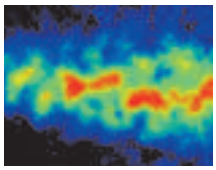
Chapter 12



The Auditory System

Overview

The auditory system is one of the engineering masterpieces of the human body. At the heart of the system is an array of miniature acoustical detectors packed into a space no larger than a pea. These detectors can faithfully transduce vibrations as small as the diameter of an atom, and they can respond a thousand times faster than visual photoreceptors. Such rapid auditory responses to acoustical cues facilitate the initial orientation of the head and body to novel stimuli, especially those that are not initially within the field of view. Although humans are highly visual creatures, much human communication is mediated by the auditory system; indeed, loss of hearing can be more socially debilitating than blindness. From a cultural perspective, the auditory system is essential not only to understanding speech, but also to music, one of the most aesthetically sophisticated forms of human expression. For these and other reasons, audition represents a fascinating and especially important mode of sensation.

Sound

In physical terms, *sound* refers to pressure waves generated by vibrating air molecules (somewhat confusingly, sound is used more casually to refer to an auditory percept). Sound waves are much like the ripples that radiate outward when a rock is thrown in a pool of water. However, instead of occurring across a two-dimensional surface, sound waves propagate in three dimensions, creating spherical shells of alternating compression and rarefaction. Like all wave phenomena, sound waves have four major features: **waveform**, **phase**, **amplitude** (usually expressed in log units known as decibels, abbreviated dB), and **frequency** (expressed in cycles per second or Hertz, abbreviated Hz). For human listeners, the amplitude and frequency of a sound pressure change at the ear roughly correspond to **loudness** and **pitch**, respectively.

The waveform of a sound stimulus is its amplitude plotted against time. It helps to begin by visualizing an acoustical waveform as a sine wave. At the same time, it must be kept in mind that sounds composed of single sine waves (i.e., pure tones) are extremely rare in nature; most sounds in speech, for example, consist of acoustically complex waveforms. Interestingly, such complex waveforms can often be modeled as the sum of sinusoidal waves of varying amplitudes, frequencies, and phases. In engineering applications, an algorithm called the Fourier transform decomposes a complex signal into its sinusoidal components. In the auditory system, as will be apparent later in the chapter, the inner ear acts as a sort of acoustical prism, decomposing complex sounds into a myriad of constituent tones.

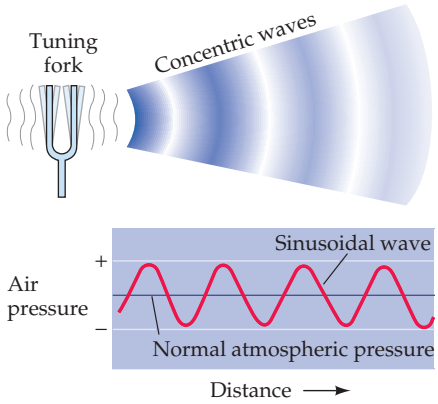


Figure 12.1 Diagram of the periodic condensation and rarefaction of air molecules produced by the vibrating tines of a tuning fork. The molecular disturbance of the air is pictured as if frozen at the instant the constituent molecules responded to the resultant pressure wave. Shown below is a plot of the air pressure versus distance from the fork. Note its sinusoidal quality.

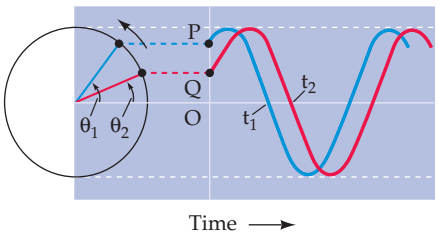


Figure 12.2 A sine wave and its projection as circular motion. The two sinusoids shown are at different phases, such that point P corresponds to phase angle θ_1 and point Q corresponds to phase angle θ_2.

Figure 12.1 diagrams the behavior of air molecules near a tuning fork that vibrates sinusoidally when struck. The vibrating tines of the tuning fork produce local displacements of the surrounding molecules, such that when the tine moves in one direction, there is molecular condensation; when it moves in the other direction, there is rarefaction. These changes in density of the air molecules are equivalent to local changes in air pressure.

Such regular, sinusoidal cycles of compression and rarefaction can be thought of as a form of circular motion, with one complete cycle equivalent to one full revolution (360°). This point can be illustrated with two sinusoids of the same frequency projected onto a circle, a strategy that also makes it easier to understand the concept of phase (Figure 12.2). Imagine that two tuning forks, both of which resonate at the same frequency, are struck at slightly different times. At a given time $t = 0$, one wave is at position P and the other at position Q. By projecting P and Q onto the circle, their respective phase angles, θ_1 and θ_2, are apparent. The sine wave that starts at P reaches a particular point on the circle, say 180°, at time t_1, whereas the wave that starts at Q reaches 180° at time t_2. Thus, phase differences have corresponding time differences, a concept that is important in appreciating how the auditory system locates sounds in space.

The human ear is extraordinarily sensitive to sound pressure. At the threshold of hearing, air molecules are displaced an average of only 10 picometers (10^{-11} m), and the intensity of such a sound is about one-trillionth of a watt per square meter! This means a listener on an otherwise noiseless planet could hear a 1-watt, 3-kHz sound source located over 450 km away (consider that even a very dim light bulb consumes more than 1 watt of power). Even dangerously high sound pressure levels (>100 dB) have power at the eardrum that is only in the milliwatt range (Box A).

The Audible Spectrum

Humans can detect sounds in a frequency range from about 20 Hz to 20 kHz. Human infants can actually hear frequencies slightly higher than 20 kHz, but lose some high-frequency sensitivity as they mature; the upper limit in average adults is closer to 15–17 kHz. Not all mammalian species are sensitive to the same range of frequencies. Most small mammals are sensitive to very high frequencies, but not to low frequencies. For instance, some species of bats are sensitive to tones as high as 200 kHz, but their lower limit is around 20 kHz—the upper limit for young people with normal hearing.

One reason for these differences is that small objects, including the auditory structures of these small mammals, resonate at high frequencies, whereas large objects tend to resonate at low frequencies—which explains why the violin has a higher pitch than the cello. Different animal species tend to emphasize frequency bandwidths in both their vocalizations and their range of hearing. In general, vocalizations by virtue of their periodicity can be distinguished from the noise "barrier" created by environmental sounds, such as wind and rustling leaves. Animals that echolocate, such as bats and dolphins, rely on very high-frequency vocal sounds to maximally resolve spatial features of the target, while animals intent on avoiding predation have auditory systems "tuned" to the low frequency vibrations that approaching predators transmit through the substrate. These behavioral differences are mirrored by a wealth of anatomical and functional specializations throughout the auditory system.

Box A

Four Causes of Acquired Hearing Loss

Acquired hearing loss is an increasingly common sensory deficit that can often lead to impaired oral communication and social isolation. Four major causes of acquired hearing loss are acoustical trauma, infection of the inner ear, ototoxic drugs, and presbyacusis (literally, the hearing of the old).

The exquisite sensitivity of the auditory periphery, combined with the direct mechanical linkage between the acoustical stimulus and the receptor cells, make the ear especially susceptible to acute or chronic acoustical trauma. Extremely loud, percussive sounds, such as those generated by explosives or gunfire, can rupture the eardrum and so severely distort the inner ear that the organ of Corti is torn. The resultant loss of hearing is abrupt and often quite severe. Less well appreciated is the fact that repeated exposure to less dramatic but nonetheless loud sounds, including those produced by industrial or household machinery or by amplified musical instruments, can also damage the inner ear. Although these sounds leave the eardrum intact, specific damage is done to the hair bundle itself; the stereocilia of cochlear hair cells of animals exposed to loud sounds shear off at their pivot points with the hair cell body, or fuse together in a platelike fashion that impedes movement. In humans, the mechanical resonance of the ear to stimulus frequencies centered about 3 kHz means that exposure to loud, broadband noises (such as those generated by jet engines) results in especially pronounced deficits near this resonant frequency.

Ototoxic drugs include aminoglycoside antibiotics (such as gentamycin and kanamycin), which directly affect hair cells, and ethacrynic acid, which poisons the potassium-extruding cells of the stria vascularis that generate the endocochlear potential. In the absence of these ion pumping cells, the endocochlear potential, which supplies the energy to drive the transduction process, is lost. Although still a matter of some debate, the relatively nonselective transduction channel apparently affords a means of entry for aminoglycoside antibiotics, which then poison hair cells by disrupting phosphoinositide metabolism. In particular, outer hair cells and those inner hair cells that transduce high-frequency stimuli are more affected, simply because of their greater energy requirements.

Finally, *presbyacusis*, the hearing loss associated with aging, may in part stem from atherosclerotic damage to the especially fine microvasculature of the inner ear, as well as from genetic predispositions to hair cell damage. Recent advances in understanding the genetic transmission of acquired hearing loss in both humans and mice point to mutations in myosin isoforms unique to hair cells as a likely culprit.

References

Holt, J. R. and D. P. Corey (1999) Ion channel defects in hereditary hearing loss. Neuron 22: 217–219.

Keats, B. J. and D. P. Corey (1999) The usher syndromes. Amer. J. Med. Gen. 89: 158–166.

Priuska, E. M. and J. Schact (1997) Mechanism and prevention of aminoglycoside ototoxicity: Outer hair cells as targets and tools. Ear, Nose, Throat J. 76: 164–171.

A Synopsis of Auditory Function

The auditory system transforms sound waves into distinct patterns of neural activity, which are then integrated with information from other sensory systems to guide behavior, including orienting movements to acoustical stimuli and intraspecies communication. The first stage of this transformation occurs at the external and middle ears, which collect sound waves and amplify their pressure, so that the sound energy in the air can be successfully transmitted to the fluid-filled cochlea of the inner ear. In the inner ear, a series of biomechanical processes occur that break up the signal into simpler, sinusoidal components, with the result that the frequency, amplitude, and phase of the original signal are all faithfully transduced by the sensory **hair cells** and encoded by the electrical activity of the **auditory nerve fibers**. One product of this process of acoustical decomposition is the systematic representation of sound frequency along the length of the cochlea, referred to as **tonotopy**, which is an important organizational feature preserved

Box B
Music

Even though we all recognize it when we hear it, the concept of music is vague. The *Oxford English Dictionary* defines it as "The art or science of combining vocal or instrumental sounds with a view toward beauty or coherence of form and expression of emotion." In terms of the present chapter, music chiefly concerns the aspect of human audition that is experienced as tones. The stimuli that give rise to tonal percepts are periodic, meaning that they repeat systematically over time, as in the sine wave in Figure 12.1. Periodic stimuli, which do not occur naturally as sine waves but rather as complex repetitions involving a number of different frequencies, give rise to a sense of harmony when sounded together in appropriate combinations, and a sense of melody when they occur sequentially.

Although we usually take the way tone-evoking stimuli are heard for granted, this aspect of audition presents some profoundly puzzling qualities. The most obvious of these is that humans perceive periodic stimuli whose fundamental frequencies have a 2:1 ratio as highly similar, and, for the most part, musically interchangeable. Thus in Western musical terminology, any two tones related by an interval of one or more octaves are given the same name (i.e., A, B, C…G), and are distinguished only by a qualifier that denotes relative ordinal position (e.g., C_1, C_2, C_3, etc.). As a result, music is framed in repeating intervals (called octaves) defined by these more or less interchangeable tones. A key question, then, is why periodic sound stimuli whose fundamentals have a 2:1 ratio are perceived as similar when there is no obvious physical or physiological basis for this phenomenon.

A second puzzling feature is that most if not all musical traditions subdivide octaves into a relatively small set of intervals for composition and performance, each interval being defined by its relationship to the lowest tone of the set. Such sets are called musical scales. The scales predominantly employed in all cultures over the centuries have used some (or occasionally all) of the 12 tonal intervals that in Western musical terminology are referred to as the chromatic scale (see figure). Moreover, some intervals of the chromatic scale, such as the fifth, the fourth, the major third, and the major sixth, are more often used in composition and performance than others. These form the majority of the intervals employed in the pentatonic and diatonic major scales, the two most frequently used scales in music world-wide. Again,

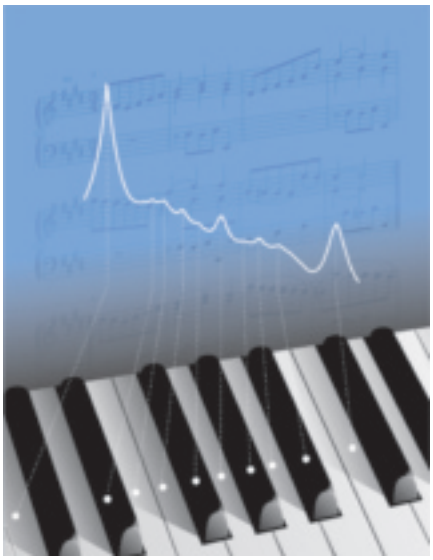


Illustration of 10 of the 12 tones in the chromatic scale, related to a piano keyboard. The function above the keyboard indicates that these tones correspond statistically to peaks of power in normalized human speech. (After Schwartz et al., 2003.)

throughout the central auditory pathways. The earliest stage of central processing occurs at the cochlear nucleus, where the peripheral auditory information diverges into a number of parallel central pathways. Accordingly, the output of the cochlear nucleus has several targets. One of these is the superior olivary complex, the first place that information from the two ears interacts and the site of the initial processing of the cues that allow listeners to localize sound in space. The cochlear nucleus also projects to the inferior colliculus of the midbrain, a major integrative center and the first place where auditory information can interact with the motor system. The inferior colliculus is an obligatory relay for information traveling to the thalamus and cortex, where additional integrative aspects (such as harmonic and temporal combinations) of sound especially germane to speech and music are processed (Box B). The large number of stations between the auditory periphery and the cortex far exceeds those in other sensory systems, providing a hint that the perception of communication and environmental sounds

there is no principled explanation of these preferences among all the possible intervals within the octave.

Perhaps the most fundamental question in music—and arguably the common denominator of all musical tonality—is why certain combinations of tones are perceived as relatively consonant or 'harmonious' and others relatively dissonant or 'inharmonious'. These perceived differences among the possible combinations of tones making up the chromatic scale are the basis for polytonal music, in which the perception of relative harmoniousness guides the composition of chords and melodic lines. The more compatible of these combinations are typically used to convey 'resolution' at the end of a musical phrase or piece, whereas less compatible combinations are used to indicate a transition, a lack of resolution, or to introduce a sense of tension in a chord or melodic sequence. Like octaves and scales, the reason for this phenomenology remains a mystery.

The classical approaches to rationalizing octaves, scales and consonance have been based on the fact that the musical intervals corresponding to octaves, fifths, and fourths (in modern musical terminology) are produced by physical sources whose relative proportions (e.g., the relative lengths of two plucked strings or their fundamental frequencies) have ratios of 2:1, 3:2, or 4:3, respectively (these relationships were first described by Pythagoras). This coincidence of numerical simplicity and perceptual effect has been so impressive over the centuries that attempts to rationalize phenomena such as consonance and scale structure in terms of mathematical relationships have tended to dominate the thinking about these issues. This conceptual framework, however, fails to account for many of the perceptual observations that have been made over the last century.

Another way to consider the problem is in terms of the biological rationale for evolving a sense of tonality in the first place. A pertinent fact in this regard is that only a small minority of naturally occurring sound stimuli are periodic. Since the auditory system evolved in the world of natural sounds, this point is presumably critical for thinking about the biological purposes of tonality and music. Indeed, the majority of periodic sounds that humans would have been exposed to during evolution are those made by the human vocal tract in the process of communication, initially prelinguistic but more recently speech sounds (see Chapter 26). Thus developing a sense of tonality would enable listeners to respond not only the distinctions among the different speech sounds that are important for understanding spoken language, but to information about the probable sex, age, and emotional state of the speaker. It may thus be that music reflects the advantage of facilitating a listener's ability to glean the linguistic intent and biological state of fellow humans through vocal utterances.

References

Burns, E. M. (1999) Intervals, scales, and tuning. In *The Psychology of Music*, D. Deutsch (ed.). New York: Academic Press, pp. 215–264.

Carterette, E. C. and R. A. Kendall (1999) Comparative music perception and cognition. In *The Psychology of Music*, D. Deutsch (ed.). New York: Academic Press.

Lewicki, M. S. (2002) Efficient coding of natural sounds. Nature Neurosci. 5: 356–363.

Pierce, J. R. (1983, 1992) *The Science of Musical Sound*. New York: W.H. Freeman and Co., Chapters 4–6.

Plomp, R. and W. J. Levelt (1965) Tonal consonance and critical bandwidth. *J. Acoust. Soc. Amer.* 28: 548–560.

Rasch, R. and R. Plomp (1999) The perception of musical tones. In *The Psychology of Music*, D. Deutsch (ed.). New York: Academic Press, pp. 89–112.

Schwartz, D. A., C. Q. Howe and D. Purves (2003) The statistical structure of human speech sounds predicts musical universals. J. Neurosci. 23: 7160–7168.

Terhardt, E. (1974) Pitch, consonance, and harmony. *J. Acoust. Soc. Amer.* 55: 1061–1069.

is an especially intensive neural process. Furthermore, both the peripheral and central auditory system are "tuned" to conspecific communication vocalizations, pointing to the interdependent evolution of neural systems used for generating and perceiving these signals.

The External Ear

The external ear, which consists of the **pinna**, **concha**, and **auditory meatus**, gathers sound energy and focuses it on the eardrum, or **tympanic membrane** (Figure 12.3). One consequence of the configuration of the human auditory meatus is that it selectively boosts the sound pressure 30- to 100-fold for frequencies around 3 kHz via passive resonance effects. This amplification makes humans especially sensitive to frequencies in the range of 2–5 kHz—and also explains why they are particularly prone to hearing loss near this frequency following exposure to loud broadband noises, such as those

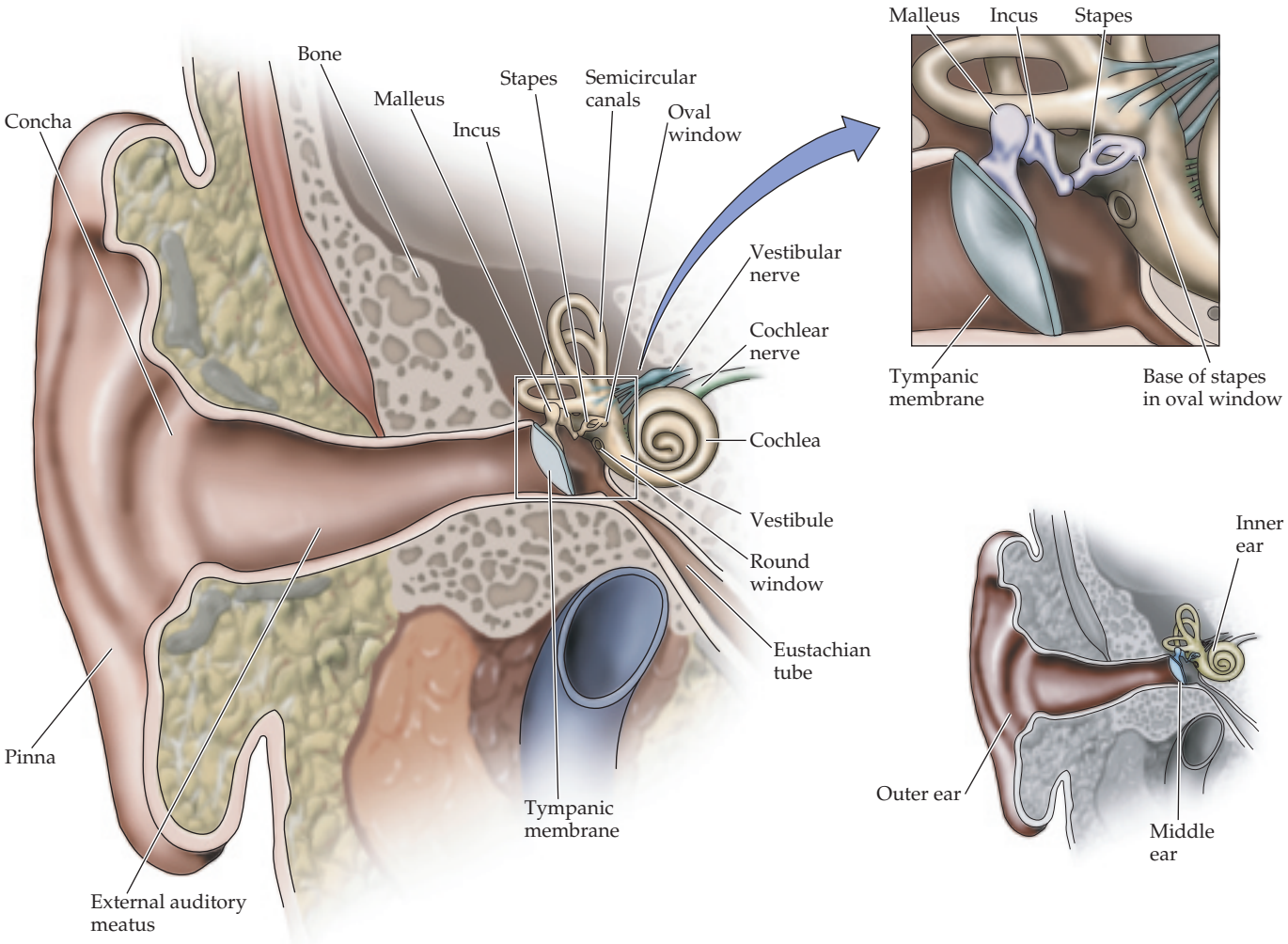


Figure 12.3 The human ear. Note the large surface area of the tympanic membrane (eardrum) relative to the oval window, a feature that facilitates transmission of airborne sounds to the fluid-filled cochlea.

generated by heavy machinery or high explosives (see Box A). The sensitivity to this frequency range in the human auditory system appears to be directly related to speech perception: although human speech is a broadband signal, the energy of the plosive consonants (e.g., *ba* and *pa*) that distinguish different phonemes (the elementary human speech sounds) is concentrated around 3 kHz (see Box A in Chapter 26). Therefore, selective hearing loss in the 2–5 kHz range disproportionately degrades speech recognition. Most vocal communication occurs in the low-kHz range to overcome environmental noise; as already noted, generation of higher frequencies is difficult for animals the size of humans.

A second important function of the pinna and concha is to selectively filter different sound frequencies in order to provide cues about the elevation of the sound source. The vertically asymmetrical convolutions of the pinna are shaped so that the external ear transmits more high-frequency components from an elevated source than from the same source at ear level. This effect can be demonstrated by recording sounds from different elevations after they have passed through an "artificial" external ear; when the recorded sounds are played back via earphones, so that the whole series is at the same elevation relative to the listener, the recordings from higher elevations are perceived as coming from positions higher in space than the recordings from lower elevations.

The Middle Ear

Sounds impinging on the external ear are airborne; however, the environment within the inner ear, where the sound-induced vibrations are converted to neural impulses, is aqueous. The major function of the middle ear is to match relatively low-impedance airborne sounds to the higher-impedance fluid of the inner ear. The term "impedance" in this context describes a medium's resistance to movement. Normally, when sound waves travel from a low-impedance medium like air to a much higher-impedance medium like water, almost all (more than 99.9%) of the acoustical energy is reflected. The middle ear (see Figure 12.3) overcomes this problem and ensures transmission of the sound energy across the air–fluid boundary by boosting the pressure measured at the tympanic membrane almost 200-fold by the time it reaches the inner ear.

Two mechanical processes occur within the middle ear to achieve this large pressure gain. The first and major boost is achieved by focusing the force impinging on the relatively large-diameter tympanic membrane on to the much smaller-diameter **oval window**, the site where the bones of the middle ear contact the inner ear. A second and related process relies on the mechanical advantage gained by the lever action of the three small interconnected middle ear bones, or **ossicles** (i.e., the malleus, incus, and stapes; see Figure 12.3), which connect the tympanic membrane to the oval window. **Conductive hearing losses**, which involve damage to the external or middle ear, lower the efficiency at which sound energy is transferred to the inner ear and can be partially overcome by artificially boosting sound pressure levels with an external hearing aid (Box C). In normal hearing, the efficiency of sound transmission to the inner ear also is regulated by two small muscles in the middle ear, the tensor tympani, innervated by cranial nerve V, and the stapedius, innervated by cranial nerve VII (see Appendix A). Flexion of these muscles, which is triggered automatically by loud noises or during self-generated vocalization, stiffens the ossicles and reduces the amount of sound energy transmitted to the cochlea, serving to protect the inner ear. Conversely, conditions that lead to flaccid paralysis of either of these muscles, such as Bell's palsy (nerve VII), can trigger a painful sensitivity to moderate or even low intensity sounds known as **hyperacusis**.

Bony and soft tissues, including those surrounding the inner ear, have impedances close to that of water. Therefore, even without an intact tympanic membrane or middle ear ossicles, acoustical vibrations can still be transferred directly through the bones and tissues of the head to the inner ear. In the clinic, bone conduction can be exploited using a simple test involving a tuning fork to determine whether hearing loss is due to conductive problems or is due to damage to the hair cells of the inner ear or to the auditory nerve itself (**sensorineural hearing loss**; see Boxes A and C)

The Inner Ear

The **cochlea** of the inner ear is arguably the most critical structure in the auditory pathway, for it is there that the energy from sonically generated pressure waves is transformed into neural impulses. The cochlea not only amplifies sound waves and converts them into neural signals, but it also acts as a mechanical frequency analyzer, decomposing complex acoustical waveforms into simpler elements. Many features of auditory perception derive from aspects of the physical properties of the cochlea; hence, it is important to consider this structure in some detail.

Box C

Sensorineural Hearing Loss and Cochlear Implants

The same features that make the auditory periphery exquisitely sensitive to detecting airborne sounds also make it highly vulnerable to damage. By far the most common forms of hearing loss involve the peripheral auditory system, namely to those structures that transmit and transduce sounds into neural impulses. Monaural hearing deficits are the defining symptom of a peripheral hearing loss, because unilateral damage at or above the auditory brainstem results in a binaural deficit (due to the extensive bilateral organization of the central auditory system). Peripheral hearing insults can be further divided into conductive hearing losses, which involve damage to the outer or middle ear, and sensorineural hearing losses, which stem from damage to the inner ear, most typically the cochlear hair cells or the VIIIth nerve itself. Although both forms of peripheral hearing loss manifest themselves as a raised threshold for hearing on the affected side, their diagnoses and treatments differ.

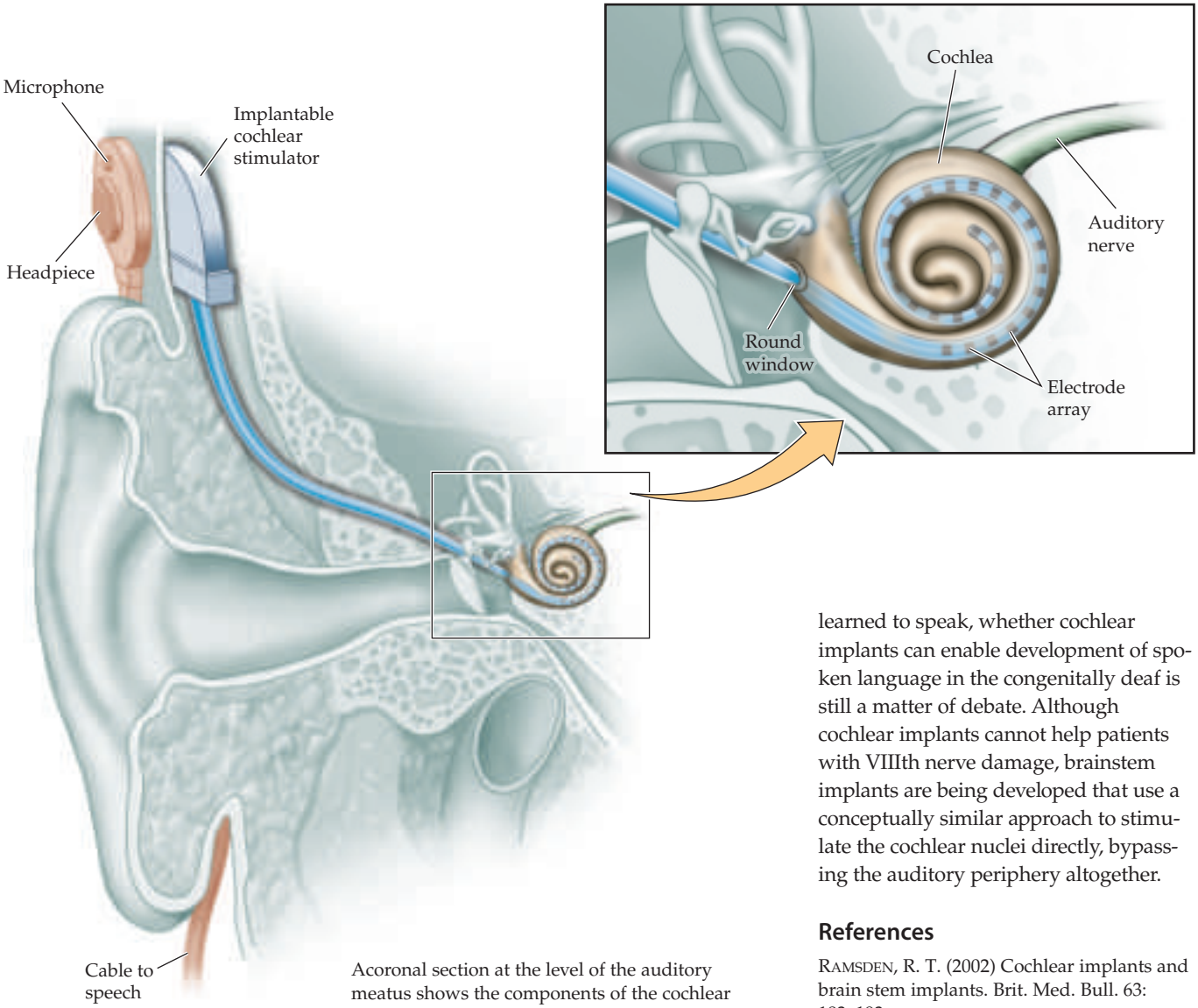
Conductive hearing loss can be due to occlusion of the ear canal by wax or foreign objects, rupture of the tympanic membrane itself, or arthritic ossification of the middle ear bones. In contrast, sensorineural hearing loss usually is due to congenital or environmental insults that lead to hair cell death (see Box A) or damage to the eighth nerve. As hair cells are relatively few in number and do not regenerate in humans, their depletion leads to a diminished ability to detect sounds. The Weber test, a simple test involving a tuning fork, can be used to distinguish between these two forms of hearing loss. If a resonating tuning fork (~256 Hz) is placed on the vertex, a patient with conductive hearing loss will report that the sound is louder in the affected ear. In the "plugged" state, sounds propagating through the skull do not dissipate so freely back out through the auditory meatus, and thus a greater amount of sound energy is transmitted to the cochlea on the blocked side. In contrast, a patient with a monaural sensorineural hearing loss will report that a Weber test sounds louder on the intact side, because even though the inner ear may vibrate equally on the two sides, the damaged side cannot transduce this vibration into a neural signal.

Treatment also differs for these two types of deafness. An external hearing aid is used to boost sounds to compensate for the reduced efficiency of the conductive apparatus in conductive hearing losses. These miniature devices are inserted in the ear canal, and contain a microphone and speaker, as well as an amplifier. One limitation of hearing aids is that they often provide rather flat amplification curves, which can interfere with listening in noisy environments; moreover, they do not achieve a high degree of directionality. The use of digital signal processing strategies partly overcomes these problems, and hearing aids obviously provide significant benefits to many people.

The treatment of sensorineural hearing loss is more complicated and invasive; conventional hearing aids are useless, because no amount of mechanical amplification can compensate for the inability to generate or convey a neural impulse from the cochlea. However, if the VIIIth nerve is intact, cochlear implants can be used to partially restore hearing. The cochlear implant consists of a peripherally mounted microphone and digital signal processor that transforms a sound into its spectral components, and additional electronics that use this information to activate different combinations of contacts on a threadlike multi-site stimulating electrode array. The electrode is inserted into the cochlea through the round window (see figure) and positioned along the length of the tonotopically organized basilar membrane and VIIIth nerve endings. This placement enables electrical stimulation of the nerve in a manner that mimics some aspects of the spectral decomposition naturally performed by the cochlea.

Cochlear implants can be remarkably effective in restoring hearing to people with hair cell damage, permitting them to engage in spoken communication. Despite such success in treating those who have lost their hearing *after* having

The cochlea (from the Latin for "snail") is a small (about 10 mm wide) coiled structure, which, were it uncoiled, would form a tube about 35 mm long (Figures 12.4 and 12.5). Both the oval window and, the **round window,** another region where the bone is absent surrounding the cochlea, are at the basal end of this tube. The cochlea is bisected from its basal almost to its apical end by the cochlear partition, which is a flexible structure that supports the **basilar membrane** and the **tectorial membrane**. There are fluid-filled chambers on each side of the cochlear partition, named the **scala vestibuli** and the **scala tympani**; a distinct channel, the **scala media**, runs within the



Acoronal section at the level of the auditory meatus shows the components of the cochlear implant, including the filamentous stimulatory electrode inserted into the cochlea through the round window.

learned to speak, whether cochlear implants can enable development of spoken language in the congenitally deaf is still a matter of debate. Although cochlear implants cannot help patients with VIIIth nerve damage, brainstem implants are being developed that use a conceptually similar approach to stimulate the cochlear nuclei directly, bypassing the auditory periphery altogether.

References

RAMSDEN, R. T. (2002) Cochlear implants and brain stem implants. Brit. Med. Bull. 63: 183–193.

RAUSCHECKER, J. P. AND R. V. SHANNON (2002) Sending sound to the brain. Science. 295: 1025–1029.

cochlear partition. The cochlear partition does not extend all the way to the apical end of the cochlea; instead there is an opening, known as the **helicotrema**, that joins the scala vestibuli to the scala tympani, allowing their fluid, known as **perilymph**, to mix. One consequence of this structural arrangement is that inward movement of the oval window displaces the fluid of the inner ear, causing the round window to bulge out slightly and deforming the cochlear partition.

The manner in which the basilar membrane vibrates in response to sound is the key to understanding cochlear function. Measurements of the vibra-

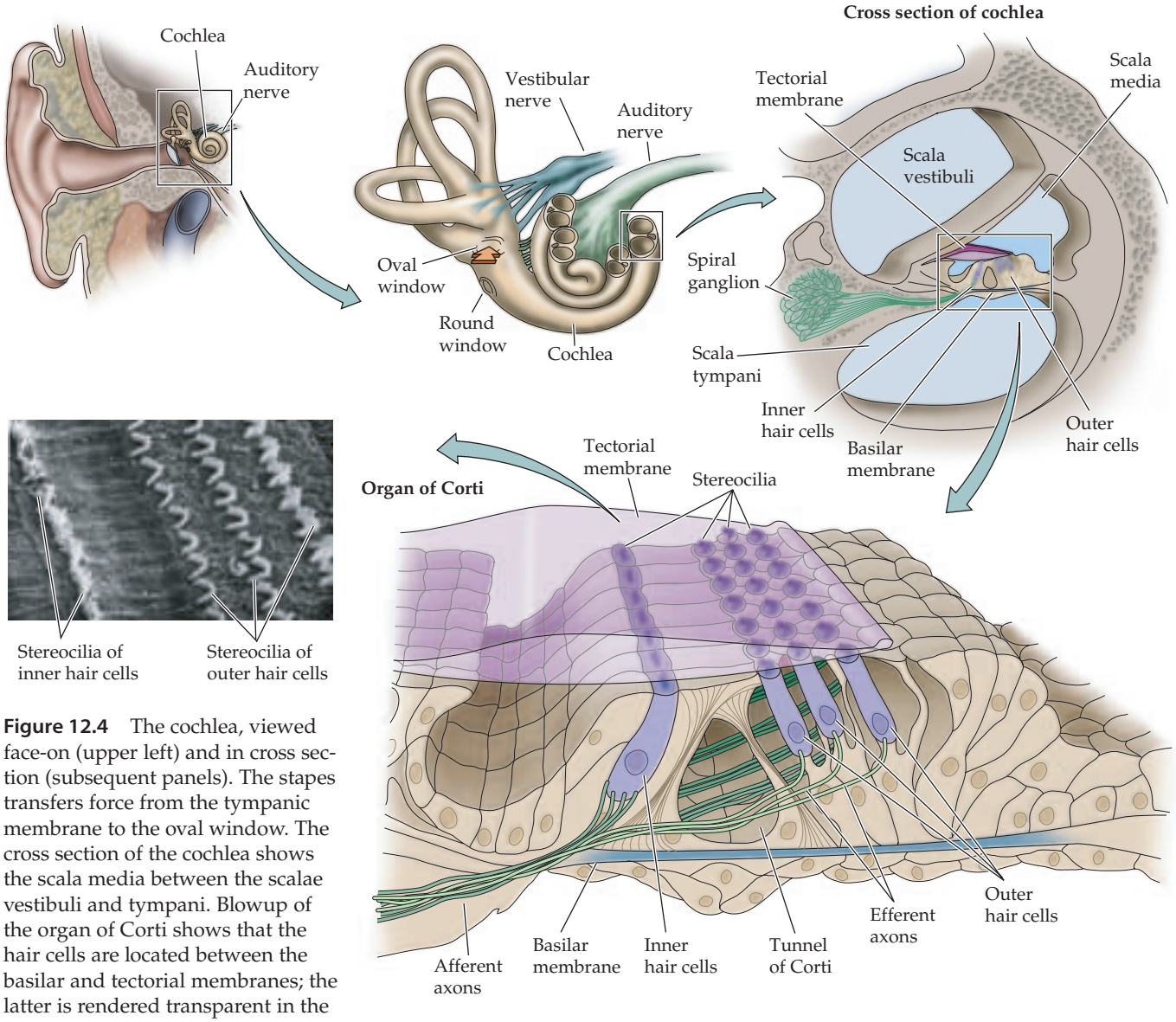


Figure 12.4 The cochlea, viewed face-on (upper left) and in cross section (subsequent panels). The stapes transfers force from the tympanic membrane to the oval window. The cross section of the cochlea shows the scala media between the scalae vestibuli and tympani. Blowup of the organ of Corti shows that the hair cells are located between the basilar and tectorial membranes; the latter is rendered transparent in the line drawing and removed in the scanning electron micrograph. The hair cells are named for their tufts of stereocilia; inner hair cells receive afferent inputs from cranial nerve VIII, whereas outer hair cells receive mostly efferent input. (Micrograph from Kessel and Kardon, 1979.)

tion of different parts of the basilar membrane, as well as the discharge rates of individual auditory nerve fibers that terminate along its length, show that both these features are highly tuned; that is, they respond most intensely to a sound of a specific frequency. Frequency tuning within the inner ear is attributable in part to the geometry of the basilar membrane, which is wider and more flexible at the apical end and narrower and stiffer at the basal end. One feature of such a system is that regardless of where energy is supplied to it, movement always begins at the stiff end (i.e., the base), and then propagates to the more flexible end (i.e., the apex). Georg von Békésy, working at Harvard University, showed that a membrane that varies systematically in its width and flexibility vibrates maximally at different positions as a function of the stimulus frequency (Figure 12.5). Using tubular models and human cochleas taken from cadavers, he found that an acoustical stimulus initiates a **traveling wave** of the same frequency in the cochlea, which propagates from the base toward the apex of the basilar membrane, growing in

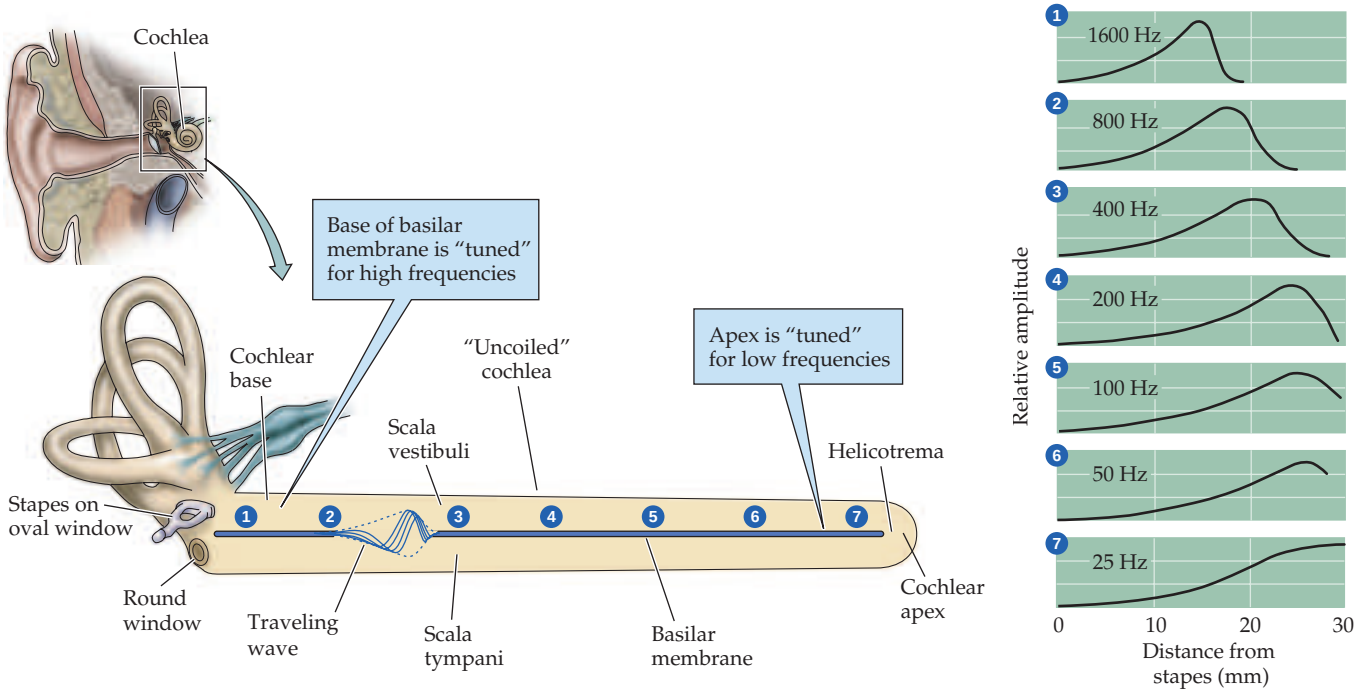


Figure 12.5 Traveling waves along the cochlea. A traveling wave is shown at a given instant along the cochlea, which has been uncoiled for clarity. The graphs on the right profile the amplitude of the traveling wave along the basilar membrane for different frequencies and show that the position (i.e., 1–7) where the traveling wave reaches its maximum amplitude varies directly with the frequency of stimulation. (Drawing after Dallos, 1992; graphs after von Békésy, 1960.)

amplitude and slowing in velocity until a point of maximum displacement is reached. This point of maximal displacement is determined by the sound frequency. The points responding to high frequencies are at the base of the basilar membrane where it is stiffer, and the points responding to low frequencies are at the apex, giving rise to a topographical mapping of frequency (that is, to **tonotopy**). An important feature is that complex sounds cause a pattern of vibration equivalent to the superposition of the vibrations generated by the individual tones making up that complex sound, thus accounting for the decompositional aspects of cochlear function mentioned earlier. This process of spectral decomposition appears to be an important strategy for detecting the various harmonic combinations that distinguish different natural sounds. Indeed, tonotopy is conserved throughout much of the auditory system, including the auditory cortex, suggesting that it is important to speech processing.

Von Békésy's model of cochlear mechanics was a passive one, resting on the premise that the basilar membrane acts like a series of linked resonators, much as a concatenated set of tuning forks. Each point on the basilar membrane was postulated to have a characteristic frequency at which it vibrated most efficiently; because it was physically linked to adjacent areas of the membrane, each point also vibrated (if somewhat less readily) at other frequencies, thus permitting propagation of the traveling wave. It is now clear, however, that the tuning of the auditory periphery, whether measured at the basilar membrane or recorded as the electrical activity of auditory nerve fibers, is too sharp to be explained by passive mechanics alone. At very low sound intensities, the basilar membrane vibrates one hundred-fold more than would be predicted by linear extrapolation from the motion measured at high intensities. Therefore, the ear's sensitivity arises from an active biomechanical process, as well as from its passive resonant properties (Box D). The outer hair cells, which together with the inner hair cells comprise the

Box D

The Sweet Sound of Distortion

As early as the first half of the eighteenth century, musical composers such as G. Tartini and W. A. Sorge discovered that upon playing pairs of tones, other tones not present in the original stimulus are also heard. These combination tones, *fc*, are mathematically related to the played tones f_1 and f_2 ($f_2 > f_1$) by the formula

$$fc = mf_1 \pm nf_2$$

where m and n are positive integers. Combination tones have been used for a variety of compositional effects, as they can strengthen the harmonic texture of a chord. Furthermore, organ builders sometimes use the difference tone ($f_2 - f_1$) created by two smaller organ pipes to produce the extremely low tones that would otherwise require building one especially large pipe.

Modern experiments suggest that this distortion product is due at least in part to the nonlinear properties of the inner ear. M. Ruggero and his colleagues placed small glass beads (10–30 mm in diameter) on the basilar membrane of an anesthetized animal and then determined the velocity of the basilar membrane in response to different combinations of tones by measuring the Doppler shift of laser light reflected from the beads. When two tones were played into the ear, the basilar membrane vibrated not only at those two frequencies, but also at other frequencies predicted by the above formula.

Related experiments on hair cells studied in vitro suggest that these nonlinearities result from the properties of the mechanical linkage of the transduction apparatus. By moving the hair bundle sinusoidally with a metal-coated glass fiber, A. J. Hudspeth and his coworkers found that the hair bundle exerts a force at the same frequency. However, when two sinusoids were applied simultaneously, the forces exerted by the hair bundle occurred not only at the primary frequencies, but at several combination frequencies as well. These distortion products are due to the transduction apparatus, since blocking the transduction channels causes the forces exerted at the combination frequencies to disappear, even though the forces at the primary frequencies remain unaffected. It seems that the tip links add a certain extra springiness to the hair bundle in the small range of motions over which the transduction channels are changing between closed and open states. If nonlinear distortions of basilar membrane vibrations arise from the properties of the hair bundle, then it is likely that hair cells can indeed influence basilar membrane motion, thereby accounting for the cochlea's extreme sensitivity. When we hear difference tones, we may be paying the price in distortion for an exquisitely fast and sensitive transduction mechanism.

References

Jaramillo, F., V. S. Markin and A. J. Hudspeth (1993) Auditory illusions and the single hair cell. Nature 364: 527–529.

Planchart, A. E. (1960) A study of the theories of Giuseppe Tartini. J. Music Theory 4: 32–61.

Robles, L., M. A. Ruggero and N. C. Rich (1991) Two-tone distortion in the basilar membrane of the cochlea. Nature 439: 413–414.

sensory cells of the inner ear, are the most likely candidates for driving this active process.

The motion of the traveling wave initiates sensory transduction by displacing the hair cells that sit atop the basilar membrane. Because these structures are anchored at different positions, the vertical component of the traveling wave is translated into a shearing motion between the basilar membrane and the overlying tectorial membrane (Figure 12.6). This motion bends the tiny processes, called **stereocilia**, that protrude from the apical ends of the hair cells, leading to voltage changes across the hair cell membrane. How the bending of stereocilia leads to receptor potentials in hair cells is considered in the following section.

Hair Cells and the Mechanoelectrical Transduction of Sound Waves

The hair cell is an evolutionary triumph that solves the problem of transforming vibrational energy into an electrical signal. The scale at which the

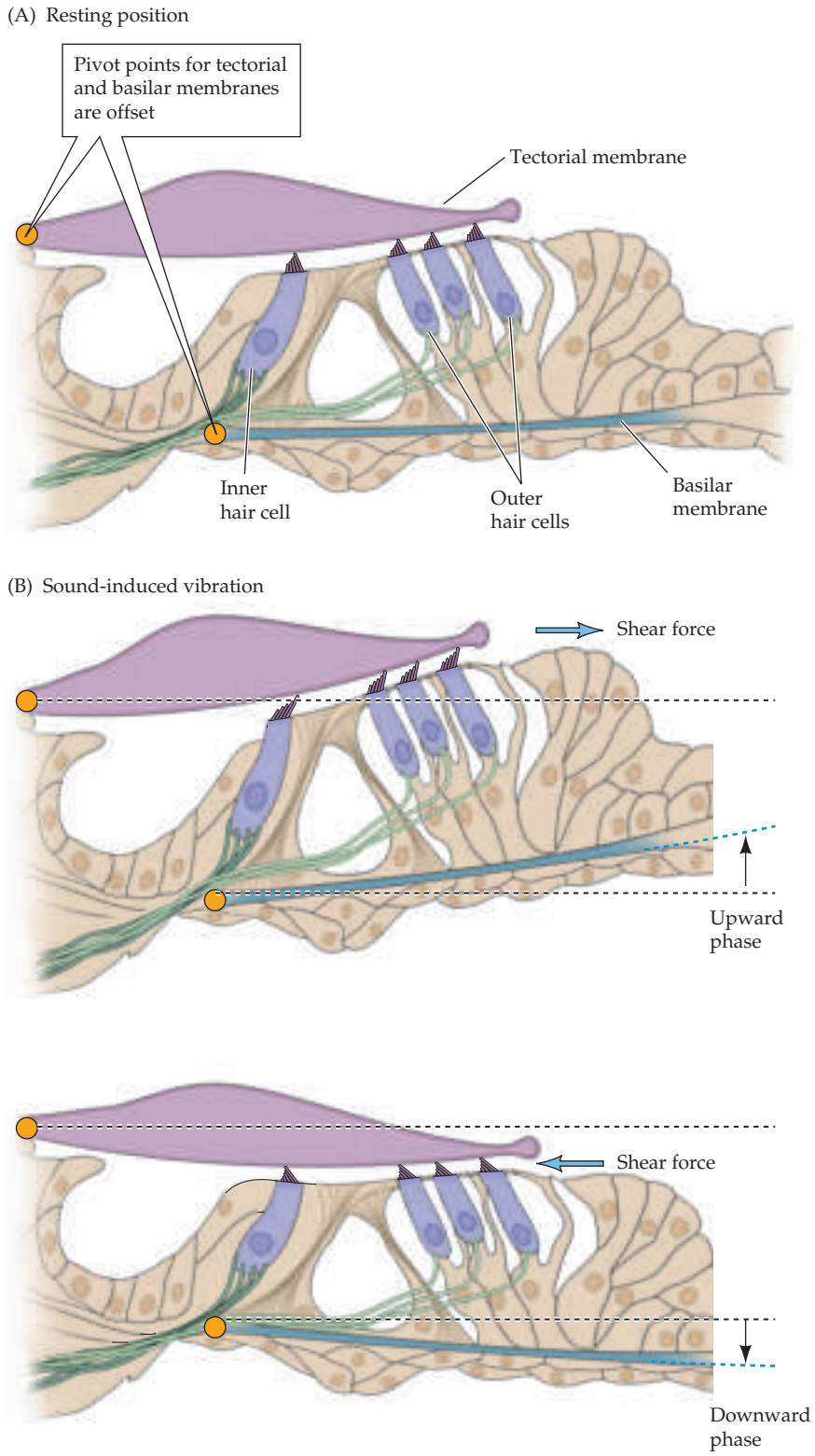


Figure 12.6 Movement of the basilar membrane creates a shearing force that bends the stereocilia of the hair cells. The pivot point of the basilar membrane is offset from the pivot point of the tectorial membrane, so that when the basilar membrane is displaced, the tectorial membrane moves across the tops of the hair cells, bending the stereocilia.

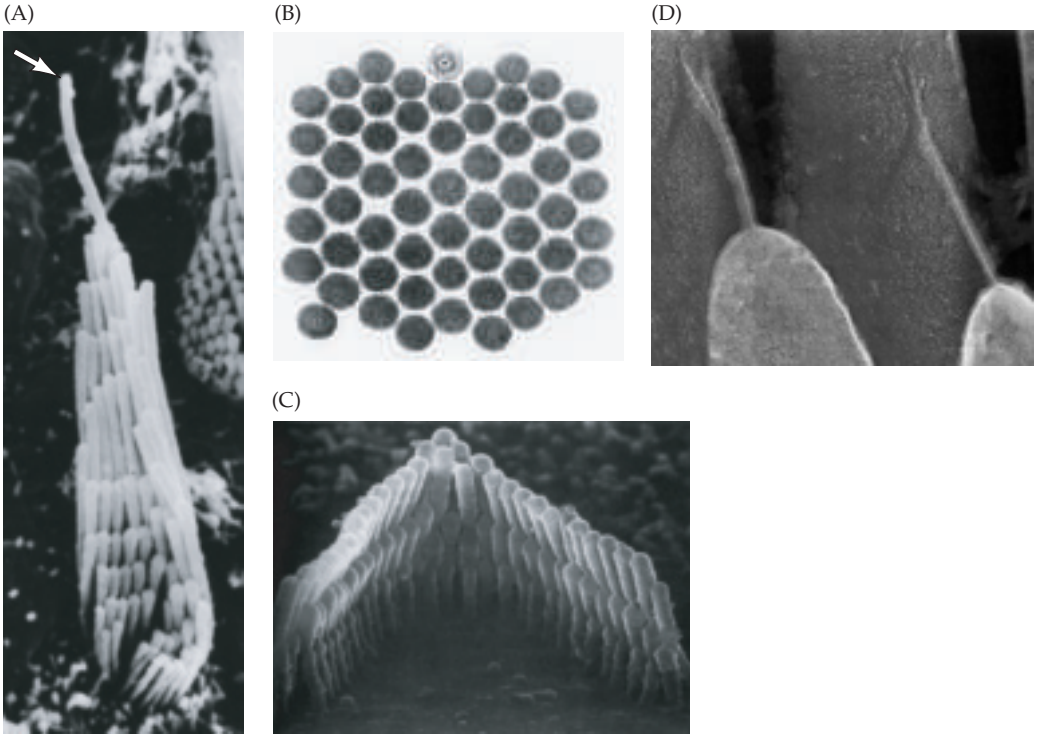


Figure 12.7 The structure and function of the hair bundle in vestibular and cochlear hair cells. The vestibular hair bundles shown here resemble those of cochlear hair cells, except for the presence of the kinocilium, which disappears in the mammalian cochlea shortly after birth. (A) The hair bundle of a guinea pig vestibular hair cell. This view shows the increasing height leading to the kinocilium (arrow). (B) Cross section through the vestibular hair bundle shows the 9 + 2 array of microtubules in the kinocilium (top), which contrasts with the simpler actin filament structure of the stereocilia. (C) Scanning electron micrograph of a guinea pig cochlear outer hair cell bundle viewed along the plane of mirror symmetry. Note the graded lengths of the stereocilia, and the absence of a kinocilium. (D) Tip links that connect adjacent stereocilia are believed to be the mechanical linkage that opens and closes the transduction channel. (A from Lindeman, 1973; B from Hudspeth, 1983; C from Pickles, 1988; D from Fain, 2003.)

hair cell operates is truly amazing: At the limits of human hearing, hair cells can faithfully detect movements of atomic dimensions and respond in the tens of microseconds! Furthermore, hair cells can adapt rapidly to constant stimuli, thus allowing the listener to extract signals from a noisy background.

The hair cell is a flask-shaped epithelial cell named for the bundle of hairlike processes that protrude from its apical end into the scala media. Each hair bundle contains anywhere from 30 to a few hundred hexagonally arranged stereocilia, with one taller **kinocilium** (Figure 12.7A). Despite their names, only the kinocilium is a true ciliary structure, with the characteristic two central tubules surrounded by nine doublet tubules that define cilia (Figure 12.7B). The function of the kinocilium is unclear, and in the cochlea of humans and other mammals it actually disappears shortly after birth (Figure 12.7C). The stereocilia are simpler, containing only an actin cytoskeleton. Each stereocilium tapers where it inserts into the apical membrane, forming a hinge about which each stereocilium pivots (Figure 12.7D). The stereocilia are graded in height and are arranged in a bilaterally symmetric fashion (in vestibular hair cells, this plane runs through the kinocilium). Displacement of the hair bundle parallel to this plane toward the tallest stereocilia depolarizes the hair cell, while movements parallel to this plane toward the shortest stereocilia cause hyperpolarization. In contrast, displacements perpendicular to the plane of symmetry do not alter the hair cell's membrane potential. The hair bundle movements at the threshold of hearing are approximately 0.3 nm, about the diameter of an atom of gold. Hair cells can convert the displacement of the stereociliary bundle into an electrical potential in as little as 10 microseconds; as described below, such speed is required for the accurate localization of the source of the sound. The need for microsecond resolution places certain constraints on the transduction mechanism, ruling out the rela-

tively slow second messenger pathways used in visual and olfactory transduction (see Chapters 7, 10, and 14); a direct, mechanically gated transduction channel is needed to operate this quickly. Evidently the filamentous structures that connect the tips of adjacent stereocilia, known as **tip links**, directly open cation-selective transduction channels when stretched, allowing K^+ ions to flow into the cell (see Figure 12.7D). As the linked stereocilia pivot from side to side, the tension on the tip link varies, modulating the ionic flow and resulting in a graded receptor potential that follows the movements of the stereocilia (Figures 12.8 and 12.9). The tip link model also explains why only deflections along the axis of the hair bundle activate transduction channels, since tip links join adjacent stereocilia along the axis directed toward the tallest stereocilia (see also Box B in Chapter 13). The exquisite mechanical sensitivity of the stereocilia also presents substantial risks: high intensity sounds can shear off the hair bundle, resulting in profound hearing deficits. Because human stereocilia, unlike those in fishes and birds, do not regenerate such damage is irreversible. The small number of hair cells (a total of about 30,000 in a human, or 15,000 per ear) further compounds the sensitivity of the inner

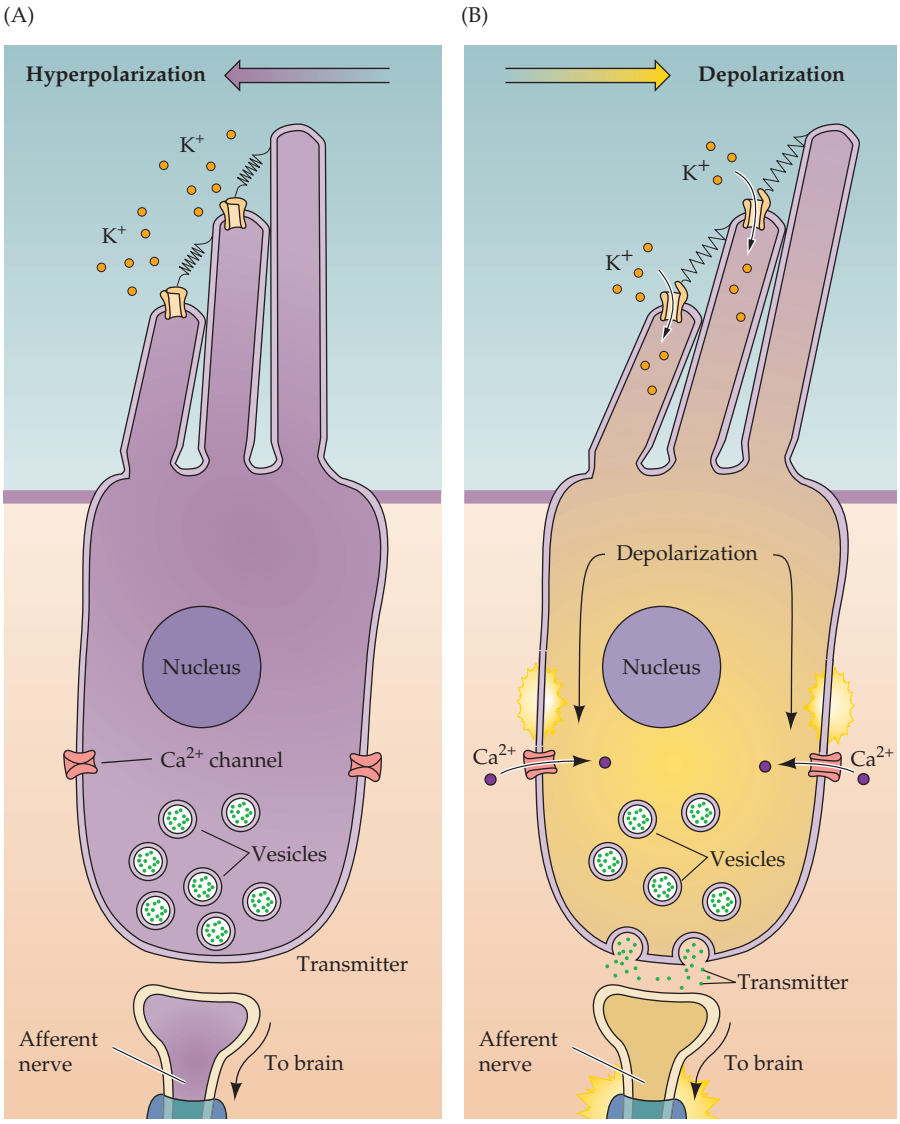


Figure 12.8 Mechanoelectrical transduction mediated by hair cells. (A,B) When the hair bundle is deflected toward the tallest stereocilium, cation-selective channels open near the tips of the stereocilia, allowing K^+ ions to flow into the hair cell down their electrochemical gradient (see text on next page for the explanation of this peculiar situation). The resulting depolarization of the hair cell opens voltage-gated Ca^{2+} channels in the cell soma, allowing calcium entry and release of neurotransmitter onto the nerve endings of the auditory nerve. (After Lewis and Hudspeth, 1983)

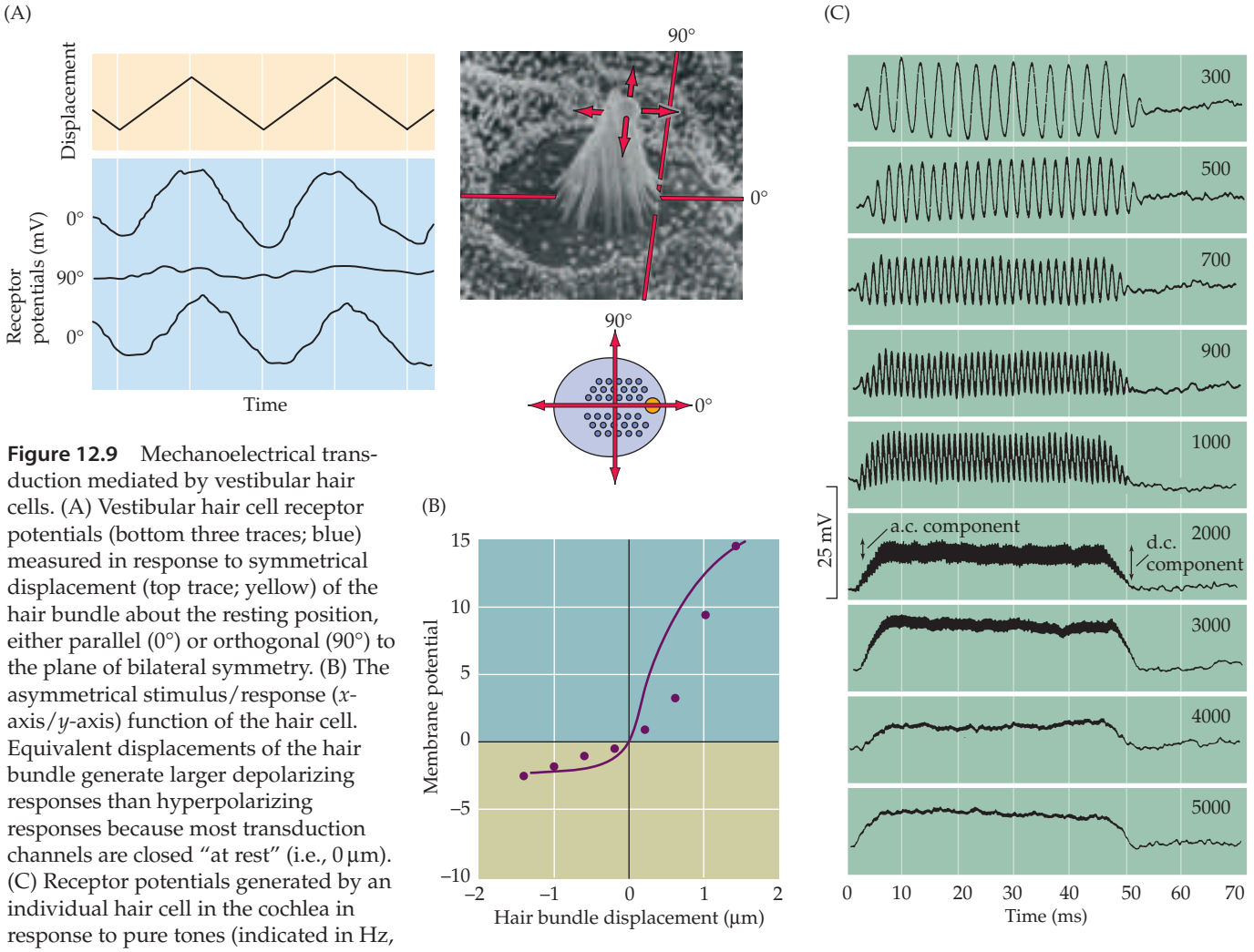


Figure 12.9 Mechanoelectrical transduction mediated by vestibular hair cells. (A) Vestibular hair cell receptor potentials (bottom three traces; blue) measured in response to symmetrical displacement (top trace; yellow) of the hair bundle about the resting position, either parallel (0°) or orthogonal (90°) to the plane of bilateral symmetry. (B) The asymmetrical stimulus/response (*x*-axis/*y*-axis) function of the hair cell. Equivalent displacements of the hair bundle generate larger depolarizing responses than hyperpolarizing responses because most transduction channels are closed "at rest" (i.e., 0 μm). (C) Receptor potentials generated by an individual hair cell in the cochlea in response to pure tones (indicated in Hz, right). Note that the hair cell potential faithfully follows the waveform of the stimulating sinusoids for low frequencies (< 3kHz), but still responds with a DC offset to higher frequencies. (A after Shotwell et al., 1981; B after Hudspeth and Corey, 1977; C after Palmer and Russell, 1986.)

ear to environmental and genetic insults. An important goal of current research is to identify the stem cells and factors that could contribute to the regeneration of human hair cells, thus affording a possible therapy for some forms of sensorineural hearing loss.

Understanding the ionic basis of hair cell transduction has been greatly advanced by intracellular recordings made from these tiny structures. The hair cell has a resting potential between –45 and –60 mV relative to the fluid that bathes the basal end of the cell. At the resting potential, only a small fraction of the transduction channels are open. When the hair bundle is displaced in the direction of the tallest stereocilium, more transduction channels open, causing depolarization as K^+ enters the cell. Depolarization in turn opens voltage-gated calcium channels in the hair cell membrane, and the resultant Ca^{2+} influx causes transmitter release from the basal end of the cell onto the auditory nerve endings (Figure 12.8A,B). Such calcium-dependent exocytosis is similar to chemical neurotransmission elsewhere in the central and peripheral nervous system (see Chapters 5 and 6); thus the hair cell has become a useful model for studying calcium-dependent transmitter release. Because some transduction channels are open at rest, the receptor potential is biphasic: Movement toward the tallest stereocilia depolarizes the cell, while move-

ment in the opposite direction leads to hyperpolarization. This situation allows the hair cell to generate a sinusoidal receptor potential in response to a sinusoidal stimulus, thus preserving the temporal information present in the original signal up to frequencies of around 3 kHz (Figure 12.9). Hair cells still can signal at frequencies above 3 kHz, although without preserving the exact temporal structure of the stimulus: the asymmetric displacement-receptor current function of the hair cell bundle is filtered by the cell's membrane time constant to produce a tonic depolarization of the soma, augmenting transmitter release and thus exciting VIIIth nerve terminals.

The high-speed demands of mechanoelectrical transduction have resulted in some impressive ionic specializations within the inner ear. An unusual adaptation of the hair cell in this regard is that K^+ serves both to depolarize *and* repolarize the cell, enabling the hair cell's K^+ gradient to be largely maintained by passive ion movement alone. As with other epithelial cells, the basal and apical surfaces of the hair cell are separated by tight junctions, allowing separate extracellular ionic environments at these two surfaces. The apical end (including the stereocilia) protrudes into the scala media and is exposed to the K^+-rich, Na^+-poor **endolymph**, which is produced by dedicated ion pumping cells in the **stria vascularis** (Figure 12.10). The basal end of the hair cell body is bathed in the same fluid that fills the scala tympani, the perilymph, which resembles other extracellular fluids in that it is K^+-poor and Na^+-rich. In addition, the compartment containing endolymph is about 80 mV more positive than the perilymph compartment (this difference is known as the endocochlear potential), while the inside of the hair cell is about 45 mV more negative than the perilymph (and 125 mV more negative than the endolymph). The resulting electrical gradient across the membrane of the stereocilia (about 125 mV; the difference between the hair cell resting potential and the endocochlear potential) drives K^+ through open transduc-

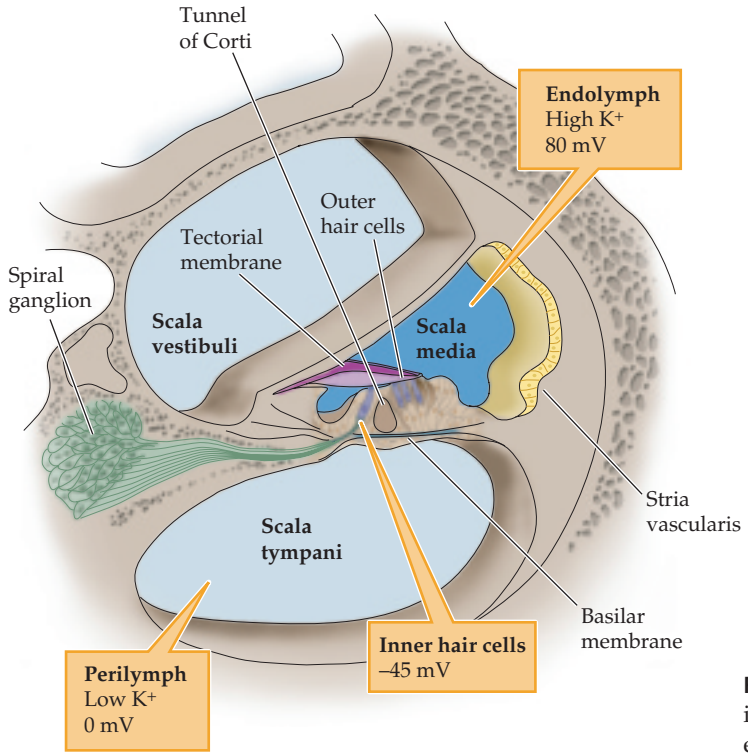


Figure 12.10 The stereocilia of the hair cells protrude into the endolymph, which is high in K^+ and has an electrical potential of +80 mV relative to the perilymph.

tion channels into the hair cell, even though these cells already have a high internal K^+ concentration. K^+ entry via the transduction channels electrotonically depolarizes the hair cell, opening voltage-gated Ca^{2+} and K^+ channels located in the membrane of the hair cell soma (see Box B in Chapter 13). The opening of *somatic* K^+ channels favors K^+ efflux, and thus repolarization; the efflux occurs because the perilymph surrounding the basal end is low in K^+ relative to the cytosol, and because the equilibrium potential for K^+ is more negative than the hair cell's resting potential ($E_{KBasal} \approx -85$ mV). Repolarization of the hair cell via K^+ efflux is also facilitated by Ca^{2+} entry. In addition to modulating the release of neurotransmitter, Ca^{2+} entry opens Ca^{2+}-dependent K^+ channels, which provide another avenue for K^+ to enter the perilymph. Indeed, the interaction of Ca^{2+} influx and Ca^{2+}-dependent K^+ efflux can lead to electrical resonances that enhance the tuning of response properties within the inner ear (also explained in Box B in Chapter 13). In essence, the hair cell operates as two distinct compartments, each dominated by its own Nernst equilibrium potential for K^+; this arrangement ensures that the hair cell's ionic gradient does not run down, even during prolonged stimulation. At the same time, rupture of Reissner's membrane, which normally separates the scalae media and vestibuli, or compounds such as ethacrynic acid (see Box A), which selectively poison the ion-pumping cells of the stria vascularis, can cause the endocochlear potential to dissipate, resulting in a sensorineural hearing deficit. In short, the hair cell exploits the different ionic milieus of its apical and basal surfaces to provide extremely fast and energy-efficient repolarization.

Two Kinds of Hair Cells in the Cochlea

The cochlear hair cells in humans consist of one row of **inner hair cells** and three rows of **outer hair cells** (see Figure 12.4). The inner hair cells are the actual sensory receptors, and 95% of the fibers of the auditory nerve that project to the brain arise from this subpopulation. The terminations on the outer hair cells are almost all from efferent axons that arise from cells in the superior olivary complex.

A clue to the significance of this efferent pathway was provided by the discovery that an active process within the cochlea, as mentioned, influences basilar membrane motion. First, it was found that the cochlea actually emits sound under certain conditions. These otoacoustical emissions can be detected by placing a sensitive microphone at the eardrum and monitoring the response after briefly presenting a tone or click, and provide a useful means to assess cochlear function in the newborn (this test is now done routinely to rule out congenital deafness). Such emissions can also occur spontaneously, especially in certain pathological states, and are thus one potential source of **tinnitus** (ringing in the ears). These observations clearly indicate that a process within the cochlea is capable of producing sound. Second, stimulation of the crossed olivocochlear bundle, which supplies efferent input to the outer hair cells, can broaden VIIIth nerve tuning curves. Third, the high sensitivity notch of VIIIth nerve tuning curves is lost when the outer hair cells are selectively inactivated. Finally, isolated outer hair cells contract and expand in response to small electrical currents, thus providing a potential source of energy to drive an active process within the cochlea. Thus, it seems likely that the outer hair cells sharpen the frequency-resolving power of the cochlea by actively contracting and relaxing, thus changing the stiffness of the tectorial membrane at particular locations. This active

process explains the nonlinear vibration of the basilar membrane at low sound intensities (see Box D).

Tuning and Timing in the Auditory Nerve

The rapid response time of the transduction apparatus allows the membrane potential of the hair cell to follow deflections of the hair bundle up to moderately high frequencies of oscillation. In humans, the receptor potentials of certain hair cells and the action potentials of their associated auditory nerve fiber can follow stimuli of up to about 3 kHz in a one-to-one fashion. Such real-time encoding of stimulus frequency by the pattern of action potentials in the auditory nerve is known as the "volley theory" of auditory information transfer. Even these extraordinarily rapid processes, however, fail to follow frequencies above 3 kHz (see Figure 12.9). Accordingly, some other mechanism must be used to transmit auditory information at higher frequencies. The tonotopically organized basilar membrane provides an alternative to temporal coding, namely a "labeled-line" coding mechanism. In this case, frequency information is specified by preserving the tonotopy of the cochlea at higher levels in the auditory pathway. Because the auditory nerve fibers associate with the inner hair cells in approximately a one-to-one ratio (although several or more VIIIth nerve fibers synapse on a single hair cell), each auditory nerve fiber transmits information about only a small part of the audible frequency spectrum. As a result, auditory nerve fibers related to the apical end of the cochlea respond to low frequencies, and fibers that are related to the basal end respond to high frequencies (see Figure 12.5). The limitations of specific fibers can be seen in electrophysiological recordings of responses to sound (Figure 12.11). These threshold functions are called **tuning curves**, and the lowest threshold of the tuning curve is called the **characteristic frequency**. Since the topographical order of the characteristic frequency of neurons is retained throughout the system, information about frequency is also preserved. Cochlear implants (see Box C) exploit the tonotopic organization of the cochlea, and particularly its eighth nerve afferents, to roughly recreate the patterns of VIIIth nerve activity elicited by sounds. In patients with damaged hair cells, such implants can effectively bypass the impaired transduction apparatus, and thus restore some degree of auditory function.

The other prominent feature of hair cells—their ability to follow the waveform of low-frequency sounds—is also important in other more subtle aspects of auditory coding. As mentioned earlier, hair cells have biphasic response properties. Because hair cells release transmitter only when depolarized, auditory nerve fibers fire only during the positive phases of low-frequency sounds (see Figure 12.11). The resultant "phase locking" that results provides temporal information from the two ears to neural centers that compare interaural time differences. The evaluation of interaural time differences provides a critical cue for sound localization and the perception of auditory "space." That auditory space can be perceived is remarkable, given that the cochlea, unlike the retina, cannot represent space directly. A final point is that VIIIth nerve activity patterns are not simply faithful neural replicas of the auditory stimulus itself. Indeed, William Bialek and his colleagues at Princeton University have shown that the VIIIth nerve in the bullfrog encodes conspecific mating calls more efficiently than artificial sounds with similar frequency and amplitude characteristics. Thus both animal and human studies support the idea that the auditory periphery is optimized to

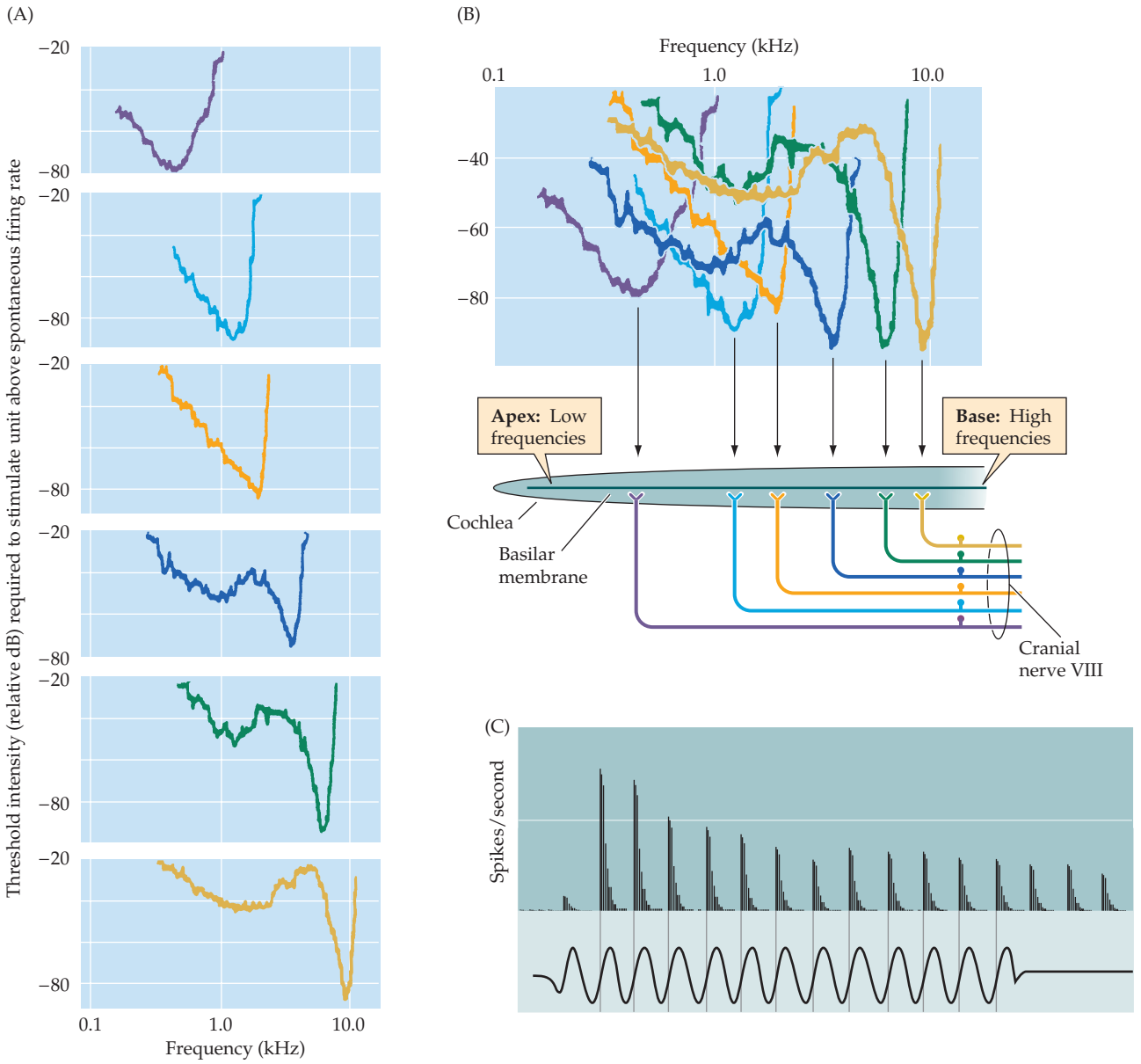


Figure 12.11 Response properties of auditory nerve fibers. (A) Frequency tuning curves of six different fibers in the auditory nerve. Each graph plots, across all frequencies to which the fiber responds, the minimum sound level required to increase the fiber's firing rate above its spontaneous firing level. The lowest point in the plot is the weakest sound intensity to which the neuron will respond. The frequency at this point is called the neuron's characteristic frequency. (B) The frequency tuning curves of auditory nerve fibers superimposed and aligned with their approximate relative points of innervation along the basilar membrane. (In the side view schematic, the basilar membrane is represented as a black line within the unrolled cochlea.) (C) Temporal response patterns of a low-frequency axon in the auditory nerve. The stimulus waveform is indicated beneath the histograms, which show the phase-locked responses to a 50-ms tone pulse of 260 Hz. Note that the spikes are all timed to the same phase of the sinusoidal stimulus. (A after Kiang and Moxon, 1972; C after Kiang, 1984.)

transmit species-typical vocal sounds, rather than simply transmitting all sounds equally well to central auditory areas.

How Information from the Cochlea Reaches Targets in the Brainstem

A hallmark of the ascending auditory system is its parallel organization. This arrangement becomes evident as soon as the auditory nerve enters the brainstem, where it branches to innervate the three divisions of the cochlear nucleus. The auditory nerve (the major component of cranial nerve VIII) comprises the central processes of the bipolar spiral ganglion cells in the cochlea (see Figure 12.4); each of these cells sends a peripheral process to contact one inner hair cell and a central process to innervate the cochlear nucleus. Within the cochlear nucleus, each auditory nerve fiber branches, sending an ascending branch to the anteroventral cochlear nucleus and a descending branch to the posteroventral cochlear nucleus and the dorsal cochlear nucleus (Figure 12.12). The tonotopic organization of the cochlea is maintained in the three parts of the cochlear nucleus, each of which contains different populations of cells with quite different properties. In addition, the patterns of termination of the auditory nerve axons differ in density and type; thus, there are several opportunities at this level for transformation of the information from the hair cells.

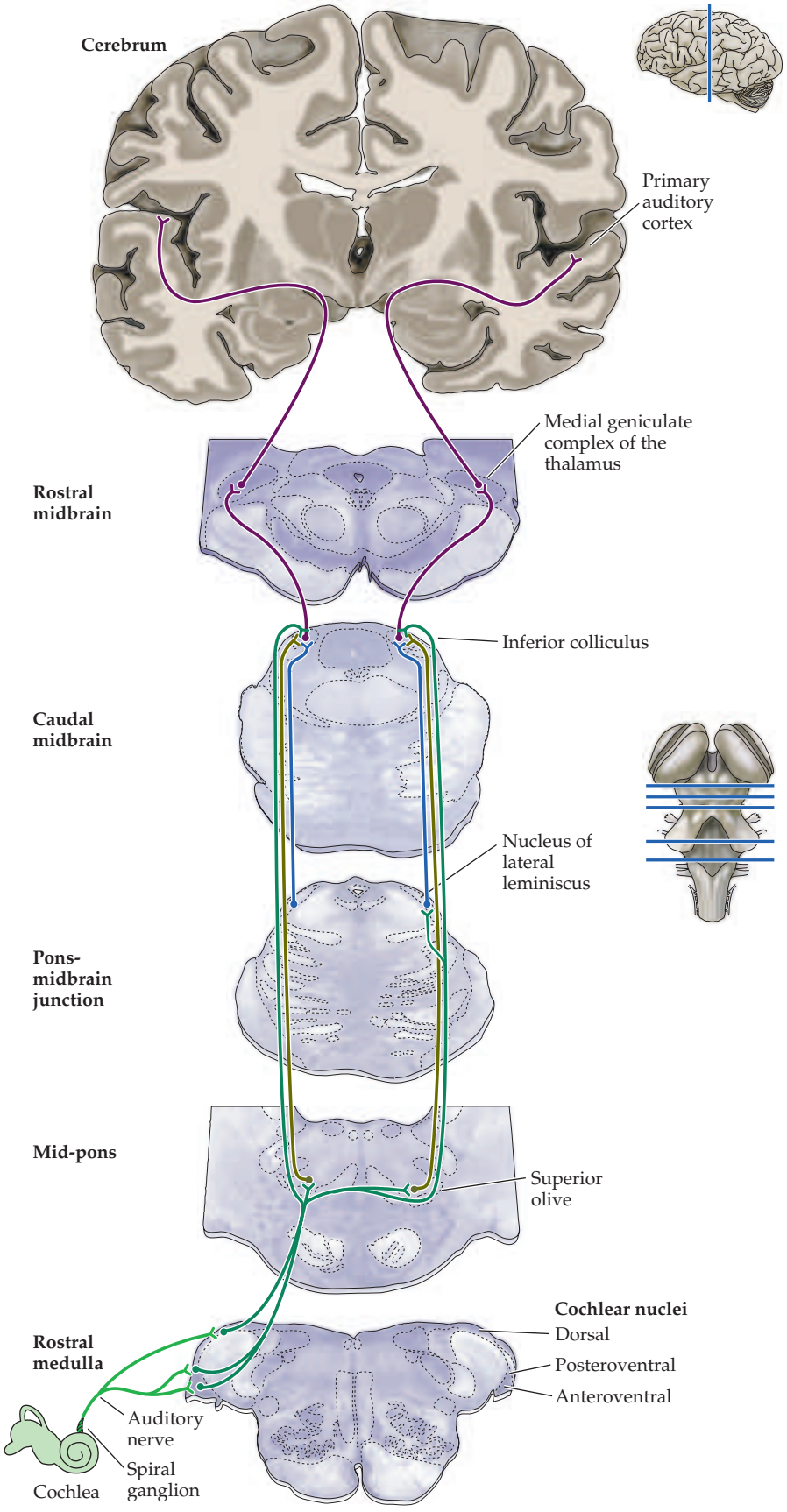
Integrating Information from the Two Ears

Just as the auditory nerve branches to innervate several different targets in the cochlear nuclei, the neurons in these nuclei give rise to several different pathways (see Figure 12.12). One clinically relevant feature of the ascending projections of the auditory brainstem is a high degree of bilateral connectivity, which means that damage to central auditory structures is almost never manifested as a monaural hearing loss. Indeed, monaural hearing loss strongly implicates unilateral peripheral damage, either to the middle or inner ear, or to the VIIIth nerve itself (see Box C). Given the relatively byzantine organization already present at the level of the auditory brainstem, it is useful to consider these pathways in the context of their functions.

The best-understood function mediated by the auditory brainstem nuclei, and certainly the one most intensively studied, is sound localization. Humans use at least two different strategies to localize the horizontal position of sound sources, depending on the frequencies in the stimulus. For frequencies below 3 kHz (which can be followed in a phase-locked manner), interaural *time* differences are used to localize the source; above these frequencies, interaural *intensity* differences are used as cues. Parallel pathways originating from the cochlear nucleus serve each of these strategies for sound localization.

The human ability to detect interaural time differences is remarkable. The longest interaural time differences, which are produced by sounds arising directly lateral to one ear, are on the order of only 700 microseconds (a value given by the width of the head divided by the speed of sound in air, about 340 m/s). Psychophysical experiments show that humans can actually detect interaural time differences as small as 10 microseconds; two sounds presented through earphones separated by such small interaural time differences are perceived as arising from the side of the leading ear. This sensitivity translates into accuracy for sound localization of about 1 degree.

Figure 12.12 Diagram of the major auditory pathways. Although many details are missing from this diagram, two important points are evident: (1) the auditory system entails several parallel pathways, and (2) information from each ear reaches both sides of the system, even at the level of the brainstem.



How is timing in the 10 microseconds range accomplished by neural components that operate in the millisecond range? The neural circuitry that computes such tiny interaural time differences consists of binaural inputs to the **medial superior olive** (**MSO**) that arise from the right and left anteroventral cochlear nuclei (Figure 12.13; see also Figure 12.12). The medial superior olive contains cells with bipolar dendrites that extend both medially and laterally. The lateral dendrites receive input from the ipsilateral anteroventral cochlear nucleus, and the medial dendrites receive input from the contralateral anteroventral cochlear nucleus (both inputs are excitatory). As might be expected, the MSO cells work as **coincidence detectors**, responding when both excitatory signals arrive at the same time. For a coincidence mechanism to be useful in localizing sound, different neurons must be maximally sensitive to different interaural time delays. The axons that project from the anteroventral cochlear nucleus evidently vary systematically in length to create delay lines. (Remember that the length of an axon divided by its conduction velocity equals the conduction time.) These anatomical differences compensate for sounds arriving at slightly different times at the two ears, so that the resultant neural impulses arrive at a particular MSO neuron simultaneously, making each cell especially sensitive to sound sources in a particular place. The mechanisms enabling MSO neurons to function as coincidence detectors at the microsecond level are still poorly understood, but certainly reflect one of the more impressive biophysical specializations in the nervous system.

Sound localization perceived on the basis of interaural time differences requires phase-locked information from the periphery, which, as already

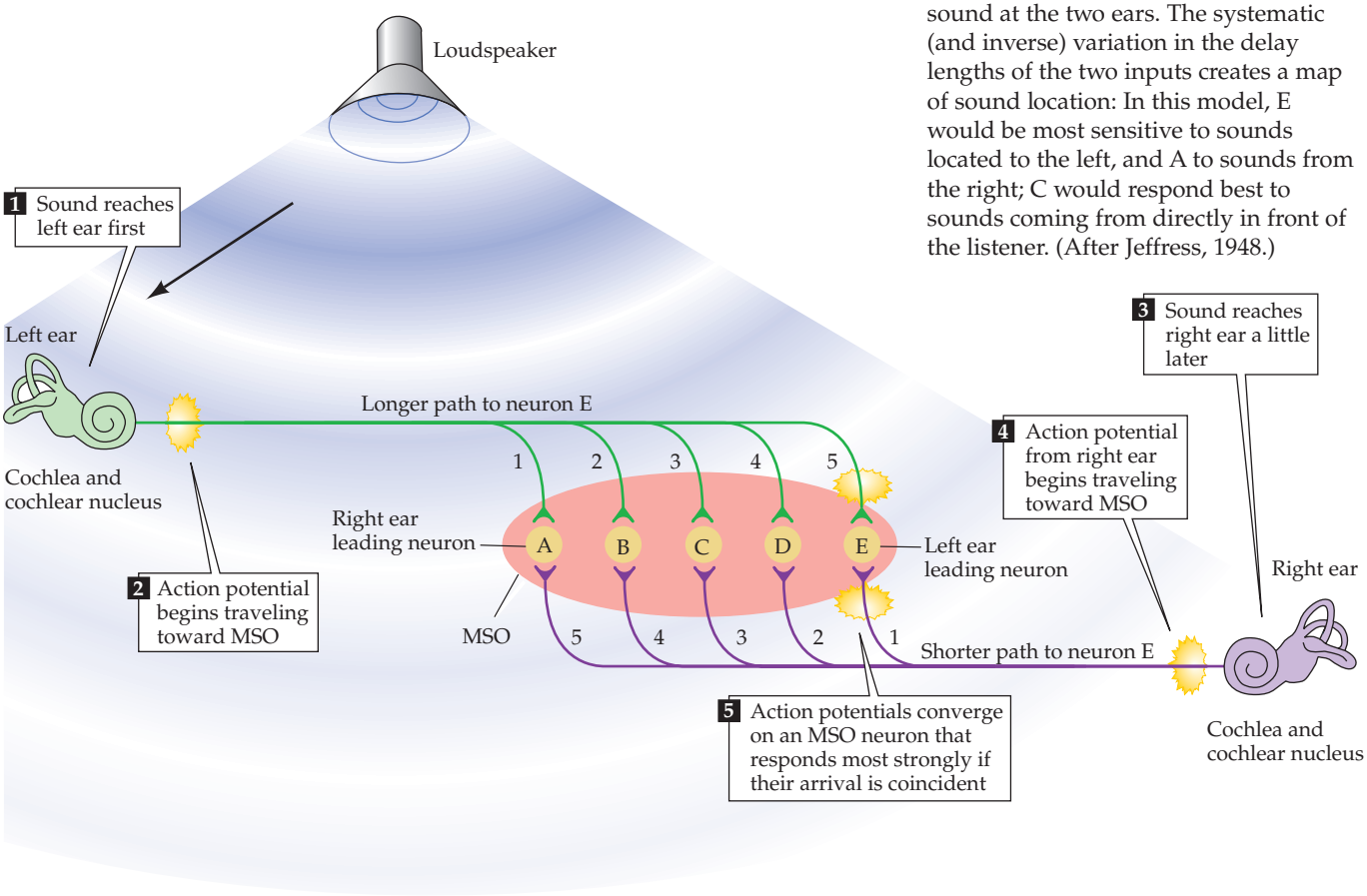


Figure 12.13 Diagram illustrating how the MSO computes the location of a sound by interaural time differences. A given MSO neuron responds most strongly when the two inputs arrive simultaneously, as occurs when the contralateral and ipsilateral inputs precisely compensate (via their different lengths) for differences in the time of arrival of a sound at the two ears. The systematic (and inverse) variation in the delay lengths of the two inputs creates a map of sound location: In this model, E would be most sensitive to sounds located to the left, and A to sounds from the right; C would respond best to sounds coming from directly in front of the listener. (After Jeffress, 1948.)

emphasized, is available to humans only for frequencies below 3 kHz. (In barn owls, the reigning champions of sound localization, phase locking occurs at up to 9 kHz.) Therefore, a second mechanism must come into play at higher frequencies. At frequencies higher than about 2 kHz, the human head begins to act as an acoustical obstacle because the wavelengths of the sounds are too short to bend around it. As a result, when high-frequency sounds are directed toward one side of the head, an acoustical "shadow" of lower intensity is created at the far ear. These intensity differences provide a second cue about the location of a sound. The circuits that compute the position of a sound source on this basis are found in the **lateral superior olive** (**LSO**) and the **medial nucleus of the trapezoid body** (**MNTB**) (Figure 12.14). Excitatory axons project directly from the ipsilateral anteroventral cochlear nucleus to the LSO (as well as to the MSO; see Figure 12.13). Note that the LSO also receives inhibitory input from the contralateral ear, via an inhibitory neuron in the MNTB. This excitatory/inhibitory interaction

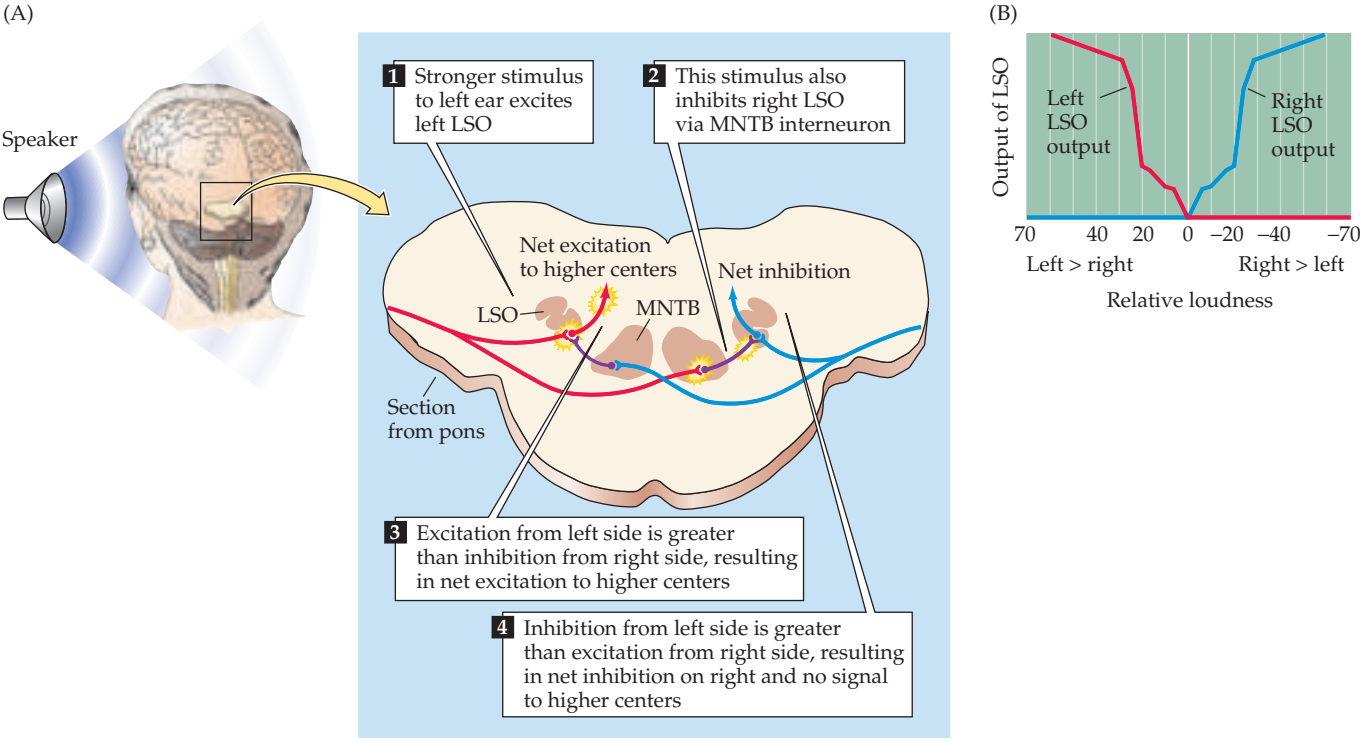


Figure 12.14 Lateral superior olive neurons encode sound location through interaural intensity differences. (A) LSO neurons receive direct excitation from the ipsilateral cochlear nucleus; input from the contralateral cochlear nucleus is relayed via inhibitory interneurons in the MNTB. (B) This arrangement of excitation–inhibition makes LSO neurons fire most strongly in response to sounds arising directly lateral to the listener on the same side as the LSO, because excitation from the ipsilateral input will be great and inhibition from the contralateral input will be small. In contrast, sounds arising from in front of the listener, or from the opposite side, will silence the LSO output, because excitation from the ipsilateral input will be minimal, but inhibition driven by the contralateral input will be great. Note that LSOs are paired and bilaterally symmetrical; each LSO only encodes the location of sounds arising on the same side of the body as its location.

results in a net excitation of the LSO on the same side of the body as the sound source. For sounds arising directly lateral to the listener, firing rates will be highest in the LSO on that side; in this circumstance, the excitation via the ipsilateral anteroventral cochlear nucleus will be maximal, and inhibition from the contralateral MNTB minimal. In contrast, sounds arising closer to the listener's midline will elicit lower firing rates in the ipsilateral LSO because of increased inhibition arising from the contralateral MNTB. For sounds arising at the midline, or from the other side, the increased inhibition arising from the MNTB is powerful enough to completely silence LSO activity. Note that each LSO only encodes sounds arising in the ipsilateral hemifield; it therefore takes both LSOs to represent the full range of horizontal positions.

In summary, there are two separate pathways—and two separate mechanisms—for localizing sound. Interaural time differences are processed in the medial superior olive, and interaural intensity differences are processed in the lateral superior olive. These two pathways are eventually merged in the midbrain auditory centers.

Monaural Pathways from the Cochlear Nucleus to the Lateral Lemniscus

The binaural pathways for sound localization are only part of the output of the cochlear nucleus. This fact is hardly surprising, given that auditory perception involves much more than locating the position of the sound source. A second major set of pathways from the cochlear nucleus bypasses the superior olive and terminates in the **nuclei of the lateral lemniscus** on the contralateral side of the brainstem (see Figure 12.12). These particular pathways respond to sound arriving at one ear only and are thus referred to as monaural. Some cells in the lateral lemniscus nuclei signal the onset of sound, regardless of its intensity or frequency. Other cells in the lateral lemniscus nuclei process other temporal aspects of sound, such as duration. The precise role of these pathways in processing temporal features of sound is not yet known. As with the outputs of the superior olivary nuclei, the pathways from the nuclei of the lateral lemniscus converge at the midbrain.

Integration in the Inferior Colliculus

Auditory pathways ascending via the olivary and lemniscal complexes, as well as other projections that arise directly from the cochlear nucleus, project to the midbrain auditory center, the **inferior colliculus**. In examining how integration occurs in the inferior colliculus, it is again instructive to turn to the most completely analyzed auditory mechanism, the binaural system for localizing sound. As already noted, space is not mapped on the auditory receptor surface; thus the perception of auditory space must somehow be synthesized by circuitry in the lower brainstem and midbrain. Experiments in the barn owl, an extraordinarily proficient animal at localizing sounds, show that the convergence of binaural inputs in the midbrain produces something entirely new relative to the periphery—namely, a computed topographical representation of auditory space. Neurons within this **auditory space map** in the colliculus respond best to sounds originating in a specific region of space and thus have both a preferred elevation and a preferred horizontal location, or azimuth. Although comparable maps of auditory space have not yet been found in mammals, humans have a clear perception of

both the elevational and azimuthal components of a sound's location, suggesting that we have a similar auditory space map.

Another important property of the inferior colliculus is its ability to process sounds with complex temporal patterns. Many neurons in the inferior colliculus respond only to frequency-modulated sounds, while others respond only to sounds of specific durations. Such sounds are typical components of biologically relevant sounds, such as those made by predators, or intraspecific communication sounds, which in humans include speech.

The Auditory Thalamus

Despite the parallel pathways in the auditory stations of the brainstem and midbrain, the **medial geniculate complex** (**MGC**) in the thalamus is an obligatory relay for all ascending auditory information destined for the cortex (see Figure 12.12). Most input to the MGC arises from the inferior colliculus, although a few auditory axons from the lower brainstem bypass the inferior colliculus to reach the auditory thalamus directly. The MGC has several divisions, including the ventral division, which functions as the major thalamocortical relay, and the dorsal and medial divisions, which are organized like a belt around the ventral division.

In some mammals, the strictly maintained tonotopy of the lower brainstem areas is exploited by convergence onto MGC neurons, generating specific responses to certain spectral combinations. The original evidence for this statement came from research on the response properties of cells in the MGC of echolocating bats. Some cells in the so-called belt areas of the bat MGC respond only to combinations of widely spaced frequencies that are specific components of the bat's echolocation signal and of the echoes that are reflected from objects in the bat's environment. In the mustached bat, where this phenomenon has been most thoroughly studied, the echolocation pulse has a changing frequency (frequency-modulated, or FM) component that includes a fundamental frequency and one or more harmonics. The fundamental frequency (FM_1) has low intensity and sweeps from 30 kHz to 20 kHz. The second harmonic (FM_2) is the most intense component and sweeps from 60 kHz to 40 kHz. Note that these frequencies do not overlap. Most of the echoes are from the intense FM_2 sound, and virtually none arise from the weak FM_1, even though the emitted FM_1 is loud enough for the bat to hear. Apparently, the bat assesses the distance to an object by measuring the delay between the FM_1 emission and the FM_2 echo. Certain MGC neurons respond when FM_2 follows FM_1 by a specific delay, providing a mechanism for sensing such frequency combinations. Because each neuron responds best to a particular delay, the population of MGC neurons encodes a range of distances.

Bat sonar illustrates two important points about the function of the auditory thalamus. First, the MGC is the first station in the auditory pathway where selectivity for combinations of frequencies is found. The mechanism responsible for this selectivity is presumably the ultimate convergence of inputs from cochlear areas with different spectral sensitivities. Second, cells in the MGC are selective not only for frequency combinations, but also for specific time intervals between the two frequencies. The principle is the same as that described for binaural neurons in the medial superior olive, but in this instance, two monaural signals with different frequency sensitivity coincide, and the time difference is in the millisecond rather than the microsecond range.

In summary, neurons in the medial geniculate complex receive convergent inputs from spectrally and temporally separate pathways. This complex, by

virtue of its convergent inputs, mediates the detection of specific spectral and temporal combinations of sounds. In many species, including humans, varying spectral and temporal cues are especially important features of communication sounds. It is not known whether cells in the human medial geniculate are selective to combinations of sounds, but the processing of speech certainly requires both spectral and temporal combination sensitivity.

The Auditory Cortex

The ultimate target of afferent auditory information is the auditory cortex. Although the auditory cortex has a number of subdivisions, a broad distinction can be made between a primary area and peripheral, or belt, areas. The **primary auditory cortex** (**A1**) is located on the superior temporal gyrus in the temporal lobe and receives point-to-point input from the ventral division of the medial geniculate complex; thus, it contains a precise tonotopic map. The **belt areas** of the auditory cortex receive more diffuse input from the belt areas of the medial geniculate complex and therefore are less precise in their tonotopic organization.

The primary auditory cortex (A1) has a topographical map of the cochlea (Figure 12.15), just as the primary visual cortex (V1) and the primary somatic sensory cortex (S1) have topographical maps of their respective sensory

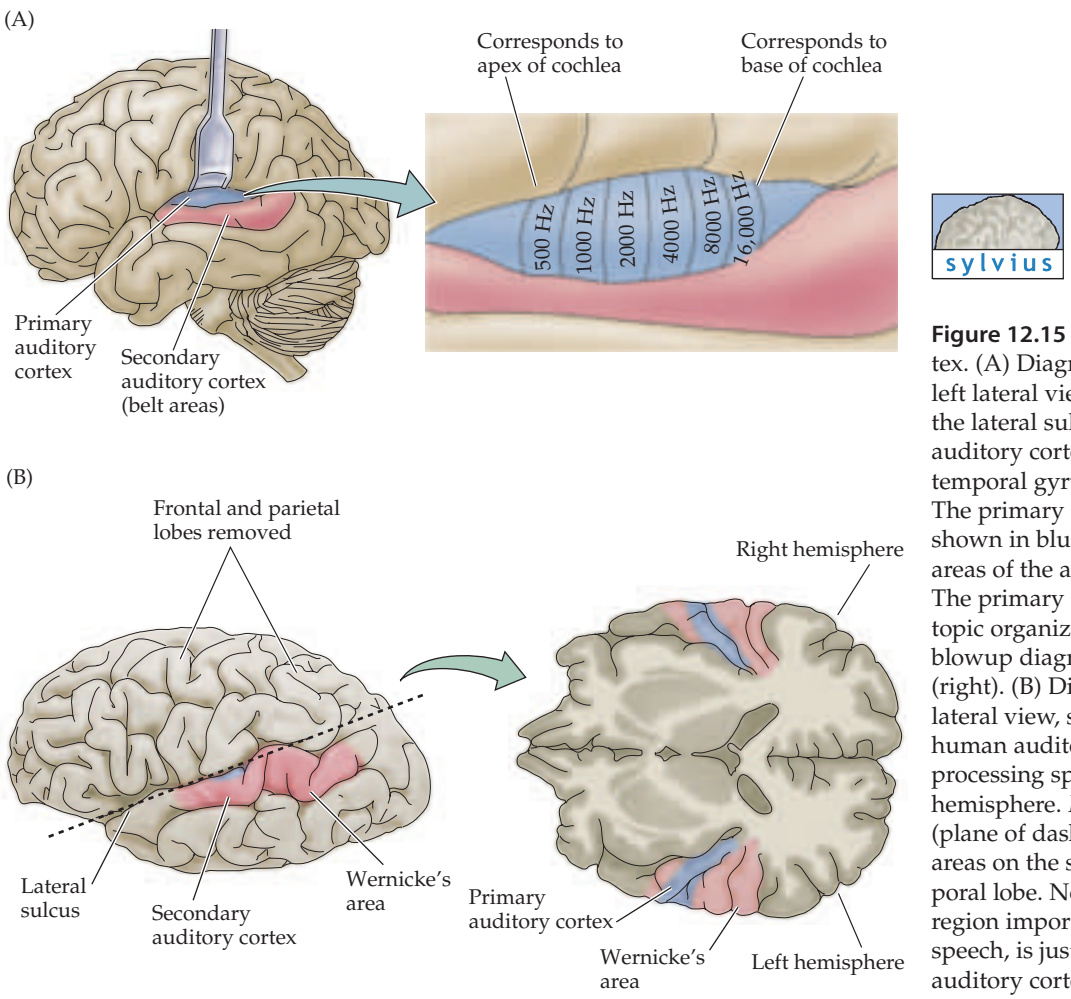
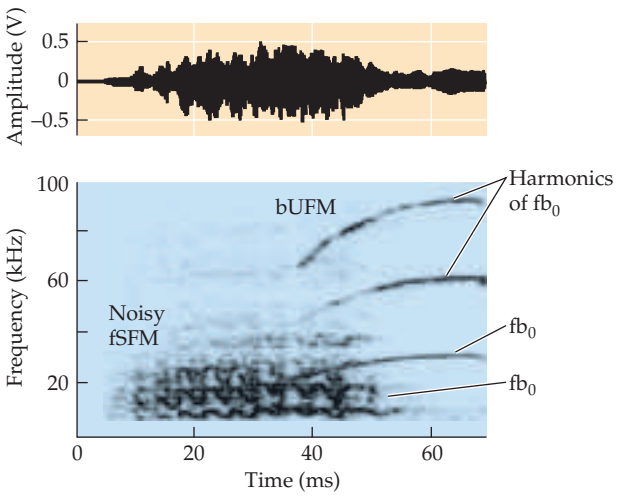


Figure 12.15 The human auditory cortex. (A) Diagram showing the brain in left lateral view, including the depths of the lateral sulcus, where part of the auditory cortex occupying the superior temporal gyrus normally lies hidden. The primary auditory cortex (A1) is shown in blue; the surrounding belt areas of the auditory cortex are in red. The primary auditory cortex has a tonotopic organization, as shown in this blowup diagram of a segment of A1 (right). (B) Diagram of the brain in left lateral view, showing locations of human auditory cortical areas related to processing speech sounds in the intact hemisphere. *Right:* An oblique section (plane of dashed line) shows the cortical areas on the superior surface of the temporal lobe. Note that Wernicke's area, a region important in comprehending speech, is just posterior to the primary auditory cortex.

Box E

Representing Complex Sounds in the Brains of Bats and Humans

Most natural sounds are complex, meaning that they differ from the pure tones or clicks that are frequently used in neurophysiological studies of the auditory system. Rather, natural sounds are tonal: they have a fundamental frequency that largely determines the "pitch" of the sound, and one or more harmonics of different intensities that contribute to the quality or "timbre" of a sound. The frequency of a harmonic is, by definition, a multiple of the fundamental frequency, and both may be modulated over time. Such *frequency-modulated* (FM) sweeps can rise or fall in frequency, or change in a sinusoidal or some other fashion. Occasionally, multiple nonharmonic frequencies may be simultaneously present in some communication or musical sounds. In some sounds, a level of spectral splatter or "broadband noise" is embedded within tonal or frequency modulated sounds. The variations in the sound spectrum are typically accompanied by a modulation of the amplitude envelop of the complex sound as well. All of these features can be visualized by performing a spectrographic analysis.

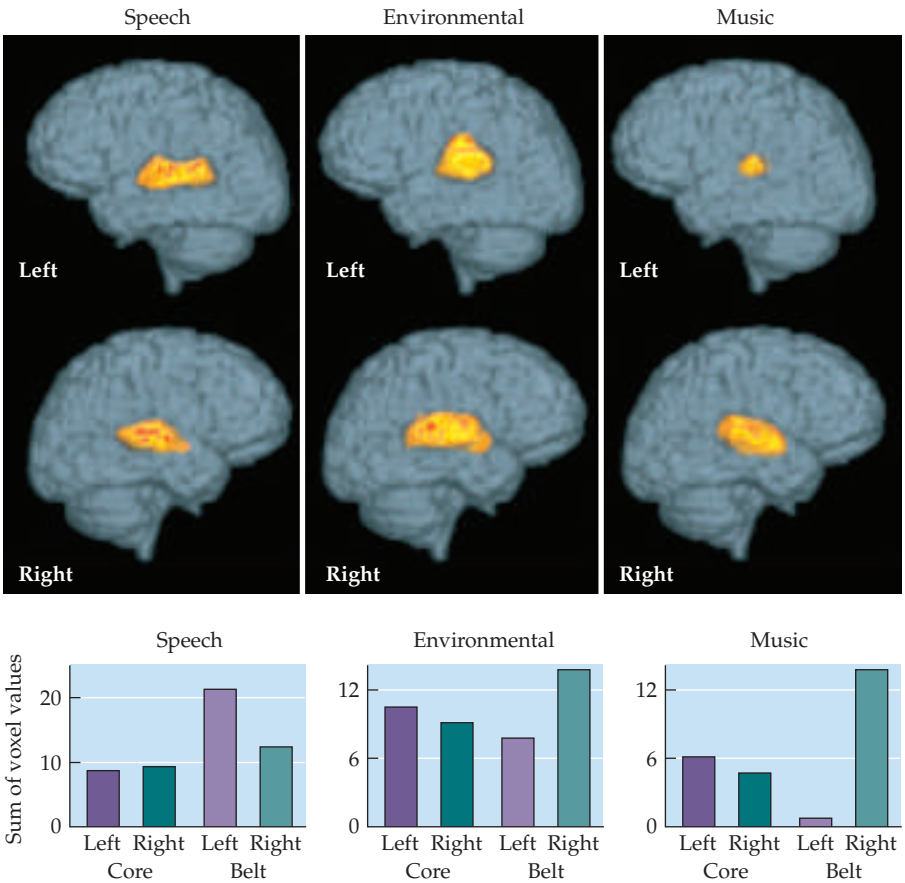


(A) Amplitude envelope (above) and spectrogram (below) of a composite syllable emitted by mustached bats for social communication. This composite consists of two simple syllables, a fixed Sinusoidal FM (fSFM) and a bent Upward FM (bUFM) that emerges from the fSFM after some overlap. Each syllable has its own fundamental (fa_0 and fb_0) and multiple harmonics. (Courtesy of Jagmeet Kanwal.)

How does the brain represent such complex natural sounds? Cognitive studies of complex sound perception provide some understanding of how a large but limited number of neurons in the brain can dynamically represent an infinite variety of natural stimuli in the sensory environment of humans and other animals. In bats, specializations for processing complex sounds are apparent. Studies in echolocating bats show that both communication and echolocation sounds (Figure A) are processed not only within some of the same areas, but also within the same neurons in the auditory cortex. In humans, multiple modes of processing are also likely, given the large overlap within the superior and middle temporal gyri in the temporal lobe for the representation of different types of complex sounds.

Asymmetrical representation is another common principle of complex sound processing that results in lateralized (though largely overlapping) representations of natural stimuli. Thus, speech sounds that are important for communication are lateralized to the left in the belt regions of the auditory cortex, whereas environmental sounds that are important for reacting to and recogniz-

epithelia. Unlike the visual and somatic sensory systems, however, the cochlea has already decomposed the acoustical stimulus so that it is arrayed tonotopically along the length of the basilar membrane. Thus, A1 is said to comprise a tonotopic map, as do most of the ascending auditory structures between the cochlea and the cortex. Orthogonal to the frequency axis of the tonotopic map is a striped arrangement of binaural properties. The neurons in one stripe are excited by both ears (and are therefore called EE cells), while the neurons in the next stripe are excited by one ear and inhibited by the other ear (EI cells). The EE and EI stripes alternate, an arrangement that is reminiscent of the ocular dominance columns in V1 (see Chapter 11).



(B) *Top:* Reconstructed functional magnetic resonance images of BOLD contrast signal change (average for 8 subjects) showing significant ($p < 0.001$) activation elicited by speech, environmental, and musical sounds on surface views of the left versus the right side of the human brain. *Bottom:* Bar graphs showing the total significant activation to each category of complex sounds in the core and belt areas of the auditory cortex for the left versus the right side. (Courtesy of Jagmeet Kanwal.)

ing aspects of the auditory environment are represented in each hemisphere (Figure B). Musical sounds that can either motivate us to march in war or to relax and meditate when coping with physical and emotional stress are highly lateralized to the right in the belt regions of the auditory cortex. The extent of lateralization for speech and possibly music may vary with sex, age, and training. In some species of bats, mice, and primates, processing of natural communication sounds appears to be lateralized to the left hemisphere. In summary, natural sounds are complex and their representation within the sensory cortex tends to be asymmetric across the two hemispheres.

References

Ehret, G. (1987) Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. Nature 325: 249–251.

Esser, K.-H., C. J. Condon, N. Suga and J. S. Kanwal (1997) Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat, *Pteronotus parnellii*. Proc. Natl. Acad. Sci. USA 94: 14019–14024.

Hauser, M. D. and K. Andersson (1994) Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: Field experiments. Proc. Natl. Acad. Sci. USA 91: 3946-3948.

Kanwal, J. S., J. Kim and K. Kamada (2000) Separate, distributed processing of environmental, speech and musical sounds in the cerebral hemispheres. J. Cog. Neurosci. (Supp.): p. 32.

Kanwal, J. S., J. S. Matsumura, K. Ohlemiller and N. Suga (1994) Acoustic elements and syntax in communication sounds emitted by mustached bats. J. Acous. Soc. Am. 96: 1229–1254.

Kanwal, J. S. and N. Suga (1995) Hemispheric asymmetry in the processing of calls in the auditory cortex of the mustached bat. Assoc. Res. Otolaryngol. 18: 104.

The auditory cortex obviously does much more than provide a tonotopic map and respond differentially to ipsi- and contralateral stimulation. Although the sorts of sensory processing that occur in the auditory cortex are not well understood, they are likely to be important to higher-order processing of natural sounds, especially those used for communication (Box E; see also Chapter 26). One clue about such processing comes from work in marmosets, a small neotropical primate with a complex vocal repertoire. The primary auditory cortex and belt areas of these animals are indeed organized tonotopically, but also contain neurons that are strongly responsive to spectral combinations that characterize certain vocalizations. The responses

of these neurons to the tonal stimuli do not accurately predict their responses to the spectral combinations, suggesting that, in accord with peripheral optimization, cortical processing is in part dedicated to detecting particular intraspecific vocalizations.

Another clue about the role of the primary auditory cortex in the processing of intraspecific communication sounds comes from work in echolocating bats. Consistent with the essential role that echolocation plays in the survival of these crepuscular animals, certain regions of the bat primary auditory cortex, like those described in the MGC, are tuned in a systematic manner to the delays between frequency modulated pulses and their echoes, thus providing information about target distance and velocity. These delay-tuned neurons can exhibit highly specific responses to intraspecific communication calls, suggesting that the same cortical neurons can serve these two distinct auditory functions (see Box E). Evidently the general ability of the mammalian auditory cortex to detect certain spectral and temporal combinations of natural sounds has been exploited in bats to serve sonar-mediated navigation, yielding these dual function neurons.

Many of the dually specialized neurons are categorized as "combination-sensitive" neurons, i.e., neurons that show a nonlinear increase in their response magnitude when presented with a combination of tones and/or noise bands in comparison to the total magnitude of the response elicited by presenting each sound element separately. Combination-sensitive neurons are tuned to more than one frequency and are specialized to recognize complex species-specific sounds and extract information that is critical for survival. This sensitivity to combinations of simple sound elements appears to be a universal property of neurons for the perception of complex sounds by many animal species, such as frogs, birds bats and nonhuman primates. Therefore, combination-sensitive neurons most likely partake in the recognition of complex sounds in the human auditory cortex as well.

Sounds that are especially important for intraspecific communication often have a highly ordered temporal structure. In humans, the best example of such time-varying signals is speech, where different phonetic sequences are perceived as distinct syllables and words (see Box A in Chapter 26). Behavioral studies in cats and monkeys show that the auditory cortex is especially important for processing temporal sequences of sound. If the auditory cortex is ablated in these animals, they lose the ability to discriminate between two complex sounds that have the same frequency components but which differ in temporal sequence. Thus, without the auditory cortex, monkeys cannot discriminate one conspecific communication sound from another. The physiological basis of such temporal sensitivity likely requires neurons that are sensitive to time-varying cues in communication sounds. Indeed, electrophysiological recordings from the primary auditory cortex of both marmosets and bats show that some neurons that respond to intraspecific communication sounds do not respond as strongly when the sounds are played in reverse, indicating sensitivity to the sounds' temporal features. Studies of human patients with bilateral damage to the auditory cortex also reveal severe problems in processing the temporal order of sounds. It seems likely, therefore, that specific regions of the human auditory cortex are specialized for processing elementary speech sounds, as well as other temporally complex acoustical signals, such as music (Box B). Thus, Wernicke's area, which is critical to the comprehension of human language, lies within the secondary auditory area (Figure 12.15; see also Chapter 26).

Summary

Sound waves are transmitted via the external and middle ear to the cochlea of the inner ear, which exhibits a traveling wave when stimulated. For high-frequency sounds, the amplitude of the traveling wave reaches a maximum at the base of the cochlea; for low-frequency sounds, the traveling wave reaches a maximum at the apical end. The associated motions of the basilar membrane are transduced primarily by the inner hair cells, while the basilar membrane motion is itself actively modulated by the outer hair cells. Damage to the outer or middle ear results in conductive hearing loss, while hair cell damage results in a sensorineural hearing deficit. The tonotopic organization of the cochlea is retained at all levels of the central auditory system. Projections from the cochlea travel via the eighth nerve to the three main divisions of the cochlear nucleus. The targets of the cochlear nucleus neurons include the superior olivary complex and nuclei of the lateral lemniscus, where the binaural cues for sound localization are processed. The inferior colliculus is the target of nearly all of the auditory pathways in the lower brainstem and carries out important integrative functions, such as processing sound frequencies and integrating the cues for localizing sound in space. The primary auditory cortex, which is also organized tonotopically, is essential for basic auditory functions, such as frequency discrimination and sound localization, and also plays an important role in processing of intraspecific communication sounds. The belt areas of the auditory cortex have a less strict tonotopic organization and also process complex sounds, such as those that mediate communication. In the human brain, the major speech comprehension areas are located in the zone immediately adjacent to the auditory cortex.

Additional Reading

Reviews

Corey, D. P. and A. J. Hudspeth (1979) Ionic basis of the receptor potential in a vertebrate hair cell. Nature 281: 675–677.

Corey, D.P. (1999) Ion channel defects in hereditary hearing loss. Neuron. 22(2):217-9.

Dallos, P. (1992) The active cochlea. J. Neurosci. 12: 4575–4585.

Garcia-Anoveros, J. and D. P. Corey (1997) The molecules of mechanosensation. Ann. Rev. Neurosci. 20: 567–597.

Heffner, H. E. and R. S. Heffner (1990) Role of primate auditory cortex in hearing. In *Comparative Perception, Volume II: Complex Signals*. W. C. Stebbins and M. A. Berkley (eds.). New York: John Wiley.

Hudspeth, A. J. (1997) How hearing happens. Neuron 19: 947–950.

Hudspeth, A. J. (2000) Hearing and deafness. Neurobiol. Dis. 7: 511–514.

Hudspeth, A. J. and M. Konishi (2000) Auditory neuroscience: Development, transduction, and integration. Proc. Natl. Acad. Sci. USA 97: 11690–11691.

Hudspeth, A. J., Y. Choe, A. D. Mehta and P. Martin (2000) Putting ion channels to work: Mechanoelectrical transduction, adaptation, and amplification by hair cells. Proc. Natl. Acad. Sci. USA 97: 11765-11772.

Kiang, N. Y. S. (1984) Peripheral neural processing of auditory information. In *Handbook of Physiology*, Section 1: *The Nervous System*, Volume III. *Sensory Processes*, Part 2. J. M. Brookhart, V. B. Mountcastle, I. Darian-Smith and S. R. Geiger (eds.). Bethesda, MD: American Physiological Society.

Neff, W. D., I. T. Diamond and J. H. Casseday (1975) Behavioral studies of auditory discrimination. In *Handbook of Sensory Physiology*, Volumes V–II. W. D. Keidel and W. D. Neff (eds.). Berlin: Springer-Verlag.

Nelken, I. (2002) Feature detection by the auditory cortex. In *Integrative Functions in the Mammalian Auditory Pathway, Springer Handbook of Auditory Research, Volume 15.* D. Oertel, R. Fay and A. N. Popper (eds.). New York: Springer-Verlag, pp. 358–416.

Suga, N. (1990) Biosonar and neural computation in bats. Sci. Am. 262 (June): 60–68.

Important Original Papers

Barbour, D. L. and X. Wang (2003) Contrast tuning in auditory cortex. Science. 299: 1073–1075.

Crawford, A. C. and R. Fettiplace (1981) An electrical tuning mechanism in turtle cochlear hair cells. J. Physiol. 312: 377–412.

Fitzpatrick, D. C., J. S. Kanwal, J. A. Butman and N. Suga (1993) Combination-sensitive neurons in the primary auditory cortex of the mustached bat. J. Neurosci. 13: 931–940.

Corey, D. P. and A. J. Hudspeth (1979) Ionic basis of the receptor potential in a vertebrate hair cell. Nature 281: 675–677.

Middlebrooks, J. C., A. E. Clock, L. Xu and D. M. Green (1994) A panoramic code for sound location by cortical neurons. Science 264: 842–844.

Knudsen, E. I. and M. Konishi (1978) A neural map of auditory space in the owl. Science 200: 795–797.

Jeffress, L. A. (1948) A place theory of sound localization. J. Comp. Physiol. Psychol. 41: 35–39.

NELKEN, I., Y. ROTMAN AND O. BAR YOSEF (1999) Responses of auditory-cortex neurons to structural features of natural sounds. Nature 397: 154–157.

SUGA, N., W. E. O'NEILL AND T. MANABE (1978) Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat. Science 200: 778–781.

VON BÉKÉSY, G. (1960) *Experiments in Hearing*. New York: McGraw-Hill. (A collection of von Békésy's original papers.)

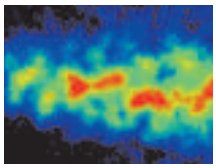
Books

PICKLES, J. O. (1988) *An Introduction to the Physiology of Hearing*. London: Academic Press.

YOST, W. A. AND G. GOUREVITCH (EDS.) (1987) *Directional Hearing*. Berlin: Springer Verlag.

YOST, W. A. AND D. W. NIELSEN (1985) *Fundamentals of Hearing*. Fort Worth: Holt, Rinehart and Winston.

Chapter 13



The Vestibular System

Overview

The vestibular system has important sensory functions, contributing to the perception of self-motion, head position, and spatial orientation relative to gravity. It also serves important motor functions, helping to stabilize gaze, head, and posture. The peripheral portion of the vestibular system includes inner ear structures that function as miniaturized accelerometers and inertial guidance devices, continually reporting information about the motions and position of the head and body to integrative centers in the brainstem, cerebellum, and somatic sensory cortices. The central portion of the system includes the vestibular nuclei, which make extensive connections with brainstem and cerebellar structures. The vestibular nuclei also directly innervate motor neurons controlling extraocular, cervical, and postural muscles. This motor output is especially important to stabilization of gaze, head orientation, and posture during movement. Although we are normally unaware of its functioning, the vestibular system is a key component in postural reflexes and eye movements. Balance, gaze stabilization during head movement, and sense of orientation in space are all adversely affected if the system is damaged. These manifestations of vestibular damage are especially important in the evaluation of brainstem injury. Because the circuitry of the vestibular system extends through a large part of the brainstem, simple clinical tests of vestibular function can be performed to determine brainstem involvement, even on comatose patients.

The Vestibular Labyrinth

The main peripheral component of the vestibular system is an elaborate set of interconnected chambers—the **labyrinth**—that has much in common, and is in fact continuous with, the cochlea (see Chapter 12). Like the cochlea, the labyrinth is derived from the otic placode of the embryo, and it uses the same specialized set of sensory cells—hair cells—to transduce physical motion into neural impulses. In the cochlea, the motion is due to airborne sounds; in the labyrinth, the motions transduced arise from head movements, inertial effects due to gravity, and ground-borne vibrations (Box A).

The labyrinth is buried deep in the temporal bone and consists of the two **otolith organs** (the **utricle** and **saccule**) and three **semicircular canals** (Figure 13.1). The elaborate and tortuous architecture of these components explains why this part of the vestibular system is called the labyrinth. The utricle and saccule are specialized primarily to respond to *linear accelerations* of the head and *static head position relative to the graviational axis*, whereas the semicircular canals, as their shapes suggest, are specialized for responding to *rotational accelerations* of the head.

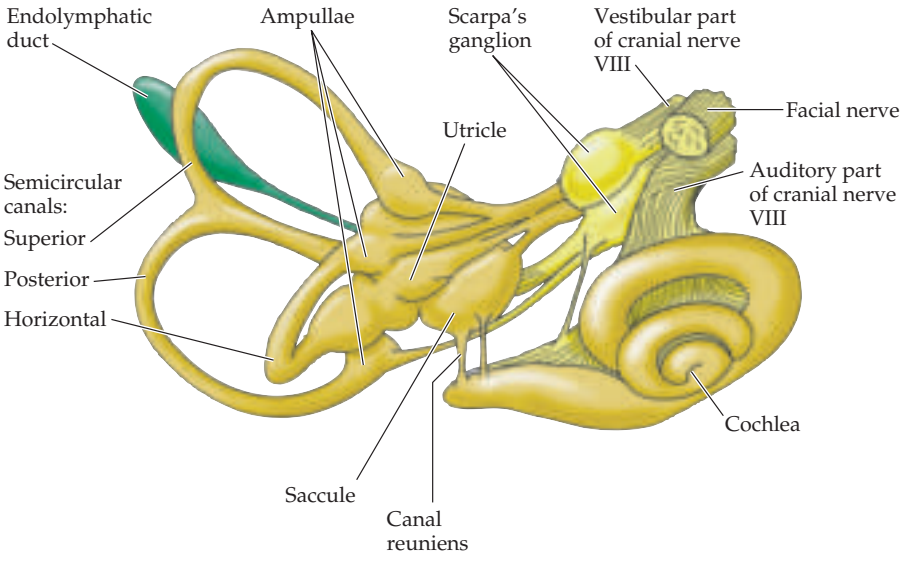


Figure 13.1 The labyrinth and its innervation. The vestibular and auditory portions of the eighth nerve are shown; the small connection from the vestibular nerve to the cochlea contains auditory efferent fibers. General orientation in head is shown in Figure 12.3; see also Figure 13.8.

The intimate relationship between the cochlea and the labyrinth goes beyond their common embryonic origin. Indeed, the cochlear and vestibular spaces are actually joined (see Figure 13.1), and the specialized ionic environments of the vestibular end organ parallel those of the cochlea. The membranous sacs within the bone are filled with fluid (endolymph) and are collectively called the membranous labyrinth. The endolymph (like the cochlear endolymph) is similar to intracellular solutions in that it is high in K^+ and low in Na^+. Between the bony walls (the osseous labyrinth) and the membranous labyrinth is another fluid, the perilymph, which is similar in composition to cerebrospinal fluid (i.e., low in K^+ and high in Na^+; see Chapter 12).

The vestibular hair cells are located in the utricle and saccule and in three juglike swellings called **ampullae**, located at the base of the semicircular canals next to the utricle. Within each ampulla, the vestibular hair cells extend their hair bundles into the endolymph of the membranous labyrinth. As in the cochlea, tight junctions seal the apical surfaces of the vestibular hair cells, ensuring that endolymph selectively bathes the hair cell bundle while remaining separate from the perilymph surrounding the basal portion of the hair cell.

Vestibular Hair Cells

The vestibular hair cells, which like cochlear hair cells transduce minute displacements into behaviorally relevant receptor potentials, provide the basis for vestibular function. Vestibular and auditory hair cells are quite similar; a detailed description of hair cell structure and function has already been given in Chapter 12. As in the case of auditory hair cells, movement of the stereocilia toward the kinocilium in the vestibular end organs opens mechanically gated transduction channels located at the tips of the stereocilia, depolarizing the hair cell and causing neurotransmitter release onto (and excitation of) the vestibular nerve fibers. Movement of the stereocilia in the direction away from the kinocilium closes the channels, hyperpolarizing the hair cell and thus reducing vestibular nerve activity. The biphasic nature of the receptor potential means that some transduction channels are open in the absence of stimulation, with the result that hair cells tonically release

transmitter, thereby generating considerable spontaneous activity in vestibular nerve fibers (see Figure 13.6). One consequence of these spontaneous action potentials is that the firing rates of vestibular fibers can increase or decrease in a manner that faithfully mimics the receptor potentials produced by the hair cells (Box B).

Importantly, the hair cell bundles in each vestibular organ have specific orientations (Figure 13.2). As a result, the organ as a whole is responsive to displacements in all directions. In a given semicircular canal, the hair cells in the ampulla are all polarized in the same direction. In the utricle and saccule, a specialized area called the **striola** divides the hair cells into two populations with opposing polarities (Figure 13.2C; see also Figure 13.4C). The directional polarization of the receptor surfaces is a basic principle of organization in the vestibular system, as will become apparent in the following descriptions of the individual vestibular organs.

The Otolith Organs: The Utricle and Saccule

Displacements and linear accelerations of the head, such as those induced by tilting or translational movements (see Box A), are detected by the two otolith organs: the saccule and the utricle. Both of these organs contain a

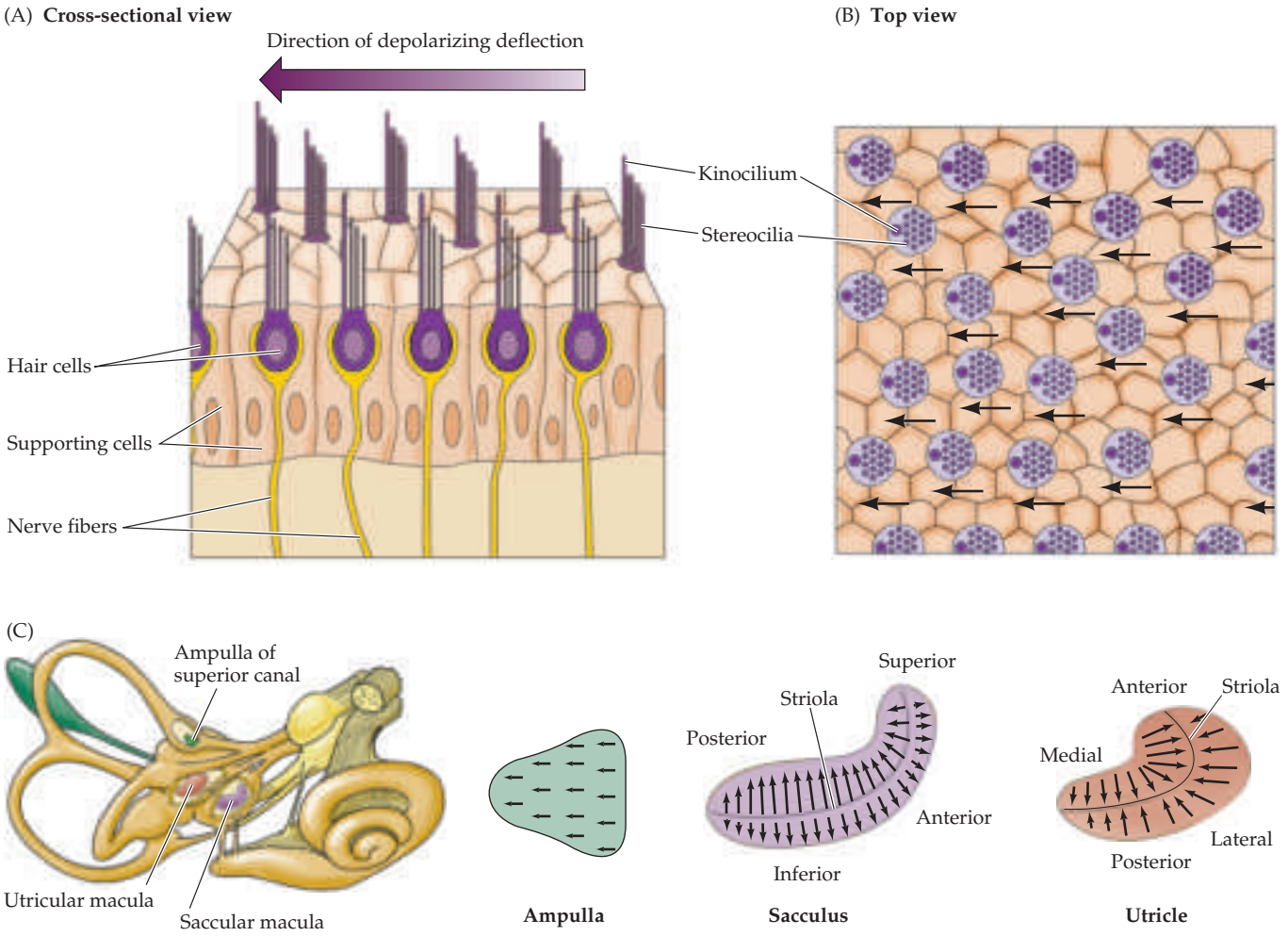
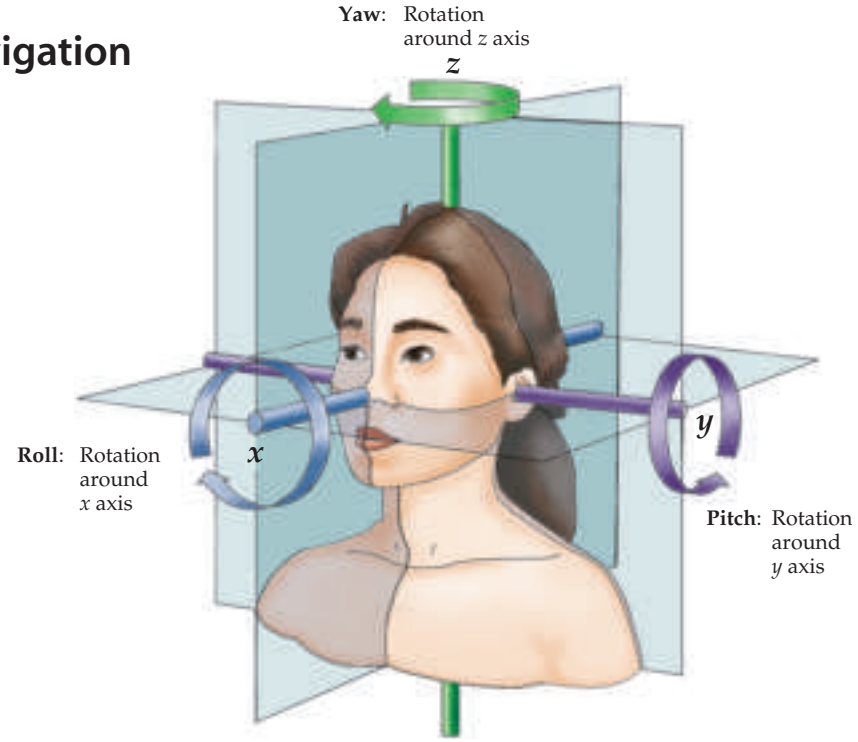


Figure 13.2 The morphological polarization of vestibular hair cells and the polarization maps of the vestibular organs. (A) A cross section of hair cells shows that the kinocilia of a group of hair cells are all located on the same side of the hair cell. The arrow indicates the direction of deflection that depolarizes the hair cell. (B) View looking down on the hair bundles. (C) In the ampulla located at the base of each semicircular canal, the hair bundles are oriented in the same direction. In the sacculus and utricle, the striola divides the hair cells into populations with opposing hair bundle polarities.

Box A

A Primer on Vestibular Navigation

The function of the vestibular system can be simplified by remembering some basic terminology of classical mechanics. All bodies moving in a three-dimensional framework have six degrees of freedom: three of these are translational and three are rotational. The translational elements refer to linear movements in the *x*, *y*, and *z* axes (the horizontal and vertical planes). Translational motion in these planes (linear acceleration and static displacement of the head) is the primary concern of the otolith organs. The three degrees of rotational freedom refer to a body's rotation relative to the *x*, *y*, and *z* axes and are commonly referred to as *roll*, *pitch*, and *yaw*. The semicircular canals are primarily responsible for sensing rotational accelerations around these three axes.



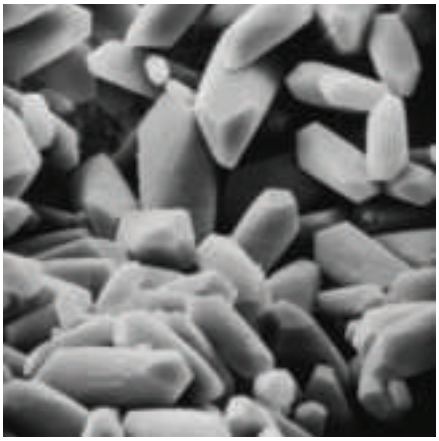


Figure 13.3 Scanning electron micrograph of calcium carbonate crystals (otoconia) in the utricular macula of the cat. Each crystal is about 50 mm long. (From Lindeman, 1973.)

sensory epithelium, the **macula**, which consists of hair cells and associated supporting cells. Overlying the hair cells and their hair bundles is a gelatinous layer; above this layer is a fibrous structure, the **otolithic membrane**, in which are embedded crystals of calcium carbonate called **otoconia** (Figures 13.3 and 13.4A). The crystals give the otolith organs their name (*otolith* is Greek for "ear stones"). The otoconia make the otolithic membrane considerably heavier than the structures and fluids surrounding it; thus, when the head tilts, gravity causes the membrane to shift relative to the sensory epithelium (Figure 13.4B). The resulting shearing motion between the otolithic membrane and the macula displaces the hair bundles, which are embedded in the lower, gelatinous surface of the membrane. This displacement of the hair bundles generates a receptor potential in the hair cells. A shearing motion between the macula and the otolithic membrane also occurs when the head undergoes linear accelerations (see Figure 13.5); the greater relative mass of the otolithic membrane causes it to lag behind the macula temporarily, leading to transient displacement of the hair bundle.

The similar effects exerted on otolithic hair cells by certain head tilts and linear accelerations would be expected to render these different stimuli perceptually equivalent when visual feedback is absent, as occurs in the dark or when the eyes are closed. Nevertheless, evidence suggests that subjects can discriminate between these two stimulus categories, apparently through combined activity of the otolith organs and the semicircular canals.

As already mentioned, the orientation of the hair cell bundles is organized relative to the striola, which demarcates the overlying layer of otoco-

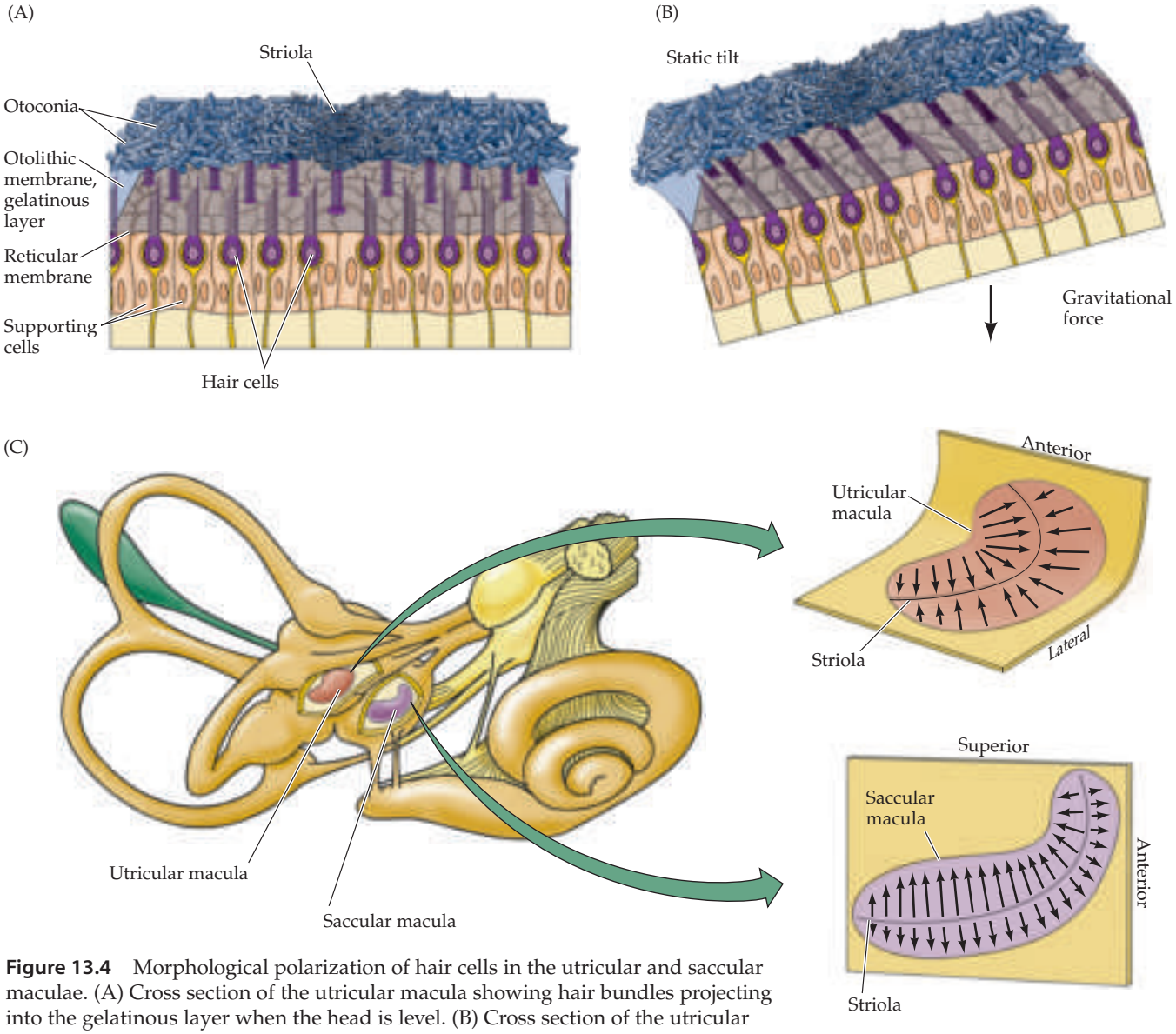


Figure 13.4 Morphological polarization of hair cells in the utricular and saccular maculae. (A) Cross section of the utricular macula showing hair bundles projecting into the gelatinous layer when the head is level. (B) Cross section of the utricular macula when the head is tilted. (C) Orientation of the utricular and saccular maculae in the head; arrows show orientation of the kinocilia, as in Figure 13.2. The *saccules* on either side are oriented more or less vertically, and the *utricles* more or less horizontally. The striola is a structural landmark consisting of small otoconia arranged in a narrow trench that divides each otolith organ. In the utricular macula, the kinocilia are directed toward the striola. In the saccular macula, the kinocilia point away from the striola. Note that, given the utricle and sacculus on both sides of the body, there is a continuous representation of all directions of body movement.

nia (see Figure 13.4A). The striola forms an axis of mirror symmetry such that hair cells on opposite sides of the striola have opposing morphological polarizations. Thus, a tilt along the axis of the striola will excite the hair cells on one side while inhibiting the hair cells on the other side. The saccular macula is oriented vertically and the utricular macula horizontally, with a continuous variation in the morphological polarization of the hair cells

Box B

Adaptation and Tuning of Vestibular Hair Cells

Hair Cell Adaptation

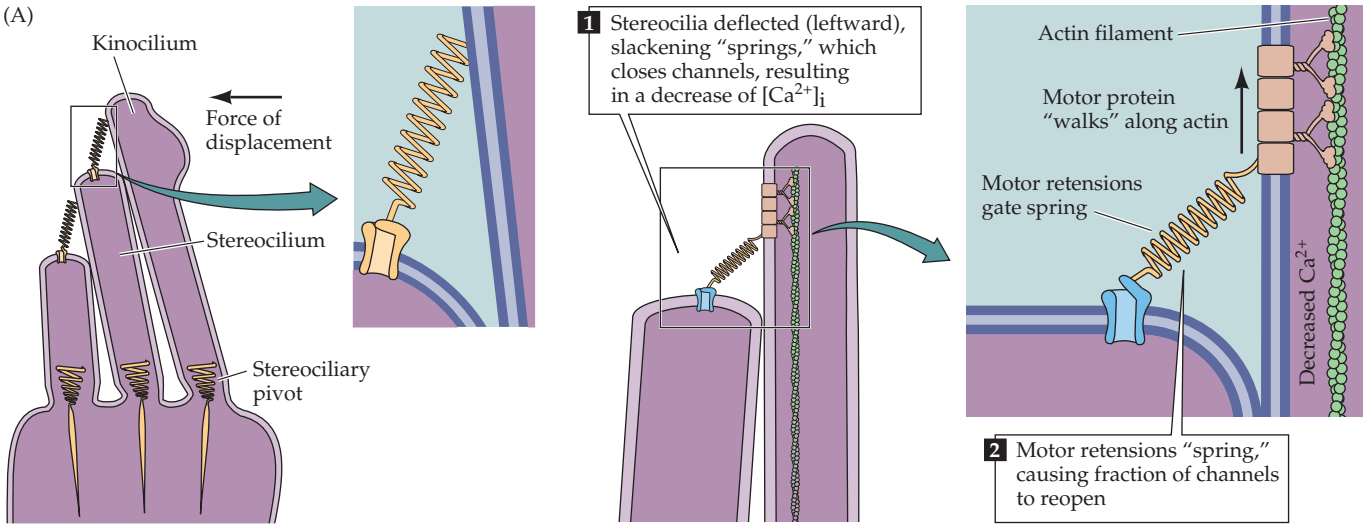
The minuscule movement of the hair bundle at sensory threshold has been compared to the displacement of the top of the Eiffel Tower by a thumb's breadth! Despite its great sensitivity, the hair cell can adapt quickly and continuously to static displacements of the hair bundle caused by large movements. Such adjustments are especially useful in the otolith organs, where adaptation permits hair cells to maintain sensitivity to small linear and angular accelerations of the head despite the constant input from gravitational forces that are over a million times greater. In other receptor cells, such as photoreceptors, adaptation is accomplished by regulating the second messenger cascade induced by the initial transduction event. The hair cell has to depend on a different strategy, however, because there is no second messenger system between the initial transduction event and the subsequent receptor potential (as might be expected for receptors that respond so rapidly).

Adaptation occurs in both directions in which the hair bundle displacement generates a receptor potential, albeit at different rates for each direction. When the hair bundle is pushed toward the kinocilium, tension is initially increased in the gating spring. During adaptation, tension decreases back to the resting level, perhaps because one end of the gating spring repositions itself along the shank of the stereocilium. When the hair bundle is displaced in the opposite direction, away from the kinocilium, tension in the spring initially decreases; adaptation then involves an increase in spring tension. One theory is that a calcium-regulated motor such as a myosin ATPase climbs along actin filaments in the stereocilium and actively resets the tension in the transduction spring. During sustained depolarization, some Ca^{2+} enters through the transduction channel, along with K^+. Ca^{2+} then causes the motor to spend a greater fraction of its time unbound from the actin, resulting in slippage of the spring down the side of the stereocilium. During sustained hyperpolarization (Figure A), Ca^{2+} levels drop below normal resting levels and the motor spends more of its time bound to the actin, thus climbing up the actin filaments and increasing the spring tension. As tension increases, some of the previously closed transduction channels open, admitting Ca^{2+} and thus slowing the motor's progress until a balance is struck between the climbing and slipping of the motor. In support of this model, when internal Ca^{2+} is reduced artificially, spring tension increases. This model of hair cell adaptation presents an elegant molecular solution to the regulation of a mechanical process.

Electrical Tuning

Although mechanical tuning plays an important role in generating frequency selectivity in the cochlea, there are other mechanisms that contribute to this process in vestibular and auditory nerve cells. These other tuning mechanisms are especially important in the otolith organs, where, unlike the cochlea, there are no

(A) Adaptation is explained in the gating spring model by adjustment of the insertion point of tips links. Movement of the insertion point up or down the shank of the stereocilium, perhaps driven by a Ca^{2+}-dependent protein motor, can continually adjust the resting tension of the tip link. (After Hudspeth and Gillespie, 1994.)



obvious macromechanical resonances to selectively filter and/or enhance biologically relevant movements. One such mechanism is an electrical resonance displayed by hair cells in response to depolarization: The membrane potential of a hair cell undergoes damped sinusoidal oscillations at a specific frequency in response to the injection of depolarizing current pulses (Figure B).

The ionic mechanism of this process involves two major types of ion channels located in the membrane of the hair cell soma. The first of these is a voltage-activated Ca^{2+} conductance, which lets Ca^{2+} into the cell soma in response to depolarization, such as that generated by the transduction current. The second is a Ca^{2+}-activated K^+ conductance, which is triggered by the rise in internal Ca^{2+} concentration. These two currents produce an interplay of depolarization and repolarization that results in electrical resonance (Figure C). Activation of the hair cell's calcium-activated K^+ conductance occurs 10 to 100 times faster than that of similar currents in other cells. Such rapid kinetics allow this conductance to generate an electrical response that usually requires the fast properties of a voltage-gated channel.

Although a hair cell responds to hair bundle movement over a wide range of frequencies, the resultant receptor potential is largest at the frequency of electrical resonance. The resonance frequency represents the characteristic frequency of the hair cell, and transduction at that frequency will be most efficient. This electrical resonance has important implications for structures like the utricle and sacculus, which may encode a range of characteristic frequencies based on the different resonance frequencies of their constituent hair cells. Thus, electrical tuning in the otolith organs can generate enhanced tuning to biologically relevant frequencies of stimulation, even in the absence of macromechanical resonances within these structures.

References

Assad, J. A. and D. P. Corey (1992) An active motor model for adaptation by vertebrate hair cells. J. Neurosci. 12: 3291–3309.

Crawford, A. C. and R. Fettiplace (1981) An electrical tuning mechanism in turtle cochlear hair cells. J. Physiol. 312: 377–412.

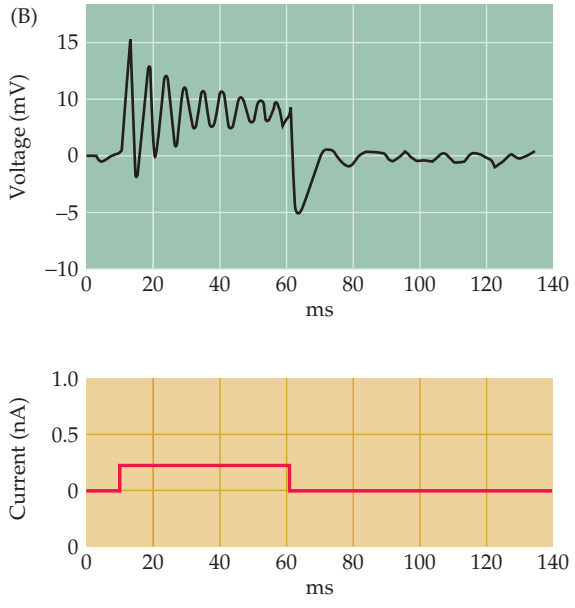
Hudspeth, A. J. (1985) The cellular basis of hearing: The biophysics of hair cells. Science 230: 745–752.

Hudspeth, A. J. and P. G. Gillespie (1994) Pulling strings to tune transduction: Adaptation by hair cells. Neuron 12: 1–9.

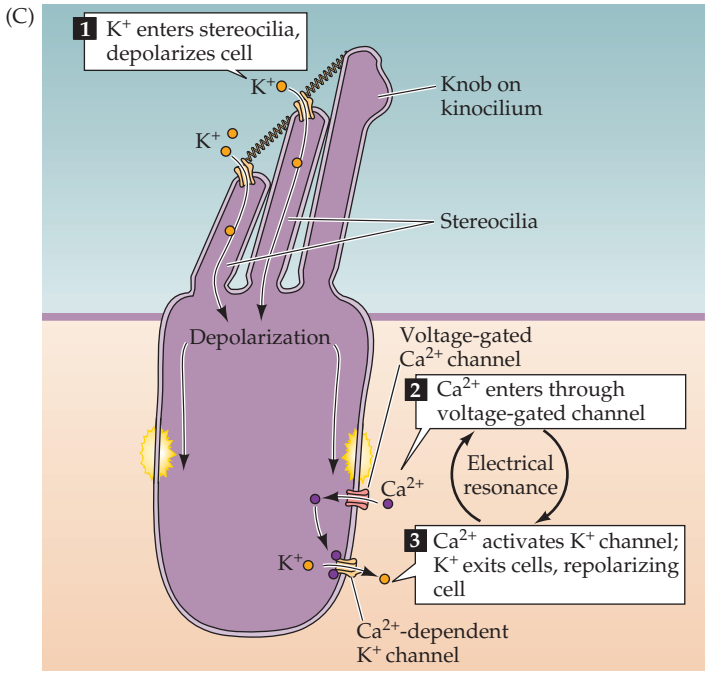
Lewis, R. S. and A. J. Hudspeth (1988) A model for electrical resonance and frequency tuning in saccular hair cells of the bull-frog, *Rana catesbeiana*. J. Physiol. 400: 275–297.

Lewis, R. S. and A. J. Hudspeth (1983) Voltage- and ion-dependent conductances in solitary vertebrate hair cells. Nature 304: 538–541.

Shepherd, G. M. G. and D. P. Corey (1994) The extent of adaptation in bullfrog saccular hair cells. J. Neurosci. 14: 6217–6229.



(B) Voltage oscillations (upper trace) in an isolated hair cell in response to a depolarizing current injection (lower trace). (After Lewis and Hudspeth, 1983.)



(C) Proposed ionic basis for electrical resonance in hair cells. (After Hudspeth, 1985.)

located in each macula (as shown in Figure 13.4C, where the arrows indicate the direction of movement that produces excitation). Inspection of the excitatory orientations in the maculae indicates that the utricle responds to movements of the head in the horizontal plane, such as sideways head tilts and rapid lateral displacements, whereas the saccule responds to movements in the vertical plane (up–down and forward–backward movements in the sagittal plane).

Note that the saccular and utricular maculae on one side of the head are mirror images of those on the other side. Thus, a tilt of the head to one side has opposite effects on corresponding hair cells of the two utricular maculae. This concept is important in understanding how the central connections of the vestibular periphery mediate the interaction of inputs from the two sides of the head (see the next section).

How Otolith Neurons Sense Linear Forces

The structure of the otolith organs enables them to sense both static displacements, as would be caused by tilting the head relative to the gravitational axis, and transient displacements caused by translational movements of the head. The mass of the otolithic membrane relative to the surrounding endolymph, as well as the otolithic membrane's physical uncoupling from the underlying macula, means that hair bundle displacement will occur transiently in response to linear accelerations and tonically in response to tilting of the head. Therefore, both tonic and transient information can be conveyed by these sense organs. Figure 13.5 illustrates some of the forces produced by head tilt and linear accelerations on the utricular macula.

These properties of hair cells are reflected in the responses of the vestibular nerve fibers that innervate the otolith organs. The nerve fibers have a

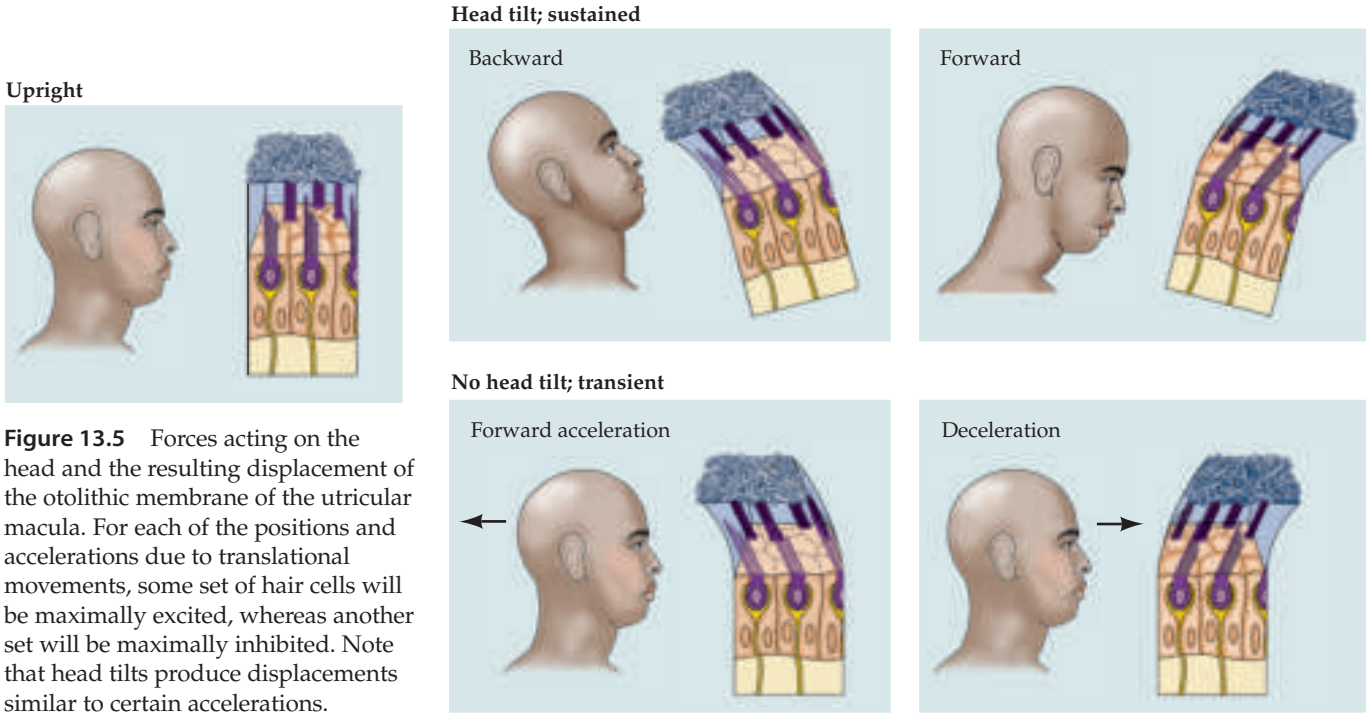


Figure 13.5 Forces acting on the head and the resulting displacement of the otolithic membrane of the utricular macula. For each of the positions and accelerations due to translational movements, some set of hair cells will be maximally excited, whereas another set will be maximally inhibited. Note that head tilts produce displacements similar to certain accelerations.

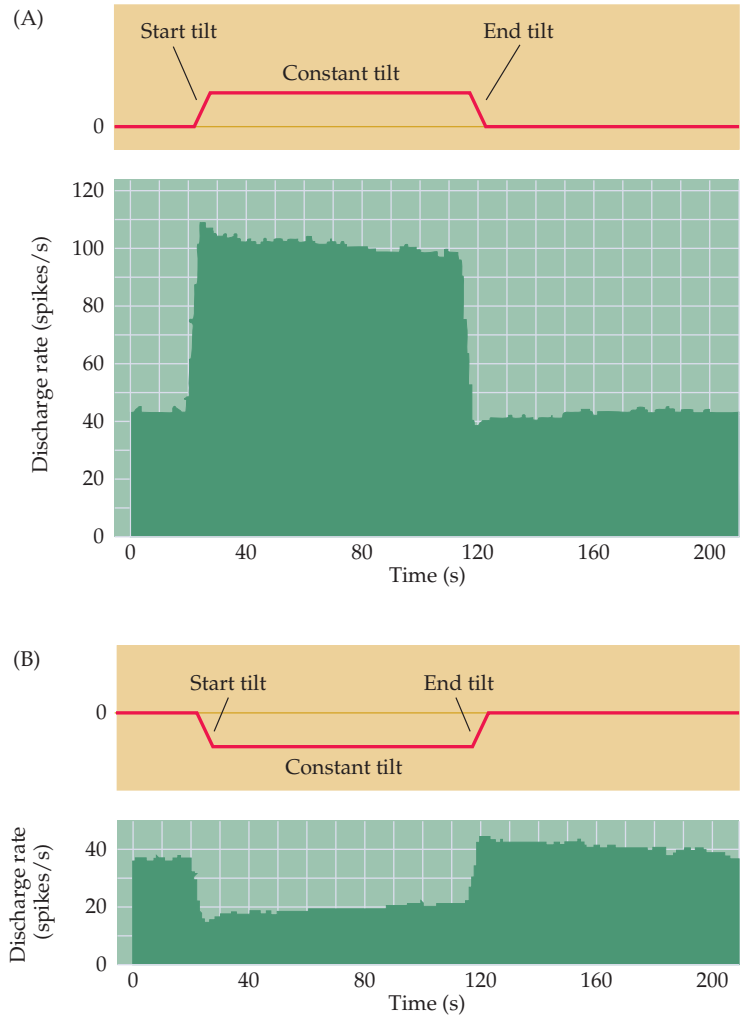


Figure 13.6 Response of a vestibular nerve axon from an otolith organ (the utricle in this example). (A) The stimulus (top) is a change in head tilt. The spike histogram shows the neuron's response to tilting in a particular direction. (B) A response of the same fiber to tilting in the opposite direction. (After Goldberg and Fernandez, 1976.)

steady and relatively high firing rate when the head is upright. The change in firing rate in response to a given movement can be either sustained or transient, thereby signaling either absolute head position or linear acceleration. An example of the sustained response of a vestibular nerve fiber innervating the utricle is shown in Figure 13.6. The responses were recorded from axons in a monkey seated in a chair that could be tilted for several seconds to produce a steady force. Prior to the tilt, the axon has a high firing rate, which increases or decreases depending on the direction of the tilt. Notice also that the response remains at a high level as long as the tilting force remains constant; thus, such neurons faithfully encode the static force being applied to the head (Figure 13.6A). When the head is returned to the original position, the firing level of the neurons returns to baseline value. Conversely, when the tilt is in the opposite direction, the neurons respond by decreasing their firing rate below the resting level (Figure 13.6B) and remain depressed as long as the static force continues. In a similar fashion, transient increases or decreases in firing rate from spontaneous levels signal the direction of linear accelerations of the head.

The range of orientations of hair bundles within the otolith organs enables them to transmit information about linear forces in every direction

the body moves (see Figure 13.4C). The utricle, which is primarily concerned with motion in the horizontal plane, and the saccule, which is concerned with vertical motion, combine to effectively gauge the linear forces acting on the head at any instant in three dimensions. Tilts of the head off the horizontal plane and translational movements of the head in any direction stimulate a distinct subset of hair cells in the saccular and utricular maculae, while simultaneously suppressing the responses of other hair cells in these organs. Ultimately, variations in hair cell polarity within the otolith organs produce patterns of vestibular nerve fiber activity that, at a population level, can unambiguously encode head position and the forces that influence it.

The Semicircular Canals

Whereas the otolith organs are primarily concerned with head translations and orientation with respect to gravity, the semicircular canals sense head *rotations*, arising either from self-induced movements or from angular accelerations of the head imparted by external forces. Each of the three semicircular canals has at its base a bulbous expansion called the **ampulla** (Figure 13.7), which houses the sensory epithelium, or **crista**, that contains the hair cells. The structure of the canals suggests how they detect the angular accelerations that arise through rotation of the head. The hair bundles extend out of the crista into a gelatinous mass, the **cupula**, that bridges the width of the ampulla, forming a fluid barrier through which endolymph cannot circulate. As a result, the relatively compliant cupula is distorted by movements of the endolymphatic fluid. When the head turns in the plane of one of the semicircular canals, the inertia of the endolymph produces a force across the cupula, distending it away from the direction of head movement and causing a displacement of the hair bundles within the crista (Figure 13.8A,B). In contrast, linear accelerations of the head produce equal forces on the two sides of the cupula, so the hair bundles are not displaced.

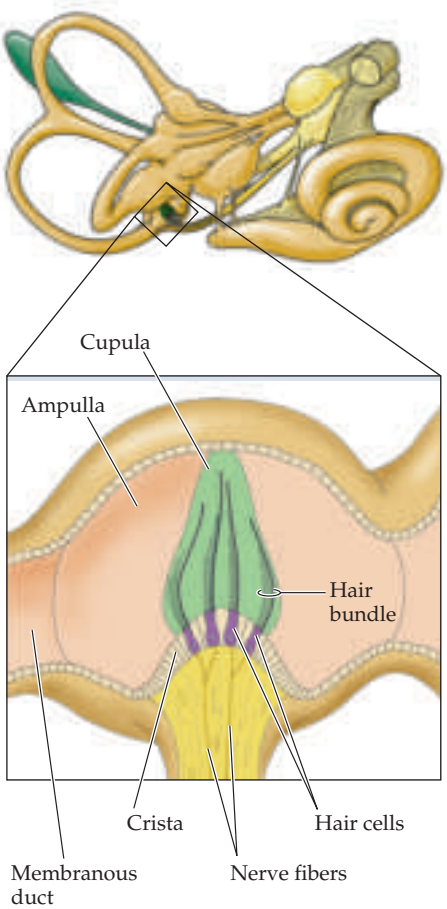


Figure 13.7 The ampulla of the posterior semicircular canal showing the crista, hair bundles, and cupula. The cupula is distorted by the fluid in the membranous canal when the head rotates.

Unlike the saccular and utricular maculae, all of the hair cells in the crista within each semicircular canal are organized with their kinocilia pointing in the same direction (see Figure 13.2C). Thus, when the cupula moves in the appropriate direction, the entire population of hair cells is depolarized and activity in all of the innervating axons increases. When the cupula moves in the opposite direction, the population is hyperpolarized and neuronal activity decreases. Deflections orthogonal to the excitatory–inhibitory direction produce little or no response.

Each semicircular canal works in concert with the partner located on the other side of the head that has its hair cells aligned oppositely. There are three such pairs: the two pairs of horizontal canals, and the superior canal on each side working with the posterior canal on the other side (Figure 13.8C). Head rotation deforms the cupula in opposing directions for the two partners, resulting in opposite changes in their firing rates (Box C). Thus, the orientation of the horizontal canals makes them selectively sensitive to rotation in the horizontal plane. More specifically, the hair cells in the canal towards which the head is turning are depolarized, while those on the other side are hyperpolarized.

For example, when the head accelerates to the left, the cupula is pushed toward the kinocilium in the left horizontal canal, and the firing rate of the relevant axons in the left vestibular nerve increases. In contrast, the cupula in the right horizontal canal is pushed away from the kinocilium, with a concomitant decrease in the firing rate of the related neurons. If the head movement is

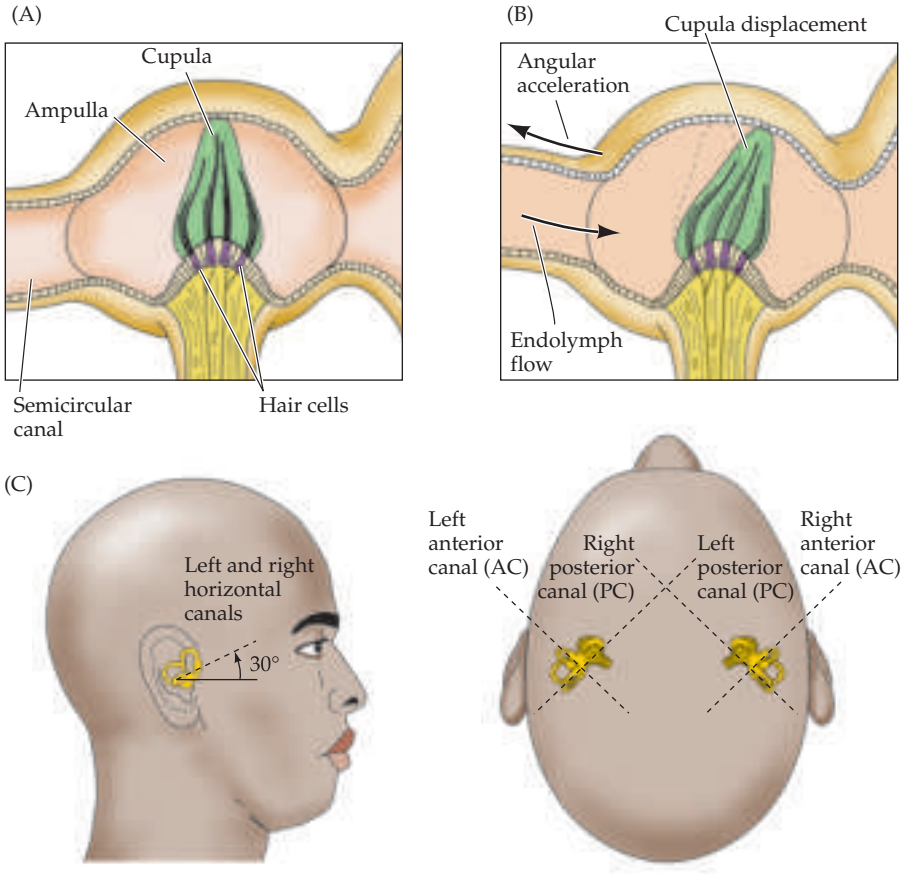


Figure 13.8 Functional organization of the semicircular canals. (A) The position of the cupula without angular acceleration. (B) Distortion of the cupula during angular acceleration. When the head is rotated in the plane of the canal (arrow outside canal), the inertia of the endolymph creates a force (arrow inside the canal) that displaces the cupula. (C) Arrangement of the canals in pairs. The two horizontal canals form a pair; the right anterior canal (AC) and the left posterior canal (PC) form a pair; and the left AC and the right PC form a pair.

to the right, the result is just the opposite. This push–pull arrangement operates for all three pairs of canals; the pair whose activity is modulated is in the plane of the rotation, and the member of the pair whose activity is increased is on the side toward which the head is turning. The net result is a system that provides information about the rotation of the head in any direction.

How Semicircular Canal Neurons Sense Angular Accelerations

Like axons that innervate the otolith organs, the vestibular fibers that innervate the semicircular canals exhibit a high level of spontaneous activity. As a result, they can transmit information by either increasing or decreasing their firing rate, thus more effectively encoding head movements (see above). The bidirectional responses of fibers innervating the hair cells of the semicircular canal have been studied by recording the axonal firing rates in a monkey's

Box C

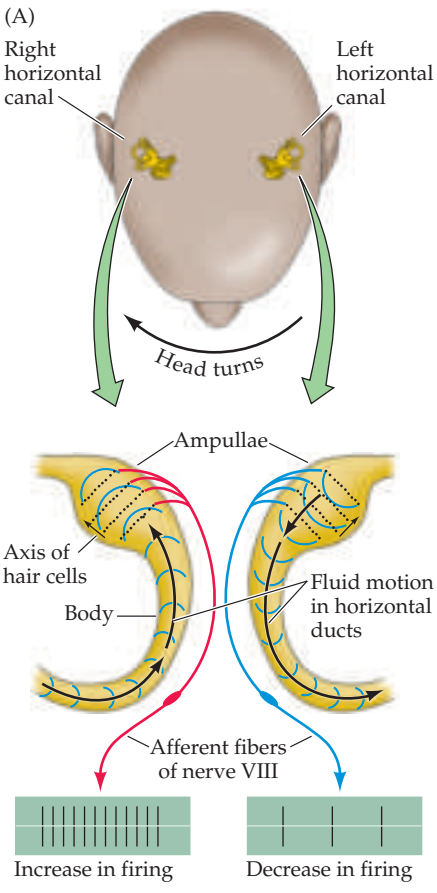
Throwing Cold Water on the Vestibular System

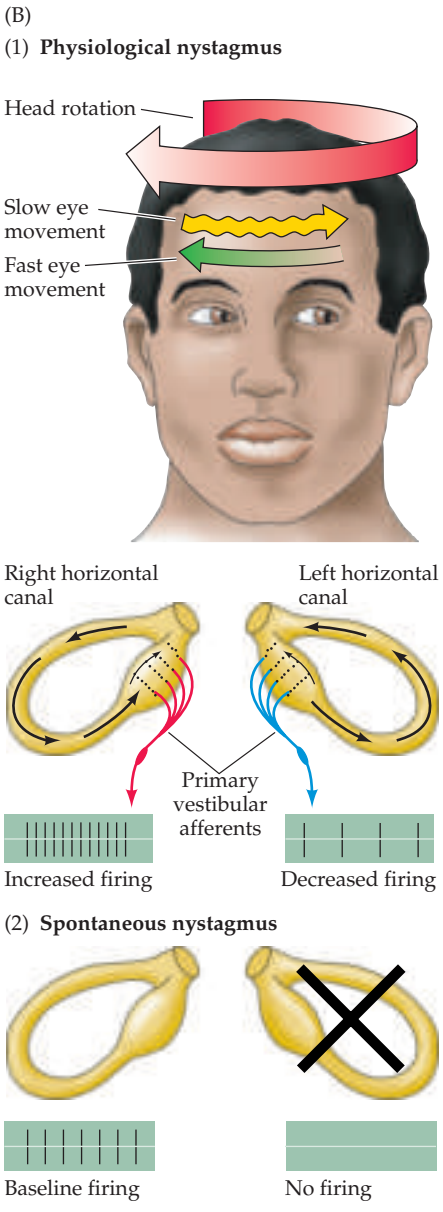
Testing the integrity of the vestibular system can indicate much about the condition of the brainstem, particularly in comatose patients.

Normally, when the head is not being rotated, the output of the nerves from the right and left sides are equal; thus, no eye movements occur. When the head is rotated in the horizontal plane, the vestibular afferent fibers on the side toward the turning motion increase their firing rate, while the afferents on the opposite side decrease their firing rate (Figures A and B). The net difference in firing rates then leads to slow movements of the eyes counter to the turning motion. This reflex response generates the slow component of a normal eye movement pattern called physiological nystagmus, which means "nodding" or oscillatory movements of the eyes (Figure B1). (The fast component is a saccade that resets the eye position; see Chapter 19.)

Pathological nystagmus can occur if there is unilateral damage to the vestibular system. In this case, the silencing of the spontaneous output from the damaged side results in an unphysiological difference in firing rate because the spontaneous discharge from the intact side remains (Figure B2). The difference in firing rates causes nystagmus, even though no head movements are being made.

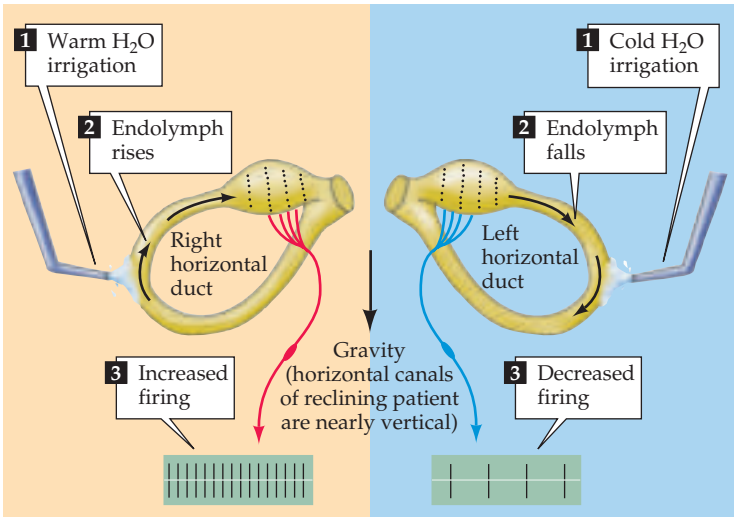
Responses to vestibular stimulation are thus useful in assessing the integrity of the brainstem in unconscious patients. If the individual is placed on his or her back and the head is elevated to about 30° above horizontal, the horizontal semicircular canals lie in an almost vertical orientation. Irrigating one ear with cold water will then lead to spontaneous eye movements because convection currents in the canal mimic rotatory head movements away from the irrigated ear (Figure C). In normal individuals, these eye movements consist of a slow movement toward the irrigated ear and a fast movement away from it. The fast movement is most readily detected by the observer, and the significance of its direction can be kept in mind by using the





(A) View looking down on the top of a person's head illustrates the fluid motion generated in the left and right horizontal canals, and the changes in vestibular nerve firing rates when the head turns to the right. (B) In normal individuals, rotating the head elicits physiological nystagmus (1), which consists of a slow eye movement counter to the direction of head turning. The slow component of the eye movements is due to the net differences in left and right vestibular nerve firing rates acting via the central circuit diagrammed in Figure 13.10. Spontaneous nystagmus (2), where the eyes move rhythmically from side to side in the absence of any head movements, occurs when one of the canals is damaged. In this situation, net differences in vestibular nerve firing rates exist even when the head is stationary because the vestibular nerve innervating the intact canal fires steadily when at rest, in contrast to a lack of activity on the damaged side.

(C)

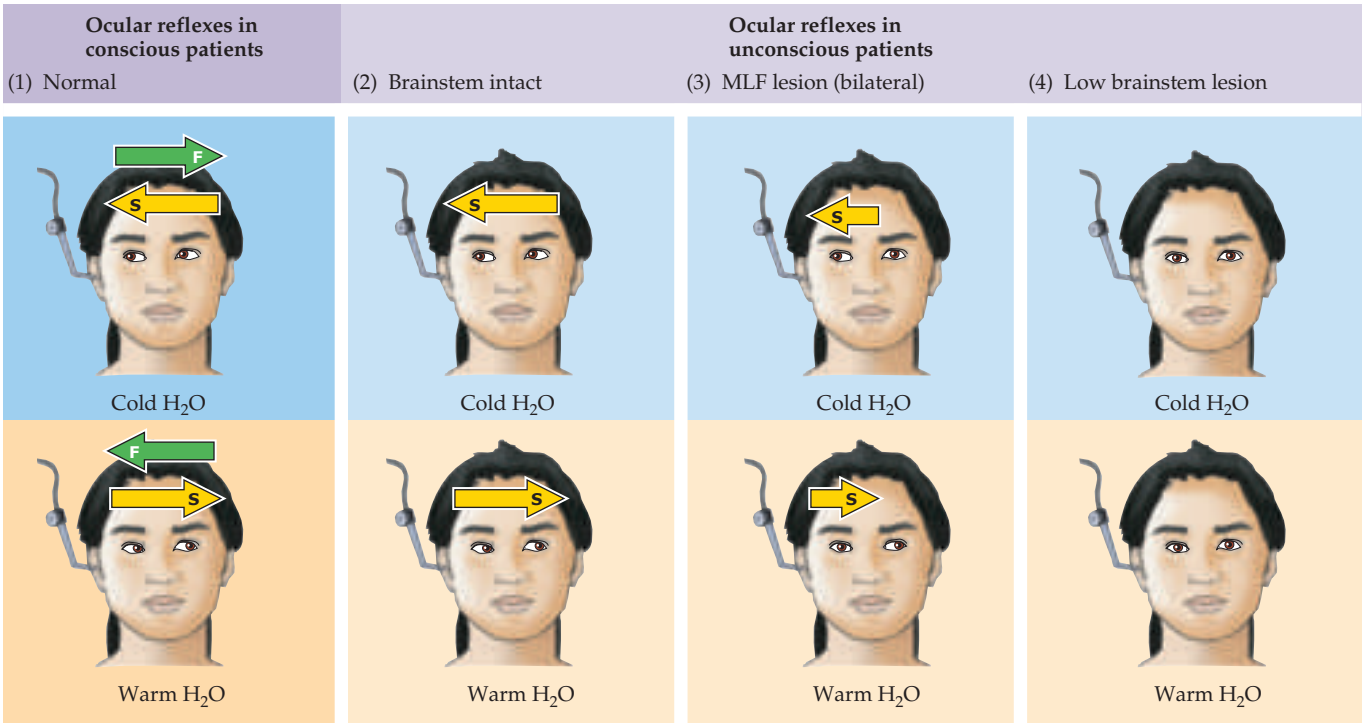


(C) Caloric testing of vestibular function is possible because irrigating an ear with water slightly warmer than body temperature generates convection currents in the canal that mimic the endolymph movement induced by turning the head to the irrigated side. Irrigation with cold water induces the opposite effect. These currents result in changes in the firing rate of the associated vestibular nerve, with an increased rate on the warmed side and a decreased rate on the chilled side. As in head rotation and spontaneous nystagmus, net differences in firing rates generate eye movements.

mnemonic COWS ("Cold Opposite, Warm Same"). This same test can also be used in unconscious patients. In patients who are comatose due to dysfunction of both cerebral hemispheres but whose brainstem is intact, saccadic movements are no longer made and the response to cold water consists of only the slow movement component of the eyes to side of the irrigated ear (Figure D). In the presence of brainstem lesions involving either the vestibular nuclei themselves, the connections from the vestibular nuclei to oculomotor nuclei (the third, fourth, or sixth cranial nerves), or the peripheral nerves exiting these nuclei, vestibular responses are abolished (or altered, depending on the severity of the lesion).

(D) Caloric testing can be used to test the function of the brainstem in an unconscious patient. The figures show eye movements resulting from cold or warm water irrigation in one ear for (1) a normal subject, and in three different conditions in an unconscious patient: (2) with the brainstem intact; (3) with a lesion of the medial longitudinal fasciculus (MLF; note that irrigation in this case results in lateral movement of the eye only on the less active side); and (4) with a low brainstem lesion (see Figure 13.10).

(D)



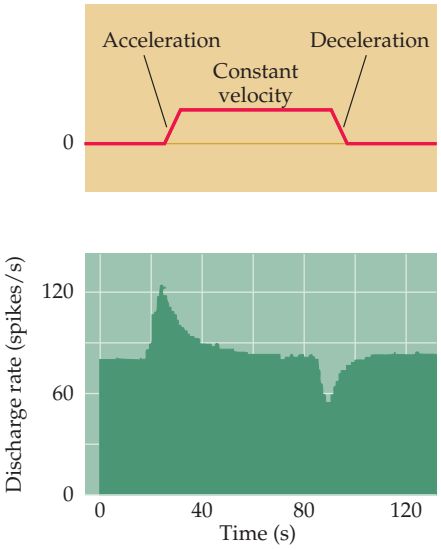


Figure 13.9 Response of a vestibular nerve axon from the semicircular canal to angular acceleration. The stimulus (top) is a rotation that first accelerates, then maintains constant velocity, and then decelerates the head. The axon increases its firing above resting level in response to the acceleration, returns to resting level during constant velocity, then decreases its firing rate below resting level during deceleration; these changes in firing rate reflect inertial effects on the displacement of the cupula. (After Goldberg and Fernandez, 1971.)

vestibular nerve. Seated in a chair, the monkey was rotated continuously in one direction during three phases: an initial period of acceleration, then a periord of several seconds at constant velocity, and finally a period of sudden deceleration to a stop (Figure 13.9). The maximum firing rates observed correspond to the period of acceleration; the maximum inhibition corresponds to the period of deceleration. During the constant-velocity phase, the response adapts so that the firing rate subsides to resting level; after the movement stops, the neuronal activity decreases transiently before returning to the resting level.

Neurons innervating paired semicircular canals have a complementary response pattern. Note that the rate of adaptation (on the order of tens of seconds) corresponds to the time it takes the cupula to return to its undistorted state (and for the hair bundles to return to their undeflected position); adaptation therefore can occur while the head is still turning, as long as a constant angular velocity is maintained. Such constant forces are rare in nature, although they are encountered on ships, airplanes, and space vehicles, where prolonged acceleratory arcs are sometimes described.

Central Pathways for Stabilizing Gaze, Head, and Posture

The vestibular end organs communicate via the vestibular branch of cranial nerve VIII with targets in the brainstem and the cerebellum that process much of the information necessary to compute head position and motion. As with the cochlear nerve, the vestibular nerves arise from a population of bipolar neurons, the cell bodies of which in this instance reside in the **vestibular nerve ganglion** (also called **Scarpa's ganglion**; see Figure 13.1). The distal processes of these cells innervate the semicircular canals and the otolith organs, while the central processes project via the vestibular portion of cranial nerve VIII to the **vestibular nuclei** (and also directly to the cerebellum; Figure 13.10). The vestibular nuclei are important centers of integration, receiving input from the vestibular nuclei of the opposite side, as well as from the cerebellum and the visual and somatic sensory systems. Because vestibular and auditory fibers run together in the eighth nerve, damage to this structure often results in both auditory and vestibular disturbances.

The central projections of the vestibular system participate in three major classes of reflexes: (1) helping to maintain equilibrium and gaze during movement, (2) maintaining posture, and (3) maintaining muscle tone. The first of these reflexes helps coordinate head and eye movements to keep gaze fixated on objects of interest during movements (other functions include protective or escape reactions; see Box D). The **vestibulo-ocular reflex** (**VOR**) in particular is a mechanism for producing eye movements that counter head movements, thus permitting the gaze to remain fixed on a particular point (Box C; see also Chapter 19). For example, activity in the left horizontal canal induced by leftward rotary acceleration of the head excites neurons in the left vestibular nucleus and results in compensatory eye movements to the right. This effect is due to excitatory projections from the vestibular nucleus to the contralateral nucleus abducens that, along with the oculomotor nucleus, help execute conjugate eye movements.

For instance, horizontal movement of the two eyes toward the right requires contraction of the left medial and right lateral rectus muscles. Vestibular nerve fibers originating in the left horizontal semicircular canal project to the medial and lateral vestibular nuclei (see Figure 13.10). Excitatory fibers from the medial vestibular nucleus cross to the contralateral abducens nucleus, which has two outputs. One of these is a motor pathway

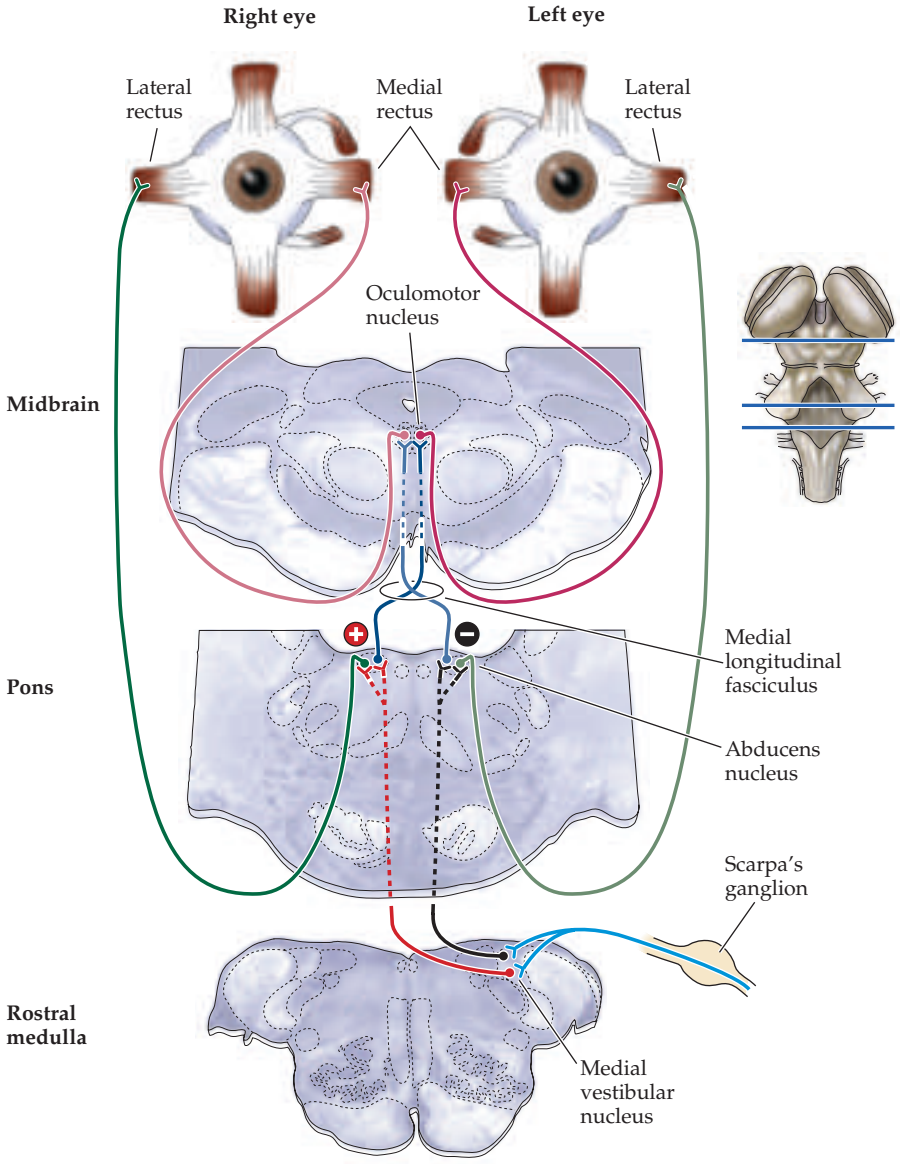


Figure 13.10 Connections underlying the vestibulo-ocular reflex. Projections of the vestibular nucleus to the nuclei of cranial nerves III (oculomotor) and VI (abducens). The connections to the oculomotor nucleus and to the contralateral abducens nucleus are excitatory (red), whereas the connections to ipsilateral abducens nucleus are inhibitory (black). There are connections from the oculomotor nucleus to the medial rectus of the left eye and from the adbucens nucleus to the lateral rectus of the right eye. This circuit moves the eyes to the right, that is, in the direction away from the left horizontal canal, when the head rotates to the left. Turning to the right, which causes increased activity in the right horizontal canal, has the opposite effect on eye movements. The projections from the right vestibular nucleus are omitted for clarity.

that causes the lateral rectus of the right eye to contract; the other is an excitatory projection that crosses the midline and ascends via the **medial longitudinal fasciculus** to the left oculomotor nucleus, where it activates neurons that cause the medial rectus of the left eye to contract. Finally, inhibitory neurons project from the medial vestibular nucleus to the left abducens nucleus, directly causing the motor drive on the lateral rectus of the left eye to decrease and also indirectly causing the right medial rectus to relax. The consequence of these several connections is that excitatory input from the horizontal canal on one side produces eye movements toward the opposite side. Therefore, turning the head to the left causes eye movements to the right.

In a similar fashion, head turns in other planes activate other semicircular canals, causing other appropriate compensatory eye movements. Thus, the VOR also plays an important role in vertical gaze stabilization in response to

the linear vertical head oscillations that accompany locomotion, and in response to vertical angular accelerations of the head, as can occur when riding on a swing. The rostrocaudal set of cranial nerve nuclei involved in the VOR (i.e., the vestibular, abducens, and oculomotor nuclei), as well as the VOR's persistence in the unconscious state, make this reflex especially useful for detecting brainstem damage in the comatose patient (see Box C).

Loss of the VOR can have severe consequences. A patient with vestibular damage finds it difficult or impossible to fixate on visual targets while the head is moving, a condition called **oscillopsia** ("bouncing vision"). If the damage is unilateral, the patient usually recovers the ability to fixate objects during head movements. However, a patient with bilateral loss of vestibular function has the persistent and disturbing sense that the world is moving when the head moves. The underlying problem in such cases is that information about head and body movements normally generated by the vestibular organs is not available to the oculomotor centers, so that compensatory eye movements cannot be made.

Descending projections from the vestibular nuclei are essential for postural adjustments of the head, mediated by the vestibulo-cervical reflex (VCR), and body, mediated by the vestibulo-spinal reflex (VSR). As with the VOR, these postural reflexes are extremely fast, in part due to the small number of synapses interposed between the vestibular organ and the relevant motor neurons (Box D). Like the VOR, the VCR and the VSR are both compromised in patients with bilateral vestibular damage. Such patients exhibit diminished head and postural stability, resulting in gait deviations; they also have difficulty balancing. These balance defects become more pronounced in low light or while walking on uneven surfaces, indicating that balance normally is the product of vestibular, visual, and proprioceptive inputs.

The anatomical substrate for the VCR involves the medial vestibular nucleus, axons from which descend in the medial longitudinal fasciculus to reach the upper cervical levels of the spinal cord (Figure 13.11). This pathway regulates head position by reflex activity of neck muscles in response to stimulation of the semicircular canals from rotational accelerations of the head. For example, during a downward pitch of the body (e.g., tripping), the superior canals are activated and the head muscles reflexively pull the head up. The dorsal flexion of the head initiates other reflexes, such as forelimb extension and hindlimb flexion, to stabilize the body and protect against a fall (see Chapter 16).

The VSR is mediated by a combination of pathways, including the lateral and medial vestibulospinal tracts and the reticulospinal tract. The inputs from the otolith organs project mainly to the lateral vestibular nucleus, which in turn sends axons in the lateral vestibulospinal tract to the spinal cord (see Figure 13.11). These axons terminate monosynaptically on extensor motor neurons, and they disynaptically inhibit flexor motor neurons; the net result is a powerful excitatory influence on the extensor (antigravity) muscles. When hair cells in the otolith organs are activated, signals reach the medial part of the ventral horn. By activating the ipsilateral pool of motor neurons innervating extensor muscles in the trunk and limbs, this pathway mediates balance and the maintenance of upright posture.

Decerebrate rigidity, characterized by rigid extension of the limbs, arises when the brainstem is transected above the level of the vestibular nucleus. Decerebrate rigidity in experimental animals is relieved when the vestibular nuclei are lesioned, underscoring the importance of the vestibular system to the maintenance of muscle tone. The tonic activation of extensor muscles in

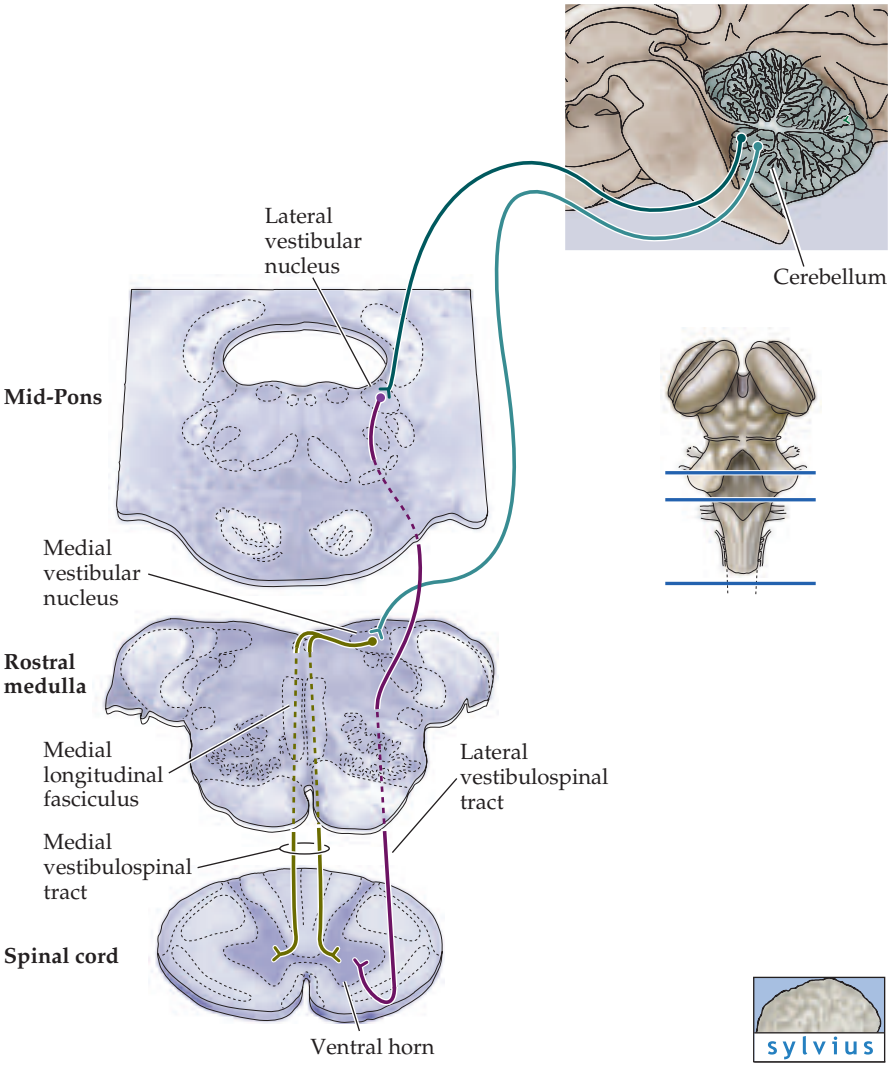


Figure 13.11 Descending projections from the medial and lateral vestibular nuclei to the spinal cord underlie the VCR and VSR. The medial vestibular nuclei project bilaterally in the medial longitudinal fasciculus to reach the medial part of the ventral horns and mediate head reflexes in response to activation of semicircular canals. The lateral vestibular nucleus sends axons via the lateral vestibular tract to contact anterior horn cells innervating the axial and proximal limb muscles. Neurons in the lateral vestibular nucleus receive input from the cerebellum, allowing the cerebellum to influence posture and equilibrium.

decerebrate rigidity suggests further that the vestibulospinal pathway is normally suppressed by descending projections from higher levels of the brain, especially the cerebral cortex (see also Chapter 16).

Vestibular Pathways to the Thalamus and Cortex

In addition to these several descending projections, the superior and lateral vestibular nuclei send axons to the ventral posterior nuclear complex of the thalamus, which in turn projects to two cortical areas relevant to vestibular

Box D

Mauthner Cells in Fish

A primary function of the vestibular system is to provide information about the direction and speed of ongoing movements, ultimately enabling rapid, coordinated reflexes to compensate for both self-induced and externally generated forces. One of the most impressive and speediest vestibular-mediated reflexes is the tail-flip escape behavior of fish (and larval amphibians), a stereotyped response that allows a potential prey to elude its predators (Figure A; tap on the side of a fish tank if you want to observe the reflex). In response to a perceived risk, fish flick their tail and are thus propelled laterally away from the approaching threat.

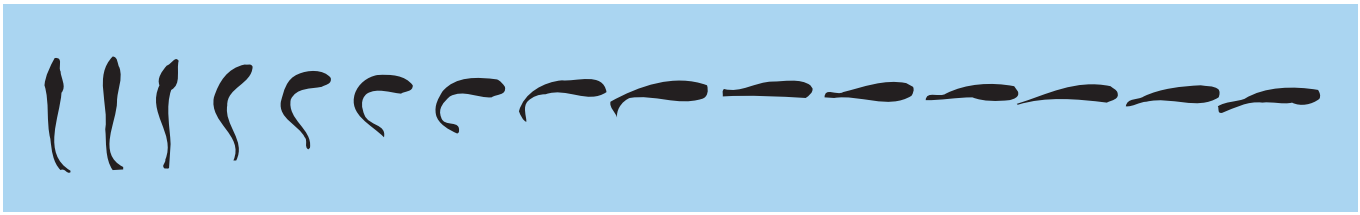
The circuitry underlying the tail-flip escape reflex includes a pair of giant medullary neurons called Mauthner cells, their vestibular inputs, and the spinal cord motor neurons to which the Mauthner cells project. (In most fish, there is one pair of Mauthner cells in a stereotypic location. Thus, these cells can be consistently visualized and studied from animal to animal.) Movements in the water, such as might be caused by an approaching predator, excite saccular hair cells in the vestibular labyrinth. These receptor potentials are transmitted via the central processes of vestibular ganglion cells in cranial nerve VIII to the two Mauthner cells in the brainstem. As in the vestibulo-spinal pathway in humans, the Mauthner cells project directly to spinal motor neurons. The small number of synapses intervening between the receptor cells and the motor neurons is one of the ways that this circuit has been optimized for speed by natural selection, an arrangement evident in humans as well. The large size of the Mauthner axons is another; the axons from these cells in a goldfish are about 50 µm in diameter.

The optimization for speed and direction in the escape reflex also is reflected in the synapses vestibular nerve afferents make on each Mauthner cell (Figure B). These connections are electrical synapses that allow rapid and faithful transmission of the vestibular signal.

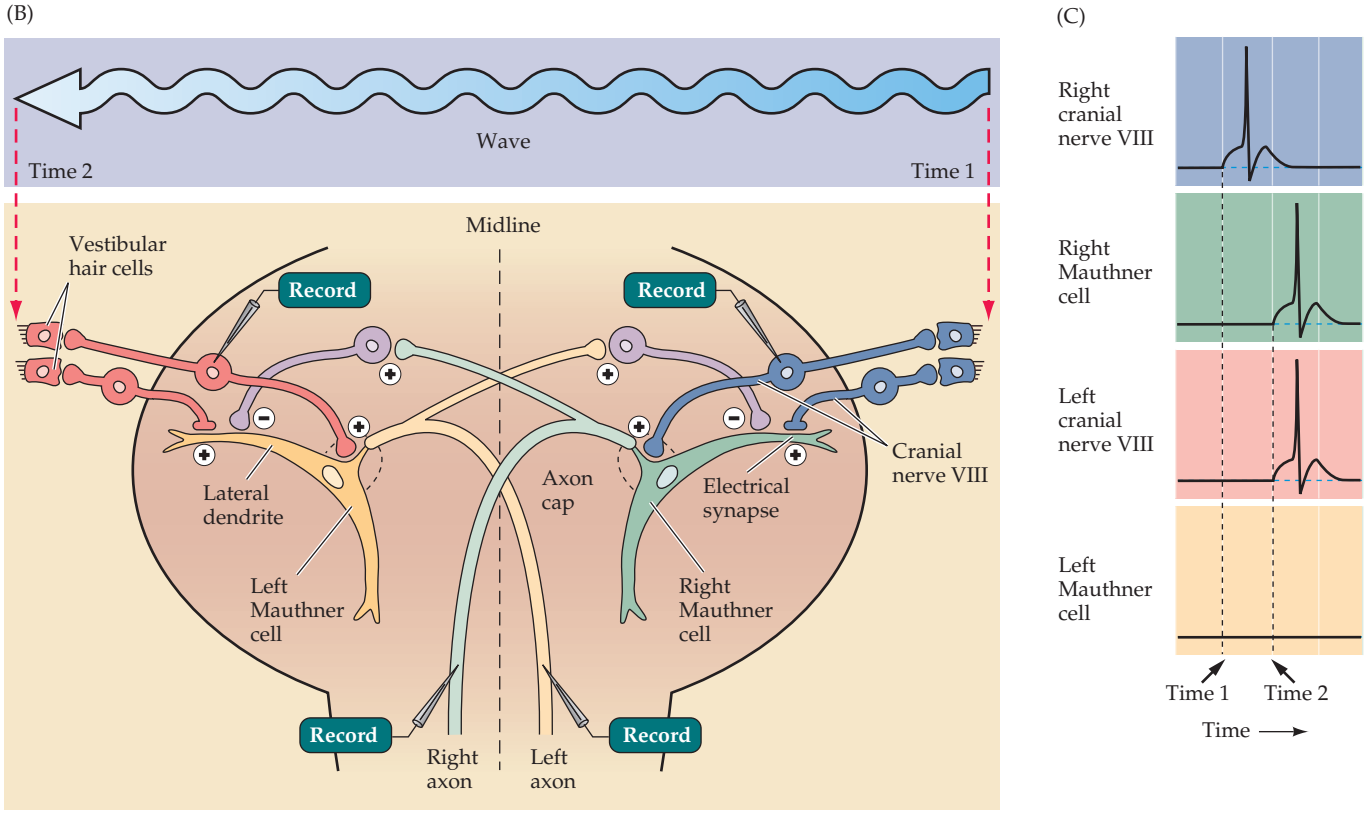
An appropriate direction for escape is promoted by two features: (1) each Mauthner cell projects only to contralateral motor neurons; and (2) a local network of bilaterally projecting interneurons inhibits activity in the Mauthner cell away from the side on which the vestibular activity originates. In this way, the Mauthner cell on one side faithfully generates action potentials that command contractions of contralateral tail musculature, thus moving the fish out of the path of the oncoming predator. Conversely, the Mauthner cell on the opposite side is silenced by the local inhibitory network during the response (Figure C).

(A) Bird's-eye view of the sequential body orientations of a fish engaging in a tail-flip escape behavior, with time progressing from left to right. This behavior is largely mediated by vestibular inputs to Mauthner cells.

(A)



sensations (Figure 13.12). One of these cortical targets is just posterior to the primary somatosensory cortex, near the representation of the face; the other is at the transition between the somatic sensory cortex and the motor cortex (Brodmann's area 3a; see Chapter 8). Electrophysiological studies of individual neurons in these areas show that the relevant cells respond to proprioceptive and visual stimuli as well as to vestibular stimuli. Many of these neurons are activated by moving visual stimuli as well as by rotation of the body (even with the eyes closed), suggesting that these cortical regions are involved in the perception of body orientation in extrapersonal space. Con-



(B) Diagram of synaptic events in the Mauthner cells of a fish in response to a disturbance in the water coming from the right. (C) Complementary responses of the right and left Mauthner cells mediating the escape response. Times 1 and 2 correspond to those indicated in Figure B. (After Furshpan and Furukuwa, 1962.)

The Mauthner cells in fish are analogous to the reticulospinal and vestibulospinal pathways that control balance, posture, and orienting movements in mammals. The equivalent behavioral responses in humans are evident in a friendly game of tag, or more serious endeavors.

References

Eaton, R. C., R. A. Bombardieri and D. L. Meyer (1977) The Mauthner-initiated startle response in teleost fish. J. Exp. Biol. 66: 65–81.

Furshpan, E. J. and T. Furukawa (1962) Intracellular and extracellular responses of the several regions of the Mauthner cell of the goldfish. J. Neurophysiol. 25:732–771.

Jontes, J. D., J. Buchanan and S. J. Smith (2000) Growth cone and dendrite dynamics in zebrafish embryos: Early events in synaptogenesis imaged *in vivo*. Nature Neurosci. 3: 231–237.

O'Malley, D. M., Y. H. Kao and J. R. Fetcho (1996) Imaging the functional organization of zebrafish hindbrain segments during escape behaviors. Neuron 17: 1145–1155.

sistent with this interpretation, patients with lesions of the right parietal cortex suffer altered perception of personal and extra-personal space, as discussed in greater detail in Chapter 25.

Summary

The vestibular system provides information about the motion and position of the body in space. The sensory receptor cells of the vestibular system are located in the otolith organs and the semicircular canals of the inner ear. The

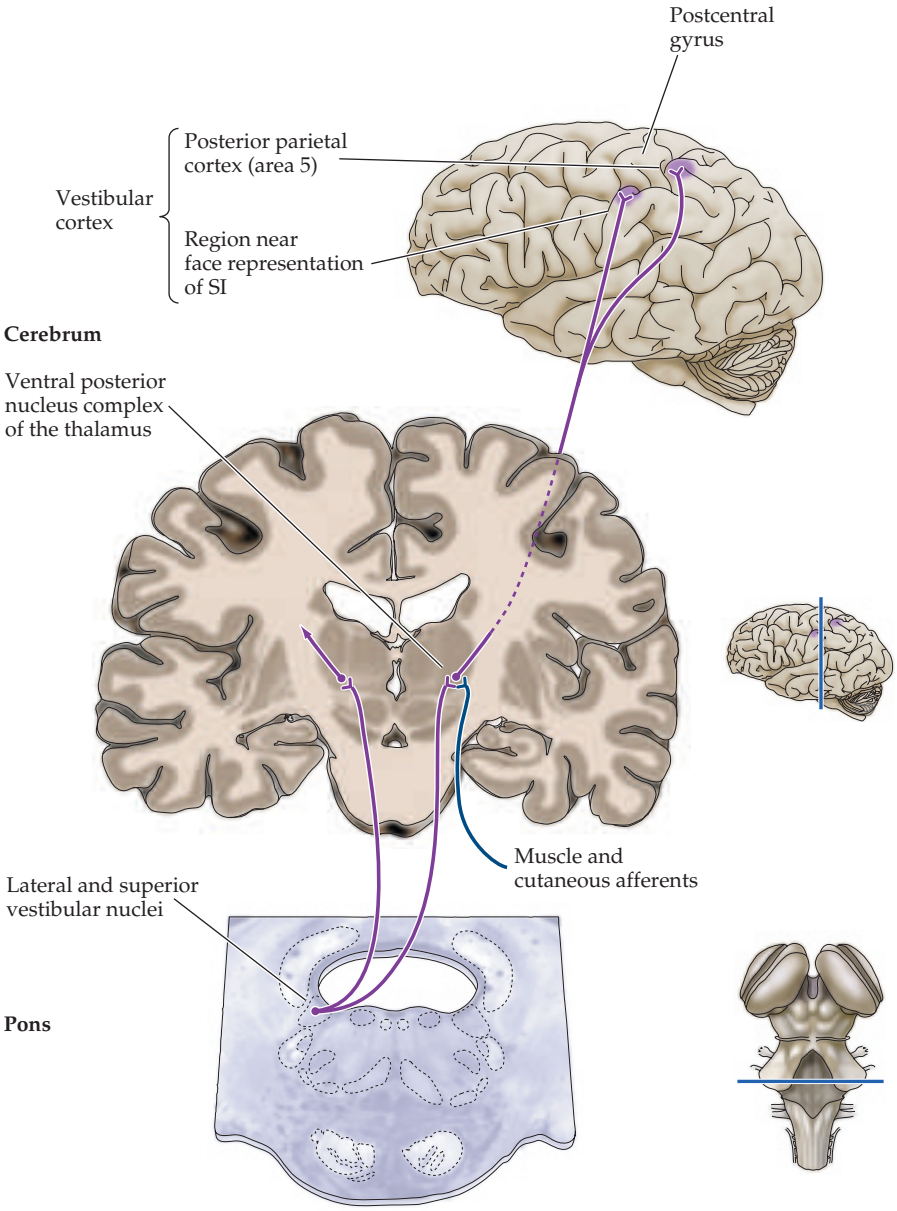


Figure 13.12 Thalamocortical pathways carrying vestibular information. The lateral and superior vestibular nuclei project to the thalamus. From the thalamus, the vestibular neurons project to the vicinity of the central sulcus near the face representation. Sensory inputs from the muscles and skin also converge on thalamic neurons receiving vestibular input (see Chapter 9).

otolith organs provide information necessary for postural adjustments of the somatic musculature, particularly the axial musculature, when the head tilts in various directions or undergoes linear accelerations. This information represents linear forces acting on the head that arise through static effects of gravity or from translational movements. The semicircular canals, in contrast, provide information about rotational accelerations of the head. This latter information generates reflex movements that adjust the eyes, head, and body during motor activities. Among the best studied of these reflexes are eye movements that compensate for head movements, thereby stabilizing the visual scene when the head moves. Input from all the vestibular organs is integrated with input from the visual and somatic sensory systems to provide perceptions of body position and orientation in space.

Additional Reading

Reviews

BENSON, A. (1982) The vestibular sensory system. In *The Senses*, H. B. Barlow and J. D. Mollon (eds.). New York: Cambridge University Press.

BRANDT, T. (1991) Man in motion: Historical and clinical aspects of vestibular function. A review. Brain 114: 2159–2174.

FURMAN, J. M. AND R. W. BALOH (1992) Otolith-ocular testing in human subjects. Ann. New York Acad. Sci. 656: 431–451.

GOLDBERG, J. M. (1991) The vestibular end organs: Morphological and physiological diversity of afferents. Curr. Opin. Neurobiol. 1: 229–235.

GOLDBERG, J. M. AND C. FERNANDEZ (1984) The vestibular system. In *Handbook of Physiology*, Section 1: *The Nervous System*, Volume III: *Sensory Processes*, Part II, J. M. Brookhart, V. B. Mountcastle, I. Darian-Smith and S. R. Geiger (eds.). Bethesda, MD: American Physiological Society.

HESS, B. J. (2001) Vestibular signals in self-orientation and eye movement control. News Physiolog. Sci. 16: 234–238.

RAPHAN, T. AND B. COHEN. (2002) The vestibulo-ocular reflex in three dimensions. Exp. Brain Res. 145: 1–27.

Important Original Papers

GOLDBERG, J. M. AND C. FERNANDEZ (1971) Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey, Parts 1, 2, 3. J. Neurophysiol. 34: 635–684.

GOLDBERG, J. M. AND C. FERNANDEZ (1976) Physiology of peripheral neurons innervating otolith organs of the squirrel monkey, Parts 1, 2, 3. J. Neurophysiol. 39: 970–1008.

LINDEMAN, H. H. (1973) Anatomy of the otolith organs. Adv. Oto.-Rhino.-Laryng. 20: 405–433.

Books

BALOH, R. W. AND V. HONRUBIA (2001) *Clinical Neurophysiology of the Vestibular System*, 3rd Ed. New York: Oxford University Press.

BALOH, R. W. (1998) *Dizziness, Hearing Loss, and Tinnitus.* Philadelphia: F. A. Davis Company.