent of target frequency. *Hearing Research*, *262*(1–2), 19–25. https://doi.org/10.1016/j.heares.2010.01.016
25. Zambarbieri, D., Schmid, R., Magenes, G., & Prablanc, C. (1982). Saccadic responses evoked by presentation of visual and auditory targets. *Experimental Brain Research*, *47*(3). https://doi.org/10.1007/BF00239359
26. Schooler, C., Roberts, B., & Cohen, R. (2008). Context, complexity, and cognitive processing in schizophrenia. *Cognitive Neuropsychiatry*, *13*(3), 250–266. https://doi.org/10.1080/13546800802058658
27. Fecteau, J., & Munoz, D. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. https://doi.org/10.1016/j.tics.2006.06.011
28. Amlôt, R., & Walker, R. (2011). *Multisensory saccade generation*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0027
29. Shelton, J., & Kumar, G. P. (2010). Comparison between Auditory and Visual Simple Reaction Times. *Neuroscience and Medicine*, *01*(01), 30–32. https://doi.org/10.4236/nm.2010.11004
30. Jain, A., Bansal, R., Kumar, A., & Singh, K. (2015). A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied and Basic Medical Research*, *5*(2), 124. https://doi.org/10.4103/2229-516X.157168
31. Frens, M. A., & Van Opstal, A. J. (1994). Auditory-Evoked Saccades in Two Dimensions: Dynamical Characteristics, Influence of Eye Position and Sound Spectrum. In *Information Processing Underlying Gaze Control* (pp. 329–339). Elsevier. https://doi.org/10.1016/B978-0-08-042506-1.50037-0
32. Zambarbieri, D. (2002). The latency of saccades toward auditory targets in humans. In *Progress in Brain Research* (Vol. 140, pp. 51–59). Elsevier. https://doi.org/10.1016/S0079-6123(02)40041-6

33. Cohen, M. E., & Ross, L. E. (1977). Saccade latency in children and adults: Effects of warning interval and target eccentricity. *Journal of Experimental Child Psychology*, *23*(3), 539–549. https://doi.org/10.1016/0022-0965(77)90044-3
34. Lueck, C. J., Crawford, T. J., Savage, C. J., & Kennard, C. (1990). Auditory-visual interaction in the generation of saccades in man. *Experimental Brain Research*, *82*(1). https://doi.org/10.1007/BF00230846
35. Werner-Reiss, U., Kelly, K. A., Trause, A. S., Underhill, A. M., & Groh, J. M. (2003). Eye Position Affects Activity in Primary Auditory Cortex of Primates. *Current Biology*, *13*(7), 554–562. https://doi.org/10.1016/S0960-9822(03)00168-4
36. 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*(4), 2331–2352. https://doi.org/10.1152/jn.00021.2005
37. Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, *309*(5966), 345–347. https://doi.org/10.1038/309345a0
38. Jay, M. F., & Sparks, D. L. (1987). Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *Journal of Neurophysiology*, *57*(1), 22–34. https://doi.org/10.1152/jn.1987.57.1.22
39. Lee, J., & Groh, J. M. (2012). Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus. *Journal of Neurophysiology*, *108*(1), 227–242. https://doi.org/10.1152/jn.00706.2011
40. 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*(2), 509–518. https://doi.org/10.1016/S0896-6273(01)00222-7
41. Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.
42. 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*(4), 2622–2627. https://doi.org/10.1152/jn.00326.2004

43. Saslow, M. G. (1967). Effects of Components of Displacement-Step Stimuli Upon Latency for Saccadic Eye Movement. *Journal of the Optical Society of America*, *57*(8), 1024. https://doi.org/10.1364/JOSA.57.001024
44. Shafiq, R., Stuart, G. W., Sandbach, J., Maruff, P., & Currie, J. (1998). The gap effect and express saccades in the auditory modality. *Experimental Brain Research*, *118*(2), 221–229. https://doi.org/10.1007/s002210050275
45. Taylor, T. L., Klein, R. M., & Munoz, D. P. (1999). Saccadic Performance as a Function of the Presence and Disappearance of Auditory and Visual Fixation Stimuli. *Journal of Cognitive Neuroscience*, *11*(2), 206–213. https://doi.org/10.1162/089892999563337
46. Kingstone, A., & Klein, R. M. (1993). Visual offsets facilitate saccadic latency: Does predisengagement of visuospatial attention mediate this gap effect? *Journal of Experimental Psychology: Human Perception and Performance*, *19*(6), 1251–1265. https://doi.org/10.1037/0096-1523.19.6.1251
47. Fischer, B., & Weber, H. (1993). Express saccades and visual attention. *Behavioral and Brain Sciences*, *16*(3), 553–567. https://doi.org/10.1017/S0140525X00031575
48. Klein, R. M., Taylor, T. L., & Kingstone, A. (1995). Against a role for attentional disengagement in the gap effect: A friendly amendment to Tam and Stelmach (1993). *Perception & Psychophysics*, *57*(4), 573–577. https://doi.org/10.3758/BF03213082
49. Munoz, D. P., & Wurtz, R. H. (1992). Role of the rostral superior colliculus in active visual fixation and execution of express saccades. *Journal of Neurophysiology*, *67*(4), 1000–1002. https://doi.org/10.1152/jn.1992.67.4.1000
50. Goffart, L., Hafed, Z. M., & Krauzlis, R. J. (2012). Visual Fixation as Equilibrium: Evidence from Superior Colliculus Inactivation. *Journal of Neuroscience*, *32*(31), 10627–10636. https://doi.org/10.1523/JNEUROSCI.0696-12.2012
51. Schuller, G., & Pollak, G. (1979). Disproportionate frequency representation in the inferior colliculus of doppler-compensating Greater Horseshoe bats: Evidence

for an acoustic fovea. *Journal of Comparative Physiology ? A*, *132*(1), 47–54. https://doi.org/10.1007/BF00617731

52. Huette, S., Winter, B., Matlock, T., Ardell, D. H., & Spivey, M. (2014). Eye movements during listening reveal spontaneous grammatical processing. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00410
53. Graupner, S., Velichkovsky, B. M., Pannasch, S., & Marx, J. (2007). Surprise, surprise: Two distinct components in the visually evoked distractor effect. *Psychophysiology*, *44*(2), 251–261. https://doi.org/10.1111/j.1469-8986.2007.00504.x
54. Pannasch, S., Dornhoefer, S. M., Unema, P. J. A., & Velichkovsky, B. M. (2001). The omnipresent prolongation of visual fixations: Saccades are inhibited by changes in situation and in subject's activity. *Vision Research*, *41*(25–26), 3345–3351. https://doi.org/10.1016/S0042-6989(01)00207-3
55. Abeles, D., Amit, R., Tal-Perry, N., Carrasco, M., & Yuval-Greenberg, S. (2020). Oculomotor inhibition precedes temporally expected auditory targets. *Nature Communications*, *11*(1), 3524. https://doi.org/10.1038/s41467-020-17158-9
56. Hughes, H. C., Nelson, M. D., & Aronchick, D. M. (1998). Spatial characteristics of visual-auditory summation in human saccades. *Vision Research*, *38*(24), 3955–3963. https://doi.org/10.1016/S0042-6989(98)00036-4
57. Diederich, A., & Colonius, H. (2007). Why two "Distractors" are better than one: Modeling the effect of non-target auditory and tactile stimuli on visual saccadic reaction time. *Experimental Brain Research*, *179*(1), 43–54. https://doi.org/10.1007/s00221-006-0768-0
58. Raab, D. H. (1962). Division of psychology: statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*(5 Series II), 574–590. https://doi.org/10.1111/j.2164-0947.1962.tb01433.x
59. Miller, J. (2016). Statistical facilitation and the redundant signals effect: What are race and coactivation models? *Attention, Perception, & Psychophysics*, *78*(2), 516–519. https://doi.org/10.3758/s13414-015-1017-z

60. Gondan, M., & Minakata, K. (2016). A tutorial on testing the race model inequality. *Attention, Perception, & Psychophysics*, *78*(3), 723–735. https://doi.org/10.3758/s13414-015-1018-y
61. Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. In *Progress in Brain Research* (Vol. 154, pp. 177–192). Elsevier. https://doi.org/10.1016/S0079-6123(06)54009-9
62. Engbert, R., & Kliegl, R. (2004). Microsaccades Keep the Eyes' Balance During Fixation. *Psychological Science*, *15*(6), 431–431. https://doi.org/10.1111/j.0956-7976.2004.00697.x
63. Rolfs, M., Engbert, R., & Kliegl, R. (2005). Crossmodal coupling of oculomotor control and spatial attention in vision and audition. *Experimental Brain Research*, *166*(3–4), 427–439. https://doi.org/10.1007/s00221-005-2382-y
64. Rolfs, M., Kliegl, R., & Engbert, R. (2008). Toward a model of microsaccade generation: The case of microsaccadic inhibition. *Journal of Vision*, *8*(11), 5–5. https://doi.org/10.1167/8.11.5
65. Hafed, Z. M. (2011). Mechanisms for generating and compensating for the smallest possible saccades. *European Journal of Neuroscience*, *33*(11), 2101–2113. https://doi.org/10.1111/j.1460-9568.2011.07694.x
66. Valsecchi, M., & Turatto, M. (2009). Microsaccadic responses in a bimodal oddball task. *Psychological Research*, *73*(1), 23–33. https://doi.org/10.1007/s00426-008-0142-x
67. Widmann, A., Engbert, R., & Schroger, E. (2014). Microsaccadic Responses Indicate Fast Categorization of Sounds: A Novel Approach to Study Auditory Cognition. *Journal of Neuroscience*, *34*(33), 11152–11158. https://doi.org/10.1523/JNEUROSCI.1568-14.2014
68. Yuval-Greenberg, S., & Deouell, L. Y. (2011). Scalp-Recorded Induced Gamma-Band Responses to Auditory Stimulation and Its Correlations with Saccadic Muscle-Activity. *Brain Topography*, *24*(1), 30–39. https://doi.org/10.1007/s10548-010-0157-7

69. Kadosh, O., & Bonneh, Y. S. (2022). Involuntary oculomotor inhibition markers of saliency and deviance in response to auditory sequences. *Journal of Vision*, *22*(5), 8. https://doi.org/10.1167/jov.22.5.8
70. Zhao, S., Yum, N. W., Benjamin, L., Benhamou, E., Yoneya, M., Furukawa, S., Dick, F., Slaney, M., & Chait, M. (2019). Rapid Ocular Responses Are Modulated by Bottom-up-Driven Auditory Salience. *The Journal of Neuroscience*, *39*(39), 7703–7714. https://doi.org/10.1523/JNEUROSCI.0776-19.2019
71. Wang, C.-A., Blohm, G., Huang, J., Boehnke, S. E., & Munoz, D. P. (2017). Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biological Psychology*, *129*, 36–44. https://doi.org/10.1016/j.biopsycho.2017.07.024
72. Braun, D. I., Pracejus, L., & Gegenfurtner, K. R. (2006). Motion aftereffect elicits smooth pursuit eye movements. *Journal of Vision*, *6*(7), 1. https://doi.org/10.1167/6.7.1
73. Steinbach, M. J. (1969). Eye tracking of self-moved targets: The role of efference. *Journal of Experimental Psychology*, *82*(2), 366–376. https://doi.org/10.1037/h0028115
74. Gauthier, G. M., & Hofferer, J. M. (1976). Eye tracking of self-moved targets in the absence of vision. *Experimental Brain Research*, *26*(2). https://doi.org/10.1007/BF00238277
75. Boucher, L., Lee, A., Cohen, Y. E., & Hughes, H. C. (2004). Ocular tracking as a measure of auditory motion perception. *Journal of Physiology-Paris*, *98*(1–3), 235–248. https://doi.org/10.1016/j.jphysparis.2004.03.010
76. Berryhill, M. E., Chiu, T., & Hughes, H. C. (2006). Smooth Pursuit of Nonvisual Motion. *Journal of Neurophysiology*, *96*(1), 461–465. https://doi.org/10.1152/jn.00152.2006
77. Madelain, L., & Krauzlis, R. J. (2003). Effects of Learning on Smooth Pursuit During Transient Disappearance of a Visual Target. *Journal of Neurophysiology*, *90*(2), 972–982. https://doi.org/10.1152/jn.00869.2002

78. de Brouwer, S., Yuksel, D., Blohm, G., Missal, M., & Lefèvre, P. (2002). What Triggers Catch-Up Saccades During Visual Tracking? *Journal of Neurophysiology*, *87*(3), 1646–1650. https://doi.org/10.1152/jn.00432.2001
79. Kerzel, D., Born, S., & Souto, D. (2010). Inhibition of Steady-State Smooth Pursuit and Catch-Up Saccades by Abrupt Visual and Auditory Onsets. *Journal of Neurophysiology*, *104*(5), 2573–2585. https://doi.org/10.1152/jn.00193.2010
80. Sokolov, E. N. (1963). Higher Nervous Functions: The Orienting Reflex. *Annual Review of Physiology*, *25*(1), 545–580. https://doi.org/10.1146/annurev.ph.25.030163.002553
81. Zénon, A. (2019). Eye pupil signals information gain. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191593. https://doi.org/10.1098/rspb.2019.1593
82. Wang, C.-A., Boehnke, S. E., Itti, L., & Munoz, D. P. (2014). Transient Pupil Response Is Modulated by Contrast-Based Saliency. *The Journal of Neuroscience*, *34*(2), 408–417. https://doi.org/10.1523/JNEUROSCI.3550-13.2014
83. Liao, H.-I., Kidani, S., Yoneya, M., Kashino, M., & Furukawa, S. (2016). Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychonomic Bulletin & Review*, *23*(2), 412–425. https://doi.org/10.3758/s13423-015-0898-0
84. Rigato, S., Rieger, G., & Romei, V. (2016). Multisensory signalling enhances pupil dilation. *Scientific Reports*, *6*(1), 26188. https://doi.org/10.1038/srep26188
85. van der Stoep, N., Van Der Smagt, M. J., Notaro, C., Spock, Z., & Naber, M. (2021). The additive nature of the human multisensory evoked pupil response. *Scientific Reports*, *11*(1), 707. https://doi.org/10.1038/s41598-020-80286-1
86. de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

87. Richter, D., Ekman, M., & De Lange, F. P. (2017). *Suppressed sensory response to predictable object stimuli throughout the ventral visual stream* [Preprint]. Neuroscience. https://doi.org/10.1101/228890

88. Milne, A. E., Zhao, S., Tampakaki, C., Bury, G., & Chait, M. (2021). Sustained Pupil Responses Are Modulated by Predictability of Auditory Sequences. *The Journal of Neuroscience*, *41*(28), 6116–6127. https://doi.org/10.1523/JNEUROSCI.2879-20.2021

89. Southwell, R., Baumann, A., Gal, C., Barascud, N., Friston, K., & Chait, M. (2017). Is predictability salient? A study of attentional capture by auditory patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714), 20160105. https://doi.org/10.1098/rstb.2016.0105

90. Zhao, S., Chait, M., Dick, F., Dayan, P., Furukawa, S., & Liao, H.-I. (2019). Pupil-linked phasic arousal evoked by violation but not emergence of regularity within rapid sound sequences. *Nature Communications*, *10*(1), 4030. https://doi.org/10.1038/s41467-019-12048-1

91. Schluroff, M. (1982). Pupil responses to grammatical complexity of sentences. *Brain and Language*, *17*(1), 133–145. https://doi.org/10.1016/0093-934X(82)90010-4

92. Wendt, D., Dau, T., & Hjortkjær, J. (2016). Impact of Background Noise and Sentence Complexity on Processing Demands during Sentence Comprehension. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00345

93. Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation. *Ear & Hearing*, *36*(4), e153–e165. https://doi.org/10.1097/AUD.0000000000000145

94. Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., De Lissa, P., Graham, P., & Lyxell, B. (2017). Objective Assessment of Listening Effort: Coregistration of Pupillometry and EEG. *Trends in Hearing*, *21*, 233121651770639. https://doi.org/10.1177/2331216517706396

95. Kadem, M., Herrmann, B., Rodd, J. M., & Johnsrude, I. S. (2020). Pupil Dilation Is Sensitive to Semantic Ambiguity and Acoustic Degradation. *Trends in Hearing*, *24*, 233121652096406. https://doi.org/10.1177/2331216520964068
96. Doughty, M. J., & Naase, T. (2006). Further Analysis of the Human Spontaneous Eye Blink Rate by a Cluster Analysis-Based Approach to Categorize Individuals With 'Normal' Versus 'Frequent' Eye Blink Activity. *Eye & Contact Lens: Science & Clinical Practice*, *32*(6), 294–299. https://doi.org/10.1097/01.icl.0000224359.32709.4d
97. Sweeney, D. F., Millar, T. J., & Raju, S. R. (2013). Tear film stability: A review. *Experimental Eye Research*, *117*, 28–38. https://doi.org/10.1016/j.exer.2013.08.010
98. Fukuda, K. (1994). Analysis of Eyeblink Activity during Discriminative Tasks. *Perceptual and Motor Skills*, *79*(3_suppl), 1599–1608. https://doi.org/10.2466/pms.1994.79.3f.1599
99. Kobald, S. O., Wascher, E., Heppner, H., & Getzmann, S. (2019). Eye blinks are related to auditory information processing: Evidence from a complex speech perception task. *Psychological Research*, *83*(6), 1281–1291. https://doi.org/10.1007/s00426-017-0952-9
100. Oh, J., Jeong, S.-Y., & Jeong, J. (2012). The timing and temporal patterns of eye blinking are dynamically modulated by attention. *Human Movement Science*, *31*(6), 1353–1365. https://doi.org/10.1016/j.humov.2012.06.003
101. Bonneh, Y. S., Adini, Y., & Polat, U. (2016). Contrast sensitivity revealed by spontaneous eyeblinks: Evidence for a common mechanism of oculomotor inhibition. *Journal of Vision*, *16*(7), 1. https://doi.org/10.1167/16.7.1
102. Huber, S. E., Martini, M., & Sachse, P. (2022). Patterns of eye blinks are modulated by auditory input in humans. *Cognition*, *221*, 104982. https://doi.org/10.1016/j.cognition.2021.104982
103. Jin, P., Zou, J., Zhou, T., & Ding, N. (2018). Eye activity tracks task-relevant structures during speech and auditory sequence perception. *Nature Communications*, *9*(1), 5374. https://doi.org/10.1038/s41467-018-07773-y

104. Leigh, R. J., & Zee, D. S. (2015). *The Neurology of Eye Movements* (5th ed.). Oxford University Press. https://doi.org/10.1093/med/9780199969289.001.0001
105. Dormal, G., Rezk, M., Yakobov, E., Lepore, F., & Collignon, O. (2016). Auditory motion in the sighted and blind: Early visual deprivation triggers a large-scale imbalance between auditory and "visual" brain regions. *NeuroImage*, *134*, 630–644. https://doi.org/10.1016/j.neuroimage.2016.04.027
106. Battal, C., Rezk, M., Mattioni, S., Vadlamudi, J., & Collignon, O. (2019). Representation of Auditory Motion Directions and Sound Source Locations in the Human Planum Temporale. *The Journal of Neuroscience*, *39*(12), 2208–2220. https://doi.org/10.1523/JNEUROSCI.2289-18.2018
107. Rezk, M., Cattoir, S., Battal, C., Occelli, V., Mattioni, S., & Collignon, O. (2020). Shared Representation of Visual and Auditory Motion Directions in the Human Middle-Temporal Cortex. *Current Biology*, *30*(12), 2289-2299.e8. https://doi.org/10.1016/j.cub.2020.04.039
108. Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*(3), 218–228. https://doi.org/10.1038/nrn1345
109. Munoz, D. P., & Wurtz, R. H. (1995a). Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. *Journal of Neurophysiology*, *73*(6), 2313–2333. https://doi.org/10.1152/jn.1995.73.6.2313
110. Munoz, D. P., & Wurtz, R. H. (1995b). Saccade-related activity in monkey superior colliculus. II. Spread of activity during saccades. *Journal of Neurophysiology*, *73*(6), 2334–2348. https://doi.org/10.1152/jn.1995.73.6.2334
111. Schall, J. D. (1995). Neural Basis of Saccade Target Selection. *Reviews in the Neurosciences*, *6*(1). https://doi.org/10.1515/REVNEURO.1995.6.1.63
112. King, A. J., Schnupp, J. W. H., Carlile, S., Smith, A. L., & Thompson, I. D. (1996). Chapter 24 The development of topographically-aligned maps of visual and auditory space in the superior colliculus. In *Progress in Brain Research* (Vol. 112, pp. 335–350). Elsevier. https://doi.org/10.1016/S0079-6123(08)63340-3

113. Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662. https://doi.org/10.1152/jn.1986.56.3.640
114. Goodale, M. A., & Murison, R. C. C. (1975). The effects of lesions of the superior colliculus on locomotor orientation and the orienting reflex in the rat. *Brain Research*, *88*(2), 243–261. https://doi.org/10.1016/0006-8993(75)90388-1
115. Jiang, W., & Stein, B. E. (2003). Cortex Controls Multisensory Depression in Superior Colliculus. *Journal of Neurophysiology*, *90*(4), 2123–2135. https://doi.org/10.1152/jn.00369.2003
116. Casseday, J. H., Fremouw, T., & Covey, E. (2002). The Inferior Colliculus: A Hub for the Central Auditory System. In D. Oertel, R. R. Fay, & A. N. Popper (Eds.), *Integrative Functions in the Mammalian Auditory Pathway* (Vol. 15, pp. 238–318). Springer New York. https://doi.org/10.1007/978-1-4757-3654-0_7
117. Winer, J. A., & Schreiner, C. E. (Eds.). (2005). *The Inferior Colliculus*. Springer New York. https://doi.org/10.1007/b138578
118. Gruters, K. G., & Groh, J. M. (2012). Sounds and beyond: Multisensory and other non-auditory signals in the inferior colliculus. *Frontiers in Neural Circuits*, *6*. https://doi.org/10.3389/fncir.2012.00096
119. Wallace, M. T., Meredith, M. A., & Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *Journal of Neurophysiology*, *69*(6), 1797–1809. https://doi.org/10.1152/jn.1993.69.6.1797
120. Andersen, R. A., Roth, G. L., Aitkin, L. M., & Merzenich, M. M. (1980). The efferent projections of the central nucleus and the pericentral nucleus of the inferior collculus in the cat. *Journal of Comparative Neurology*, *194*(3), 649–662. https://doi.org/10.1002/cne.901940311
121. Harting, J., & Van Lieshout, D. P. (2000). Projections from the rostral pole of the inferior colliculus to the cat superior colliculus. *Brain Research*, *881*(2), 244–247. https://doi.org/10.1016/S0006-8993(00)02849-3

122. King, A. J., Jiang, Z. D., & Moore, D. R. (1998). Auditory brainstem projections to the ferret superior colliculus: Anatomical contribution to the neural coding of sound azimuth. *The Journal of Comparative Neurology*, *390*(3), 342–365. https://doi.org/10.1002/(SICI)1096-9861(19980119)390:3<342::AID-CNE4>3.0.CO;2-1

123. Zwiers, M. P., Versnel, H., & Van Opstal, A. J. (2004). Involvement of Monkey Inferior Colliculus in Spatial Hearing. *The Journal of Neuroscience*, *24*(17), 4145–4156. https://doi.org/10.1523/JNEUROSCI.0199-04.2004

124. Russo, G. S., & Bruce, C. J. (1994). Frontal eye field activity preceding aurally guided saccades. *Journal of Neurophysiology*, *71*(3), 1250–1253. https://doi.org/10.1152/jn.1994.71.3.1250

125. Leichnetz, G. R., Smith, D. J., & Spencer, R. F. (1984). Cortical projections to the paramedian tegmental and basilar pons in the monkey. *Journal of Comparative Neurology*, *228*(3), 388–408. https://doi.org/10.1002/cne.902280307

126. Segraves, M. A. (1992). Activity of monkey frontal eye field neurons projecting to oculomotor regions of the pons. *Journal of Neurophysiology*, *68*(6), 1967–1985. https://doi.org/10.1152/jn.1992.68.6.1967

127. Schlag-Rey, M., Schlag, J., & Dassonville, P. (1992). How the frontal eye field can impose a saccade goal on superior colliculus neurons. *Journal of Neurophysiology*, *67*(4), 1003–1005. https://doi.org/10.1152/jn.1992.67.4.1003

128. Meredith, M. A. (1999). The frontal eye fields target multisensory neurons in cat superior colliculus. *Experimental Brain Research*, *128*(4), 460–470. https://doi.org/10.1007/s002210050869

129. Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., & Valero-Cabré, A. (2014). Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Frontiers in Integrative Neuroscience*, *8*. https://doi.org/10.3389/fnint.2014.00066

130. Garg, A., Schwartz, D., & Stevens, A. A. (2007). Orienting auditory spatial attention engages frontal eye fields and medial occipital cortex in congenitally

blind humans. *Neuropsychologia*, *45*(10), 2307–2321. https://doi.org/10.1016/j.neuropsychologia.2007.02.015

131. Barbas, H., & Mesulam, M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. *Journal of Comparative Neurology*, *200*(3), 407–431. https://doi.org/10.1002/cne.902000309
132. Leszczynski, M., Bickel, S., Nentwich, M., Russ, B. E., Parra, L., Lakatos, P., Mehta, A., & Schroeder, C. E. (2023). Saccadic modulation of neural excitability in auditory areas of the neocortex. *Current Biology*, *33*(7), 1185-1195.e6. https://doi.org/10.1016/j.cub.2023.02.018
133. Pierrot-Deseilligny, Ch., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of reflexive visually-guided saccades. *Brain*, *114*(3), 1473–1485. https://doi.org/10.1093/brain/114.3.1473
134. Hasegawa, R. P., Peterson, B. W., & Goldberg, M. E. (2004). Prefrontal Neurons Coding Suppression of Specific Saccades. *Neuron*, *43*(3), 415–425. https://doi.org/10.1016/j.neuron.2004.07.013
135. de Souza, J. F. X., Menon, R. S., & Everling, S. (2003). Preparatory Set Associated With Pro-Saccades and Anti-Saccades in Humans Investigated With Event-Related fMRI. *Journal of Neurophysiology*, *89*(2), 1016–1023. https://doi.org/10.1152/jn.00562.2002
136. Gaymard, B., François, C., Ploner, C. J., Condy, C., & Rivaud-Péchoux, S. (2003). A direct prefrontotectal tract against distractibility in the human brain. *Annals of Neurology*, *53*(4), 542–545. https://doi.org/10.1002/ana.10560
137. Johnston, K., & Everling, S. (2006). Monkey Dorsolateral Prefrontal Cortex Sends Task-Selective Signals Directly to the Superior Colliculus. *The Journal of Neuroscience*, *26*(48), 12471–12478. https://doi.org/10.1523/JNEUROSCI.4101-06.2006
138. Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the Basal Ganglia in the Control of Purposive Saccadic Eye Movements. *Physiological Reviews*, *80*(3), 953–978. https://doi.org/10.1152/physrev.2000.80.3.953

139. Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709
140. Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology*, *52*(1), 77–86. https://doi.org/10.1016/j.ijpsycho.2003.12.005
141. Sara, S. J., & Bouret, S. (2012). Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. *Neuron*, *76*(1), 130–141. https://doi.org/10.1016/j.neuron.2012.09.011
142. Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028
143. Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. https://doi.org/10.1002/hbm.22466
144. McGinley, M. J., David, S. V., & McCormick, D. A. (2015). Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron*, *87*(1), 179–192. https://doi.org/10.1016/j.neuron.2015.05.038
145. Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron*, *84*(2), 355–362. https://doi.org/10.1016/j.neuron.2014.09.033
146. Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences*, *24*(6), 466–480. https://doi.org/10.1016/j.tics.2020.03.005
147. Wang, C.-A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, *33*, 134–140. https://doi.org/10.1016/j.conb.2015.03.018

148. Ebitz, R. B., & Moore, T. (2017). Selective Modulation of the Pupil Light Reflex by Microstimulation of Prefrontal Cortex. *The Journal of Neuroscience*, *37*(19), 5008–5018. https://doi.org/10.1523/JNEUROSCI.2433-16.2017

149. Gottlieb, J. P., MacAvoy, M. G., & Bruce, C. J. (1994). Neural responses related to smooth-pursuit eye movements and their correspondence with electrically elicited smooth eye movements in the primate frontal eye field. *Journal of Neurophysiology*, *72*(4), 1634–1653. https://doi.org/10.1152/jn.1994.72.4.1634

150. Newsome, W., & Pare, E. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, *8*(6), 2201–2211. https://doi.org/10.1523/JNEUROSCI.08-06-02201.1988

151. Zeki, S., Watson, J., Lueck, C., Friston, K., Kennard, C., & Frackowiak, R. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience*, *11*(3), 641–649. https://doi.org/10.1523/JNEUROSCI.11-03-00641.1991

152. Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H.-J., & Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature*, *400*(6746), 724–726. https://doi.org/10.1038/23390

153. Griffiths, T. D., Bench, C. J., & Frackowiak, R. S. J. (1994). Human cortical areas selectively activated by apparent sound movement. *Current Biology*, *4*(10), 892–895. https://doi.org/10.1016/S0960-9822(00)00198-6

154. Griffiths, T. D., Rees, G., Rees, A., Green, G. G. R., Witton, C., Rowe, D., Büchel, C., Turner, R., & Frackowiak, R. S. J. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nature Neuroscience*, *1*(1), 74–79. https://doi.org/10.1038/276

155. Jiang, F., Stecker, G. C., & Fine, I. (2014). Auditory motion processing after early blindness. *Journal of Vision*, *14*(13), 4–4. https://doi.org/10.1167/14.13.4

156. Poirier, C., Collignon, O., De Volder, A. G., Renier, L., Vanlierde, A., Tranduy, D., & Scheiber, C. (2005). Specific activation of the V5 brain area by auditory motion

processing: An fMRI study. *Cognitive Brain Research*, *25*(3), 650–658. https://doi.org/10.1016/j.cogbrainres.2005.08.015

157. Middlebrooks, J. C. (2015). Sound localization. In *Handbook of Clinical Neurology* (Vol. 129, pp. 99–116). Elsevier. https://doi.org/10.1016/B978-0-444-62630-1.00006-8
158. Chaplin, T. A., Rosa, M. G. P., & Lui, L. L. (2018). Auditory and Visual Motion Processing and Integration in the Primate Cerebral Cortex. *Frontiers in Neural Circuits*, *12*, 93. https://doi.org/10.3389/fncir.2018.00093
159. Middlebrooks, J. C., & Green, D. M. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology*, *42*(1), 135–159. https://doi.org/10.1146/annurev.ps.42.020191.001031
160. Ahissar, M., Ahissar, E., Bergman, H., & Vaadia, E. (1992). Encoding of sound-source location and movement: Activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *Journal of Neurophysiology*, *67*(1), 203–215. https://doi.org/10.1152/jn.1992.67.1.203
161. Yoon, H. W., Chung, J.-Y., Song, M.-S., & Park, H. (2005). Neural correlates of eye blinking; improved by simultaneous fMRI and EOG measurement. *Neuroscience Letters*, *381*(1–2), 26–30. https://doi.org/10.1016/j.neulet.2005.01.077
162. Stekelenburg, J. J., & Van Boxtel, A. (2002). Pericranial muscular, respiratory, and heart rate components of the orienting response. *Psychophysiology*, *39*(6), 707–722. https://doi.org/10.1111/1469-8986.3960707
163. Gruters, K. G., Murphy, D. L. K., Jenson, C. D., Smith, D. W., Shera, C. A., & Groh, J. M. (2018). The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing. *Proceedings of the National Academy of Sciences*, *115*(6). https://doi.org/10.1073/pnas.1717948115
164. Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.14, retrieved 24 May 2022 from http://www.praat.org/