

4

A Journey Through the
Auditory System

The main function of sensory systems is to get information about the outside world to the brain, where it can be used to help plan future behavior. In hearing, that information is carried by pressure variations in the air (sound waves). Someone says “I don’t like you” (information), the sound waves propagate to your ears, and you punch them on the nose (behavior). In this chapter we explore how the information in the sound waves is converted (or *transduced*) into a form that can be used by the brain; specifically in the form of electrical activity in nerve cells or *neurons*. Later in the book, we look at how our perceptions relate to the physiological mechanisms. First, however, we must learn something about the biological hardware involved; where it is, what it looks like, and what it does. Because the left and right ears are roughly mirror images of one another, I will only describe the anatomy of the right ear. It should be remembered that everything is duplicated on the opposite side of the head to produce a pleasing symmetry. Please also note that when specific numbers and dimensions are given, they refer to the *human* auditory system, unless otherwise stated.

The human ear is an exquisitely sensitive organ. We can detect displacements of the eardrum of less than one tenth the width of a hydrogen atom in response to

a 1000-Hz pure tone.¹ Natural selection has managed to engineer an instrument of such elegance and sophistication that our best efforts at sound recording and processing seem hopelessly crude in comparison. In the classic film “Fantastic Voyage,” an intrepid group of scientists travels around the blood vessels of a dying man in a miniaturized submarine, passing various organs and sites of interest as they go. In our version of the journey, we will follow a sound wave into the ear and continue, after transduction, up to the brain.

4.1 FROM AIR TO EAR

The main anatomical features of the peripheral auditory system are shown in Fig. 4.1. The peripheral auditory system is divided into the *outer ear*, *middle ear*, and *inner ear*.

4.1.1 Outer Ear

The *pinna* is the external part of the ear, that strangely shaped cartilaginous flap that you hook your sunglasses on. The pinna is the bit that gets cut off when someone has their ear cut off, although the hearing sense is affected very little by this amputation. Van Gogh did *not* make himself deaf in his left ear when he attacked his pinna with a razor in 1888. The pinnae are more important in other animals (bats, dogs, etcetera) than they are in humans. Our pinnae are too small and inflexible to be very useful for collecting sound from a particular direction, for example. They do, however, cause spectral modifications (i.e., filtering) to the sound as it enters the ear, and these modifications vary depending on the direction the sound is coming from. The spectral modifications help the auditory system determine the location of a sound source (see Section 9.2.2).

The opening in the pinna, the *concha*, leads to the *ear canal* (*external auditory meatus*) which is a short and crooked tube ending at the *eardrum* (*tympanic membrane*). The tube is about 2.5 cm long and has resonant properties like an organ pipe that is open at one end (see Section 3.1.3). Another way of thinking about this is that the ear canal acts like a broadly tuned band-pass filter. Because of the resonance of the ear canal and the concha, we are more sensitive to sound frequencies between about 1000 and 6000 Hz. The pinna, concha, and ear canal together make up the *outer ear*. The propagation of sound down the ear canal is the last stage in hearing in which sound waves are carried by the air.

¹Because the eardrum response is linear, the tiny displacements of the eardrum near hearing threshold ($<10^{-11}$ m) can be inferred from the displacements in response to higher-level sounds, such as those reported by Huber et al. (2001).

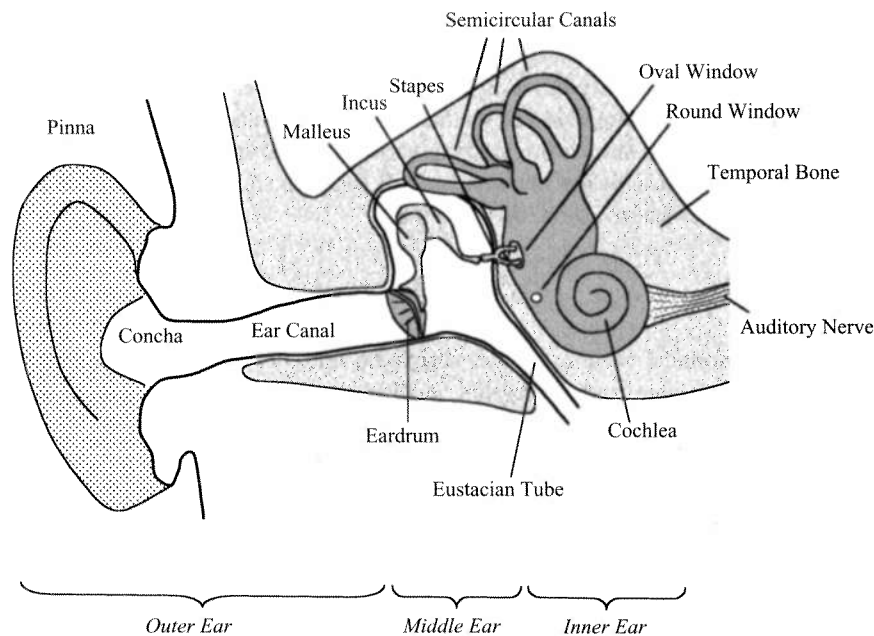


FIG. 4.1. The anatomy of the peripheral auditory system.

4.1.2 Middle Ear

The eardrum is a thin, taut, and easily punctured membrane that vibrates in response to pressure changes in the ear canal. On the other side of the eardrum from the ear canal is the *middle ear*. The middle ear is filled with air, and is connected to the back of the throat by the *eustachian tube*. Swallowing or yawning opens this tube to allow the pressure in the middle ear to equalize with the external air pressure. The pressure changes we experience as we climb rapidly in an aircraft, for instance, can cause an imbalance between the air pressures on both sides of the eardrum, causing our ears to “pop.” Swallowing helps alleviate this problem.

Although the middle ear is filled with air, the acoustic vibrations are carried from the eardrum to the cochlea (where transduction takes place) by three tiny bones—the smallest in the body—called the *malleus*, *incus*, and *stapes* (“hammer,” “anvil,” and “stirrup”). These bones are called, collectively, the *ossicles*. Their job is to transmit the pressure variations in an air-filled compartment (the ear canal) into pressure variations in a water-filled compartment (the *cochlea*) as efficiently as possible. The transmission is not as trivial as it might seem. If you shout from land at someone swimming under water, most of the sound is reflected back from the surface, because water has a much higher *impedance* than air (see Section 3.2.2). The bones in the middle ear solve this problem by concentrating the forces produced by the sound waves at the eardrum onto a smaller area (the *oval window*

in the cochlea). Because pressure equals force divided by area, the effect of this transformation is to increase the pressure by a factor of about 20. The ossicles also act as a lever system, so that large, weak vibrations at the eardrum are converted into smaller, stronger vibrations at the oval window. Finally, the eardrum itself performs a buckling motion that increases the force of the vibrations and decreases the displacement and velocity. The overall effect of all these components is to increase the pressure at the oval window to around 20–30 times that at the eardrum (see Rosowski & Relkin, 2001). This system is regarded as an *impedance-matching transformer*.

Attached to the malleus and stapes are small muscles that contract reflexively at high sound levels (above about 75 dB SPL). This increases the stiffness of the chain of ossicles and reduces the magnitude of the vibrations transmitted to the cochlea. The mechanism is most effective at reducing the level of low-frequency sounds (below about 1000 Hz), and acts like a high-pass filter. The reflex does not do much to protect the ear against high-frequency sounds, which are often the most damaging. Because the reflex involves neural circuits in the brainstem, the mechanism is also too slow (latency of 60–120 ms) to protect our ears against impulsive sounds, such as gunshots. Instead, the reflex may be involved in reducing the interference produced by intense low-frequency sounds, or in reducing the audibility of our own vocalizations (e.g., speech), which mostly reach our ears via the bones in our head.

4.2 THE COCHLEA

The cochlea, in the inner ear, is where transduction occurs. It is there that acoustic vibrations are converted into electrical neural activity. However, the cochlea is much more than a simple microphone. Structures within the cochlea perform processing on the sound waveform that is of great significance to the way we perceive sounds.

4.2.1 Anatomy

The cochlea is a fluid-filled cavity that is within the same compartment as the semi-circular canals that are involved in balance. (“Fluid” here means water with various biologically important chemicals dissolved in it.) The cochlea is a thin tube, about 3.5 cm long, with an average diameter of about 2 mm, although the diameter varies along the length of the cochlea, being greatest at the *base* (near the oval window) and least at the *apex* (the other end of the tube). The cochlea, however, is not a straight tube. The tube has been coiled up to save space. The whole structure forms a spiral, similar to a snail shell, with about two and a half turns from the base to the apex (see Fig. 4.1). The cochlea has rigid bony walls. I heard a story of a student who was under the impression that as sound enters the ear, the

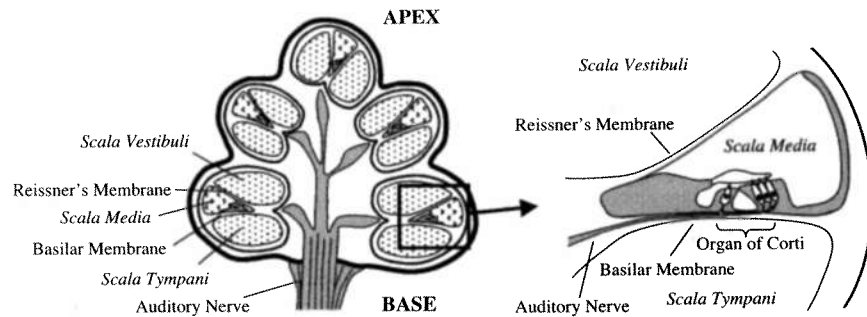


FIG. 4.2. Two magnifications of a cross-section of the cochlea. The spiral is viewed from the side, in contrast to the view from above in Figure 4.1.

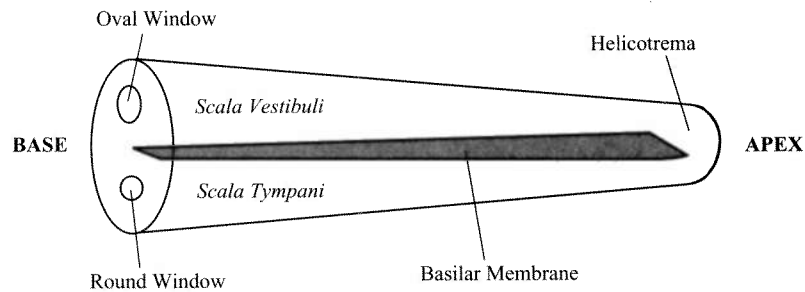


FIG. 4.3. A highly schematic illustration of the cochlea as it might appear if the spiral were unwound. The vertical dimension is exaggerated relative to the horizontal. Reissner's membrane and the scala media are not illustrated.

cochlea stretches out by uncoiling the spiral like the “blow out” horn you find at parties. Note: This does *not* happen.

A cross-section of the cochlea (see Fig. 4.2) reveals that the tube is divided along its length by two membranes, *Reissner's membrane* and the *basilar membrane*. This creates three fluid-filled compartments: the *scala vestibuli*, the *scala media*, and the *scala tympani*. The scala vestibuli and the scala tympani are connected by a small opening (the *helicotrema*) between the basilar membrane and the cochlea wall at the apex (see Fig. 4.3). The scala media, however, is an entirely separate compartment that contains a different fluid composition (*endolymph*) from that in the other two scalae (*perilymph*).

Figure 4.4 shows a further magnification of a part of the cochlear cross-section shown in Fig. 4.2. Above the basilar membrane is a gelatinous structure called the *tectorial membrane*. Just below this, and sitting on top of the basilar membrane, is the *organ of Corti*, which contains rows of *hair cells* and various supporting cells and nerve endings. Cells are the tiny little bags of biochemical machinery, held

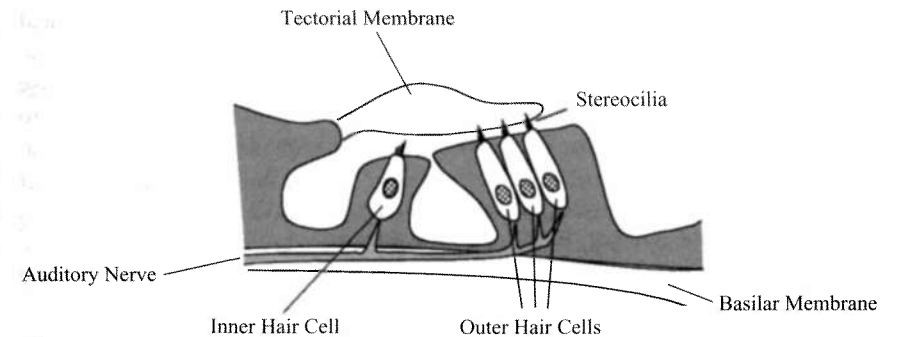


FIG. 4.4. The tectorial membrane and the organ of Corti.

together by a membrane, that make up most of the human body. Hair cells are very specialized types of cells. As their name suggests, they have minute little hairs, or more correctly, *stereocilia*, sticking out of their tops. In each human cochlea there is one row of *inner hair cells* (closest to the inside of the cochlear spiral) and up to five rows of *outer hair cells*. Along the length of the cochlea there are thought to be about 3,500 inner hair cells and about 12,000 outer hair cells (Møller, 2000). The tallest tips of the stereocilia of the outer hair cells are embedded in the tectorial membrane, whereas the stereocilia of the inner hair cells are not. The outer hair cells change the mechanical properties of the basilar membrane, as described in Chapter 5. The inner hair cells are responsible for converting the vibration of the basilar membrane into electrical activity.

4.2.2 The Basilar Membrane

Sound enters the cochlea through an opening (the oval window) covered by a membrane. The fluid in the cochlea is almost incompressible, so if the oval window moves in suddenly, due to pressure from the stapes, Reissner's membrane and the basilar membrane are pushed down, and the *round window* (a second membrane-covered opening at the other side of the base) moves out. It follows that vibration of the stapes leads to vibration of the basilar membrane.

The basilar membrane is very important to mammalian hearing. The basilar membrane separates out the *frequency components* of a sound. At the base of the cochlea, near the oval window, the basilar membrane is narrow and stiff. This area is most sensitive to high frequencies. The other end of the membrane, at the tip or apex of the cochlea, is wide and loose and is most sensitive to low frequencies. (Note that the basilar membrane becomes *wider* as the cochlea becomes *narrower*, see Fig. 4.3). The properties of the membrane vary continuously between these extremes along its length, so that each place on the basilar membrane has a particular frequency of sound, or *characteristic frequency*, to which it is most sensitive.

You can understand how this works if you are familiar with stringed musical instruments. The higher the tension in the string, the higher the frequency of the note that is produced: Stiff strings have higher *resonant frequencies* than do loose ones (see Section 3.1.4). Another simple way to understand this mechanism is to imagine a long series of springs, hanging alongside each other from a horizontal wooden rod. Each spring has a mass attached to it. (The basilar membrane is not actually composed of a series of coiled springs: This is just an analogy.) The springs at the left end (corresponding to the base of the cochlea) are very stiff. As we move along, toward the right end of the rod, the springs get looser until, at the end of the rod (corresponding to the apex of the cochlea), the springs are very loose.

If you have played with a spring, you understand some of the properties of these systems. If you attach a mass to the end of a spring, pull it down, and release it, the mass will move up and down at a particular rate, the oscillations slowly dying out over time. The rate of oscillation is the resonant frequency of the system (see Section 3.1.2). If the spring is very stiff, the mass will move up and down rapidly, i.e., at a high frequency. If the spring is very loose, the mass will move up and down slowly, i.e., at a low frequency. As described in Section 3.1.2, if you hold the end of the spring and move your hand up and down at a rate higher than the resonant frequency, the mass and spring may vibrate a little but not much movement will be obtained. If you move your hand up and down at the resonant frequency, however, then the oscillations will build and build and become much more intense.

Now imagine that the whole rod of masses and springs is moved up and down at a particular rate (this corresponds to stimulating the basilar membrane with a pure tone at a particular frequency). To be accurate, the movement should be sinusoidal. What you would see is that a small group of masses and springs vibrate very strongly over large distances. For these masses and springs, the motion of the rod is close to their resonant frequencies. If the rod is moved up and down at a slower rate, this group of springs would be located nearer to the right (apex) of the rod. If a higher rate is used, springs near the left (base) of the rod would be excited. If you could move the rod up and down at two rates at once (impose lots of high frequency wiggles on a low frequency up and down movement) a group of springs to the left would respond to the high frequency movement and a group of springs to the right would respond to the low frequency movement. We would have effectively *separated out* the two different frequencies.

Unfortunately for us, the basilar membrane and the cochlea as a whole are much more complex than this simple model might suggest. The motion of the membrane is affected by the inertia of the surrounding fluids, resonance in the tectorial membrane, and the stereocilia of the outer hair cells (see Møller, 2000, page 82), and, crucially, by the mechanical action of the outer hair cells. The outer hair cells enhance the tuning of the basilar membrane by actively influencing the motion of the membrane. We look at this important mechanism in Chapter 5.

Despite the complications, it is fair to say that the basilar membrane behaves like a continuous array of tuned resonators, and this means that it behaves as a *bank of*

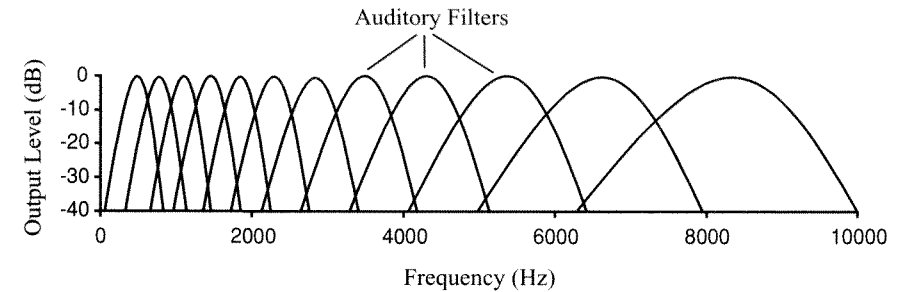


FIG. 4.5. A popular model of the cochlea, in which the frequency selectivity of the basilar membrane is represented by an array of overlapping band-pass filters. Each curve shows the relative attenuation characteristics of one auditory filter. The curves to the left show the responses of places near the apex of the cochlea, whereas those to the right show the responses of places near the base. The basilar membrane is effectively a *continuous* array of filters: many more filters, which are much more tightly spaced, than those in the figure.

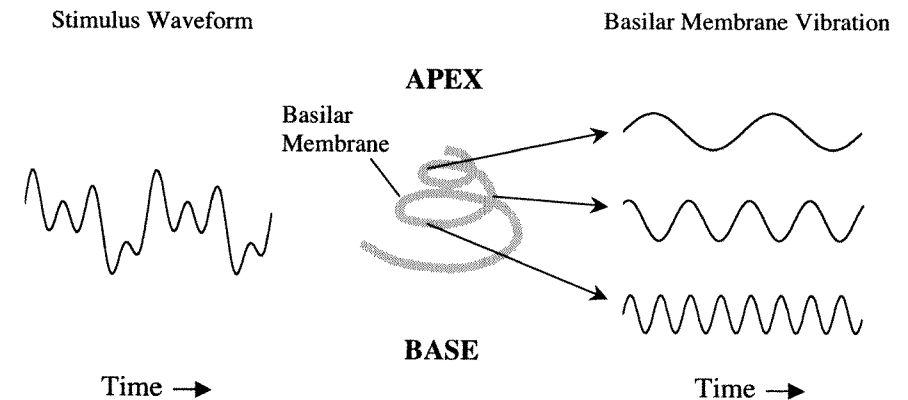


FIG. 4.6. A schematic illustration of how a complex sound waveform (left) is decomposed into its constituent frequency components (right) by the basilar membrane. The approximate locations on the basilar membrane that are vibrating with these patterns are shown in the center.

overlapping band-pass filters (see Fig. 4.5). These filters are often called *auditory filters*. Each place on the basilar membrane has a particular characteristic frequency, a bandwidth, and an impulse response. When a complex sound enters the ear, the higher-frequency components of the sound excite the basilar membrane toward the base and the lower-frequency components excite the basilar membrane toward the apex. The mechanical separation of the individual frequency components depends on their frequency separation. In this way, the basilar membrane performs a *spectral analysis* of the incoming sound (see Fig. 4.6).

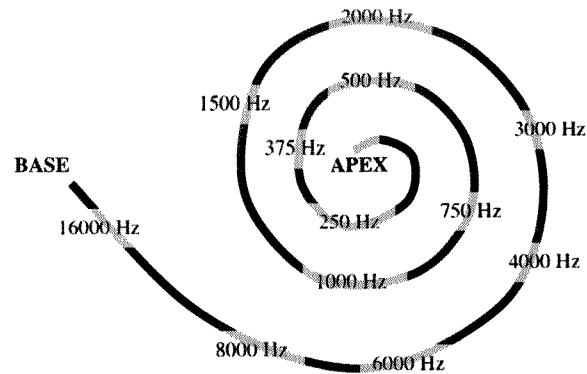


FIG. 4.7. The approximate distribution of characteristic frequencies around the human cochlea, with a viewpoint above the spiral.

The characteristic frequency of a specific place on the basilar membrane does not increase in a linear way as we go from the apex to the base: A constant distance along the basilar membrane does *not* correspond to a constant difference in characteristic frequency. The position on the membrane is more accurately a *logarithmic* function of characteristic frequency, so that high frequencies are spaced much more closely together than low frequencies. Figure 4.7 illustrates the distribution of characteristic frequencies around the cochlea.

4.2.3 The Traveling Wave

The basilar membrane in humans and other mammals is a tiny delicate structure (only 0.45 mm at its widest point) hidden within the bony walls of the cochlea. Nevertheless, by careful surgery, physiologists have been able to make direct observations of the motion of the basilar membrane in response to sound. Von Békésy (1960) was the pioneer of this line of research, observing the motion of the basilar membrane in cochleae isolated from human and animal cadavers. Actually, he observed the motion of silver particles scattered on Reissner's membrane, but since Reissner's membrane moves with the whole *cochlear partition* (the structures around the scala media, including the basilar membrane and the organ of Corti) the responses he measured apply to the basilar membrane as well.

Von Békésy observed that if a pure tone is played to the ear, a characteristic pattern of vibration is produced on the basilar membrane. If we imagine the cochlea is stretched out to form a thin, straight tube, the motion of the basilar membrane looks a bit like a water wave traveling from the base to the apex of the cochlea. This pattern of vibration is called a *traveling wave*, as illustrated in Fig. 4.8. If we follow the wave from the base to the apex, we can see that it builds up gradually until it reaches a maximum (at the place on the basilar membrane that resonates at the

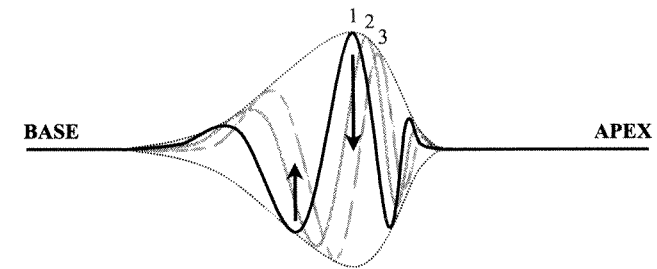


FIG. 4.8. Three time frames in the motion of the basilar membrane in response to a pure tone. The arrows show the direction of motion of the basilar membrane at two places along its length. The dotted lines show the envelope traced out by the traveling wave (i.e., the maximum displacement at each place). Compared to the real thing, these plots have been *hugely* exaggerated in the vertical direction.

frequency of the tone), before diminishing rapidly. Note also that the wavelength reduces from the base to the apex. Similar to a water wave on a pond, the traveling wave does not correspond to any movement of material from base to apex. Rather, the wave is a consequence of each place on the basilar membrane moving up and down in response to the pure-tone stimulation. It is important to remember that the frequency of vibration at each place on the basilar membrane is equal to the frequency of the pure tone.

A common misconception is that the motion of the traveling wave from base to apex is a result of the pressure variations entering the cochlea at the oval window (i.e., at the base). This is *not* the case. Sound travels very quickly in the cochlear fluids and thus all places on the basilar membrane are stimulated virtually instantaneously when there is a pressure variation at the oval window. The traveling wave would look the same if sound entered near the apex rather than the base. The characteristic motion of the traveling wave arises because there is a progressive *phase delay* from base to apex. That is, the vibration of the membrane at the apex lags behind that at the base. I will not go into the details, but this is a characteristic of the filtering properties of the basilar membrane, and results from the fact that a stiffness-limited system responds more quickly than a mass-limited system (see Pickles, 1988).

The peak of the traveling wave traces an outline, or envelope, which shows the overall region of response on the basilar membrane. Although there is a peak at one place on the basilar membrane, the region of response covers a fair proportion of the total length, especially for low-frequency sounds (see Fig. 4.9). This is because each place acts as a band-pass filter and responds to a range of frequencies. It is clearly not the case that each place on the basilar membrane responds to one frequency and one frequency only (although the response will be maximal for stimulation at the characteristic frequency). Indeed, in response to very intense

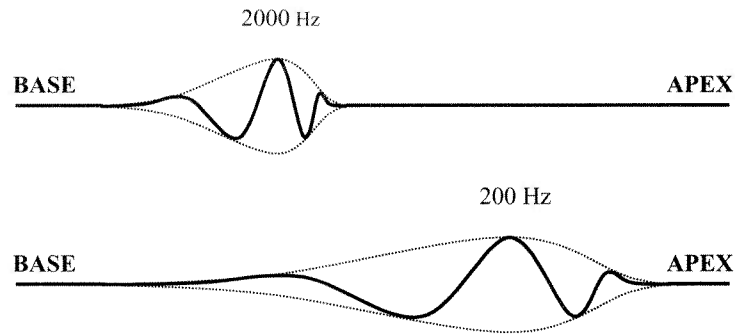


FIG. 4.9. A snapshot of the basilar membrane displacement at a single instant, in response to pure tones with two different frequencies. Based on measurements by von Békésy (see von Békésy, 1960).

low-frequency sounds, every place on the membrane produces a significant vibration, irrespective of characteristic frequency.

4.3 TRANSDUCTION

We have seen that the sound waves entering the ear produce vibrations on the basilar membrane. The different frequencies in the sound wave are separated onto different places on the basilar membrane. This is all a pointless exercise if the ear can not now tell the brain which parts of the membrane are vibrating and by how much. The ear must convert the mechanical vibrations of the basilar membrane into electrical activity in the auditory nerve. This task is accomplished by the inner hair cells.

4.3.1 How Do Inner Hair Cells Work?

On top of each hair cell are rows of stereocilia, which are like tiny hairs. When the basilar membrane and the tectorial membrane move up and down, they also move sideways relative to one another. This “shearing” motion causes the stereocilia on the hair cells to sway from side to side (see Fig. 4.10). The movement of the stereocilia is very small (the figures in this section show huge exaggerations of the actual effect). For a sound near the threshold of audibility, the displacement is only 0.3 *billionths* of a meter. If the stereocilia were the size of the Sears Tower in Chicago, then this would be equivalent to a displacement of the top of the tower by just 5 cm (Dallos, 1996).

The stereocilia are connected to one another by protein filaments called *tip links*. When the stereocilia are bent toward the scala media (i.e., toward the outside of

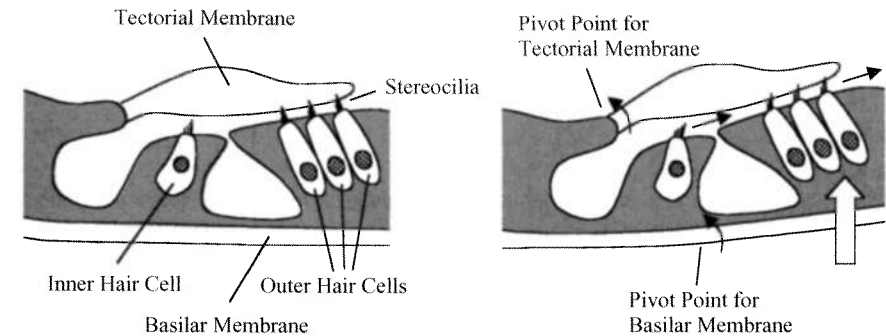


FIG. 4.10. An illustration of how displacement of the basilar membrane toward the scala vestibuli (curved arrows) produces a shearing force between the basilar membrane and the tectorial membrane, causing the stereocilia on the hair cells to be bent to the right (straight arrows). The basilar membrane and the tectorial membrane pivot about the points shown on the figure. Displacement of the basilar membrane toward the scala tympani produces the opposite effect, causing the stereocilia to be bent to the left (not shown).

the cochlea) the tip links are stretched. The stretching causes them to pull on tiny trap doors blocking channels in the membranes of the stereocilia (see Fig. 4.11). If these channels are opened up, positively charged potassium ions flow into the hair cell and produce an increase in the *electric potential* of the cell (on the order of a few thousandths of a volt, or millivolts, mV). Because the “resting” electric potential of the inner hair cell is negative (about -45 mV), the increase in potential is called *depolarization*. Depolarization causes a chemical *neurotransmitter* to be released into the tiny gap (or *synaptic cleft*) between the hair cell and the neuron in the auditory nerve (see Fig. 4.12). When the neurotransmitter arrives at the neuron, it causes electrical activity in the neuron (neural *spikes*, see Section 4.4.1). When the stereocilia are bent in the opposite direction (i.e., toward the center of the cochlea), the tip links slacken, the channels stay closed, and the release of neurotransmitter is reduced. The larger the movement of the basilar membrane, the more tip links are opened. The greater the electrical change in the hair cell, the more neurotransmitter is released, and the greater the resulting activity in the auditory nerve.

The outer hair cells are activated in the same way as the inner hair cells—by the bending of the stereocilia and the opening of ion channels. However, it is thought that the resulting changes in the electric potential of the cell produce changes in the *cell length*, thus allowing the outer hair cell to affect the motion of the basilar membrane (see Section 5.2.5). Outer hair cells are *not* involved in the transmission of information about basilar-membrane motion to the auditory nerve and to the brain.

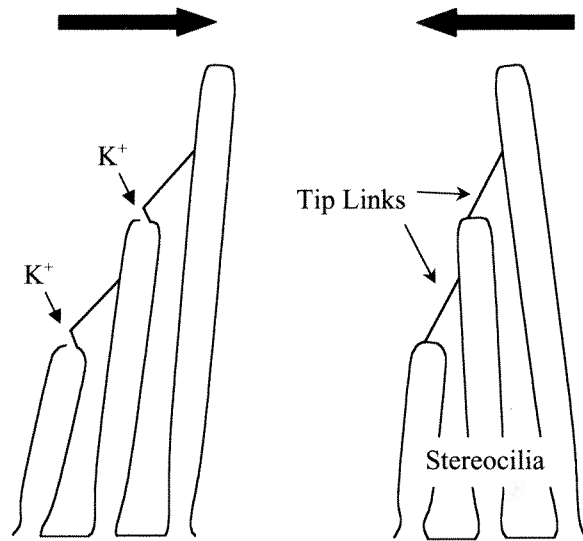


FIG. 4.11. How movement of the stereocilia causes an electrical change in the hair cell. When the stereocilia are bent to the right (toward scala media), the tip links are stretched and ion channels are opened. Positively charged potassium ions (K^+) enter the cell, causing the interior of the cell to become more positive (depolarization). When the stereocilia are bent in the opposite direction the tip links slacken and the channels close.

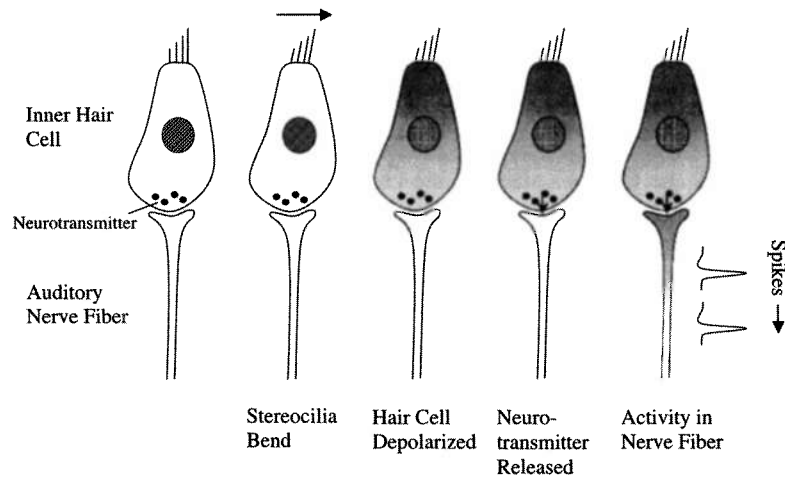


FIG. 4.12. The main stages in the transduction process. Time proceeds from left to right.

4.4 THE AUDITORY NERVE

4.4.1 Neurons

One problem with being a large organism is the difficulty of passing messages between different parts of the body. The solution for all large animals is a nervous system. A nervous system is comprised of cells called *neurons*. Neurons are responsible for rapid communication between sensory cells, muscle cells, and the brain. The human brain contains over a hundred billion neurons, each of which has hundreds of connections to other neurons. The neurons and connections form a processing network of enormous complexity and power that enables us to think, feel, and watch TV.

A neuron is composed of four main structures: the *dendrites*, the *soma* (or cell body), the *axon*, and the *terminal buttons*. Figure 4.13 shows the structures of two

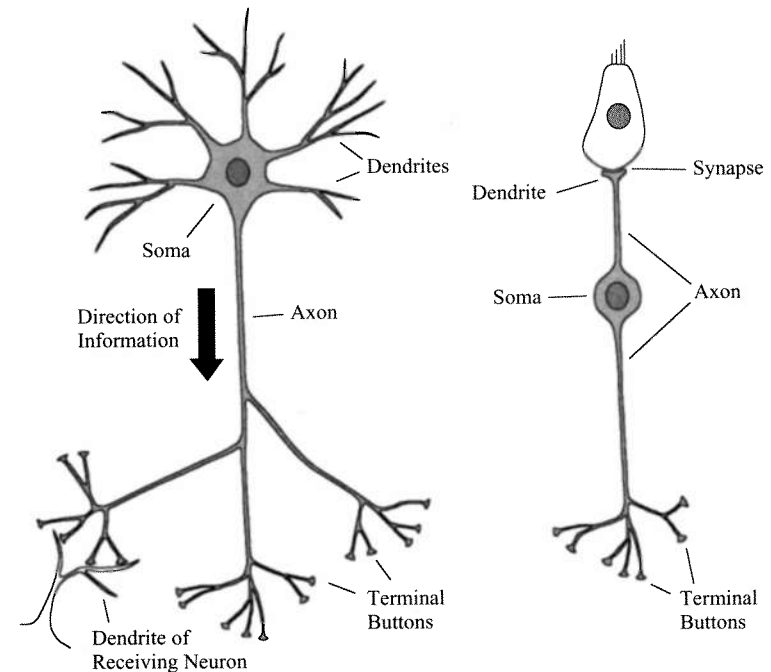


FIG. 4.13. An illustration of the structures of two neurons. On the left is the type of neuron one might find in the brain, with many dendrites and terminal buttons. At the bottom left the terminal buttons of the neuron are shown forming synapses with the dendrites of another neuron. On the right is a sensory neuron with one dendrite, in this case one of the neurons from the auditory nerve. The lengths of the axons, and the complexity of the branching dendrites and axons, have been reduced for illustrative purposes.

two typical neurons. Broadly speaking, the dendrites *receive* information (from sensory cells like the inner hair cells or from other neurons), the soma *integrates* the information, the axon *carries* the information, and the terminal buttons *pass the information on*, usually to the dendrites of another neuron. A connection between a terminal button and a dendrite, or between a sensory cell and a dendrite, is called a *synapse*. In the brain, the dendrites of a single neuron usually form synapses with the terminal buttons of hundreds of other neurons.

Axons can be quite long (almost a *meter* in length for some “motor” neurons involved in the control of muscles). They carry information in the form of electrical impulses called *action potentials* or *spikes*. The magnitude of every spike is the same (about 100 mV), so that information is carried by the *firing rate* (number of spikes per second) or *pattern* of spikes, not by variations in the magnitude of the potential change for each spike. Spikes travel along the axon at speeds of up to 120 meters per second. The change in electric potential caused by the arrival of a spike at the terminal button triggers the release of neurotransmitter that diffuses across the synaptic cleft between the two cells. The more spikes that arrive, the more neurotransmitter is released. If neurotransmitter is detected by the receiving neuron, then this may trigger—or inhibit—the production of spikes in that neuron. In other words, the connection between two neurons can be *excitatory* or *inhibitory*. Strictly speaking, neural communication is *electrochemical* in nature: Electrical impulses in one neuron lead to the release of a chemical that influences the production of electrical impulses in another neuron. That, multiplied several hundred billion times, is how your brain works.

4.4.2 Activity in the Auditory Nerve

The auditory nerve is a bundle of axons or *nerve fibers* that are connected to (*synapse with*) the hair cells. In total, there are about 30,000 neurons in the human auditory nerve. The majority of nerve fibers connect to the inner hair cells. Each inner hair cell is contacted by the dendrites of approximately 20 auditory nerve fibers (statistics from Møller, 2000). Because each inner hair cell is attached to a specific place on the basilar membrane, each neuron in the auditory nerve carries information about the vibration of the basilar membrane at a *single place* in the cochlea. Because each place in the cochlea is most sensitive to a particular characteristic frequency, each neuron in the auditory nerve is also most sensitive to a particular characteristic frequency. Figure 4.14 shows the tuning properties of neurons with a range of characteristic frequencies. The figure shows that each neuron becomes progressively less sensitive as the frequency of stimulation is moved away from the characteristic frequency, as does the place on the basilar membrane to which the neuron is connected. Tuning curves are essentially inverted versions of the filter shapes we discussed in this chapter (compare with Fig. 4.5). Tuning curves are discussed further in Section 5.3.1.

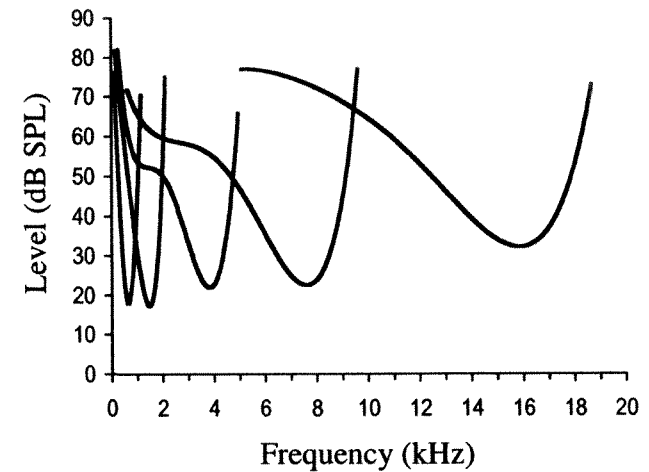


FIG. 4.14. Frequency threshold tuning curves recorded from the auditory nerve of a chinchilla. Each curve shows the level of a pure tone required to produce a just-measurable increase in the firing rate of a neuron, as a function of the frequency of the tone. Low levels indicate high sensitivity. Five curves are shown, illustrating the tuning properties of five neurons with characteristic frequencies ranging from about 500 Hz to about 16 kHz. The curves are smoothed representations of recordings made by Ruggero and Semple (see Ruggero, 1992).

In terms of spatial layout, the characteristic frequencies of the nerve fibers increase from the center to the periphery of the auditory nerve. Those fibers near the center of the auditory nerve bundle originate in the apex of the cochlea and have low characteristic frequencies, and those fibers near the periphery of the auditory nerve originate in the base of the cochlea and have high characteristic frequencies (Fig. 4.2 illustrates the pattern of innervation). The spatial frequency map in the cochlea is preserved as a spatial frequency map in the auditory nerve. The organization of frequency in terms of place is called *tonotopic organization*, and is preserved right up to the auditory cortex, part of the cerebral cortex of the brain. The place on the basilar membrane that is excited determines the place in the auditory nerve that is excited, which (via several other staging posts) determines the places on the auditory cortex that are excited. Information is carried through the auditory system in *frequency channels*.

In quiet, most fibers in the auditory nerve show a background level of firing called *spontaneous activity*. Most fibers (perhaps 90%) have *high* spontaneous rates, producing about 60 spikes per second. These fibers tend to be quite sensitive and show an increase in firing rate in response to a low stimulus level. The remaining fibers have low spontaneous rates of less than about 10 spikes per second. These fibers tend to be less sensitive. The difference in sensitivity may be related

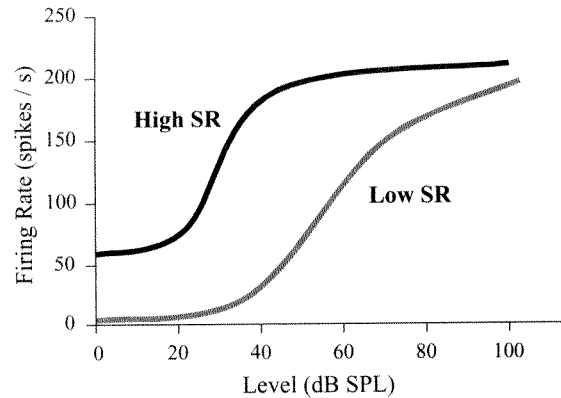


FIG. 4.15. An illustration of the relation between the level of a tone at characteristic frequency and firing rate (in spikes per second), for auditory nerve fibers with high (high SR) and low (low SR) spontaneous firing rates. Based (loosely) on recordings from the cat by Sachs and Abbas (1974).

to the location of the synapse with the inner hair cell. High spontaneous rate fibers synapse on the side of the cell closest to the outer hair cells. Low spontaneous rate fibers synapse on the opposite side of the cell (see Sewell, 1996). When stimulated with a pure tone at its characteristic frequency, a neuron will increase its firing rate as the level of the tone is increased, up to a certain maximum firing rate, at which point the response is *saturated*: Further increases in level will have no effect on the firing rate. A plot of firing rate against sound level is called a *rate-level function*. In general, high spontaneous rate fibers have steeper rate-level functions that saturate at a much lower level than do low spontaneous rate fibers (see Fig. 4.15). An explanation for this difference is provided in Section 5.3.2.

Auditory nerve fibers also show a characteristic change in firing rate with time from the onset of a sound. When a sound is turned on, the fibers produce a peak of activity (the onset response) that declines with time. In addition, when the sound is turned off, the activity in a neuron falls below its spontaneous activity for 100 milliseconds or so (see Fig. 4.16). The neuron is said to be *adapted*. Adaptation may be the result of the depletion of neurotransmitter from the inner hair cell. When an intense tone is first turned on, the inner hair cell releases a lot of neurotransmitter, and produces a large response in the auditory nerve fiber. However, the hair cell then has to replenish its supply of neurotransmitter, and until it has a chance to do so (when the tone is turned off), it cannot respond as strongly as it did at the tone's onset.

4.4.3 Place Coding

The firing rate of a neuron in the auditory nerve is determined by the magnitude of basilar-membrane vibration at the place to which it is connected. It follows

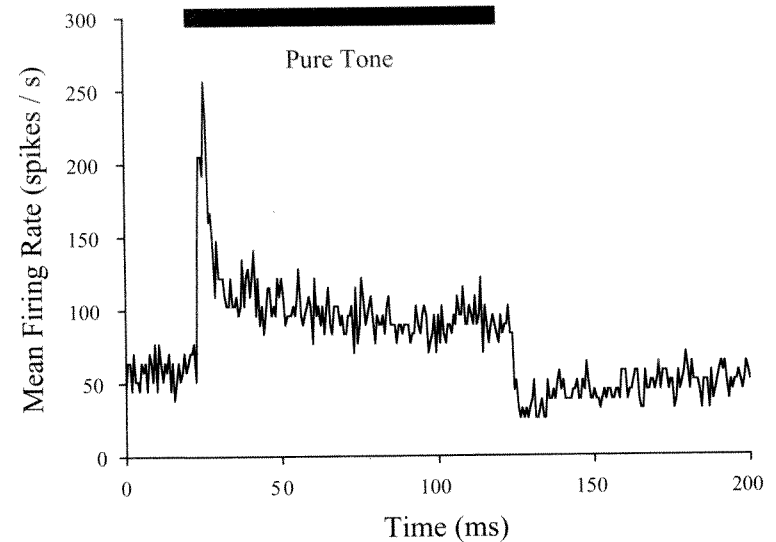


FIG. 4.16. A simulation of the activity over time of a high spontaneous rate auditory nerve fiber in response to a 100-ms pure tone (the time course of the tone is indicated by the thick black line above the plot). The vertical scale represents the mean firing rate over 1,500 repetitions of the stimulus.

that each neuron in the auditory nerve has a particular characteristic frequency, and is only sensitive to a limited range of frequencies around this frequency. In addition, increases in sound level increase the firing rate of the neuron (up to the saturation level). Therefore, one way in which the auditory system represents the spectrum of a sound is in terms of the firing rates of different neurons in the auditory nerve. If a sound with low-frequency components is presented, then neurons with low characteristic frequencies (near the center of the auditory-nerve bundle) will increase their firing rates. If a sound with high-frequency components is presented, then neurons with high characteristic frequencies (near the periphery of the auditory-nerve bundle) will increase their firing rates. Representation of spectral information in this way is called a *place code* or a *rate-place code*, because the spectral information is represented by the pattern of activity across the array of neurons.

4.4.4 Phase Locking and Temporal Coding

Place coding is not the only way in which the characteristics of sounds are represented. An electrical change in the inner hair cells occurs only when their stereocilia are bent toward the outside of the cochlea (see Section 4.3.1). If the basilar membrane is vibrating happily up and down in response to a low-frequency pure tone,

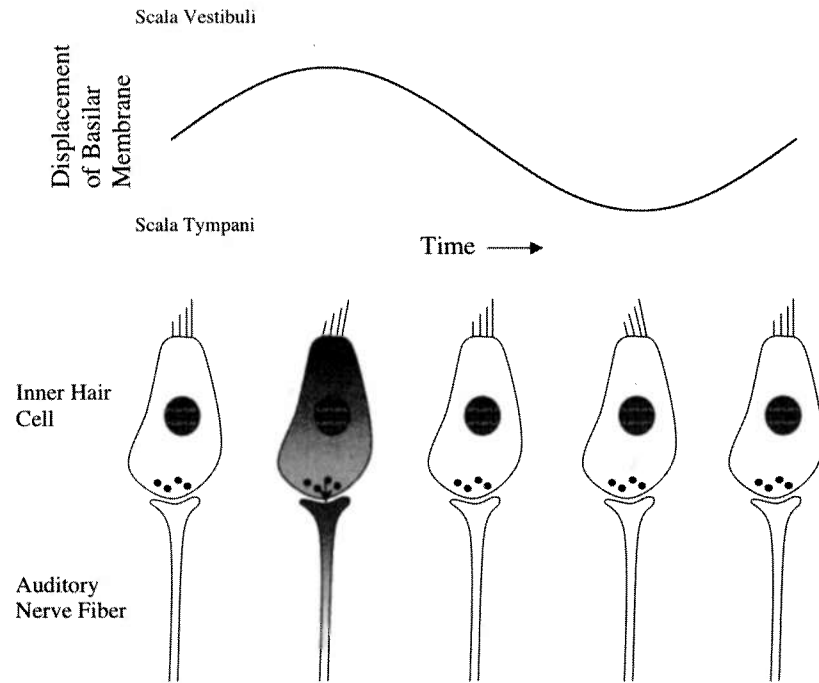


FIG. 4.17. An illustration of how electrical activity in the inner hair cells and in the auditory nerve is related to the motion of the basilar membrane. Activity is greatest at a particular phase during each cycle of basilar membrane vibration (indicated by the second hair cell from the left).

the stereocilia will bend from side to side, but the hair cells will only depolarize when the stereocilia are bent in one direction, i.e., at a particular *phase* of the vibration. This, in turn, means that neurons in the auditory nerve will tend to produce spikes at a particular phase of the waveform. This property is called *phase locking*, because the response of the neuron is locked to a particular phase of the stimulation, or more accurately, a particular phase in the *vibration of the basilar membrane*. The mechanism of phase locking is illustrated in Fig. 4.17.

The existence of phase locking immediately suggests another way in which frequency can be represented in the auditory nerve, specifically in terms of the *timing* or *synchrony* of the activity in the auditory nerve. If a 100-Hz pure tone is presented, neurons will *tend* to produce spikes that are spaced at integer multiples of the period of the waveform, in this case 10 ms (although there will be some variability in the timing of individual spikes). Neurons cannot fire at rates greater than about two hundred spikes per second, and this would seem to limit the usefulness of phase locking to a frequency of around 200 Hz. However, even

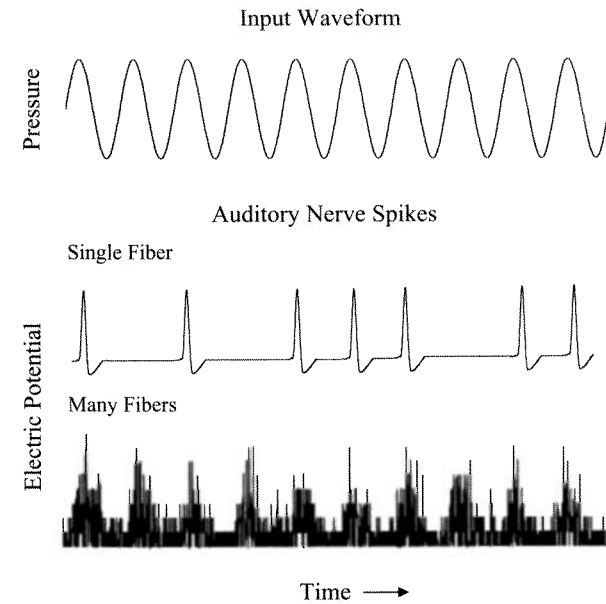


FIG. 4.18. An illustration of the auditory nerve activity in response to a 250-Hz pure tone (top panel). The middle panel shows the pattern of spikes that may be produced by a single auditory nerve fiber. The lower panel represents the combined spikes produced by 500 nerve fibers (or by one nerve fiber stimulated 500 times in succession). Note that, although there is some variability in the phase at which a neuron fires from cycle to cycle, the periodicity of the waveform is well represented across the array of fibers.

if an individual fiber cannot respond at a sufficiently high rate to represent every cycle of the incoming waveform, information may be combined across neurons to represent the frequency of high-frequency tones. If one neuron produces spikes on the first, third, fifth, etcetera cycle of the incoming pure tone, another might produce spikes on the second, fourth, sixth, etcetera cycle of the sound wave. The combined firing patterns of the two neurons reflect each cycle of the pure tone. In reality, neurons are not nearly as regular as this simplistic example might suggest, but the principle holds. Figure 4.18 illustrates a typical pattern of phase locked spikes for a single fiber, and the pattern of activity averaged across many fibers. Remember that each neuron is responding to the temporal pattern of vibration of the place on the basilar membrane to which it is connected. If pure tones with frequencies of 100 Hz and 500 Hz are presented, some neurons will phase lock to 100 Hz and some neurons will phase lock to 500 Hz, reflecting the separation of these components on the basilar membrane.

There is a limit to how rapidly the electric potential can fluctuate in an inner hair cell, and, at high stimulation frequencies, the potential does not vary up and

down with every period of the waveform. Consequently, auditory nerve fibers show a tendency to produce spikes at a particular phase of the sound waveform up to a maximum frequency of about 5000 Hz (Johnson, 1980). Up to this frequency, therefore, spectral information may be represented partly by a temporal code (the time between consecutive spikes). Above 5000 Hz, the spikes are not related to a particular phase in the *fine structure* of the waveform. However, neurons also tend to phase lock to the *envelope* of a sound, so they produce spikes at a particular phase of amplitude modulation, for example (Joris & Yin, 1992). They can do this even if the *carrier* frequency is greater than 5000 Hz. Therefore, phase locking may be a general way of representing the periodicity of waveforms such as complex tones. We see in Chapter 7 how phase locking may be the basis for pitch perception, and in Chapter 9 why phase locking is necessary for the precise localization of sounds.

4.5 FROM EAR TO BRAIN

Let's recap what we have discussed so far. Sound enters the ear canal and causes the eardrum to vibrate. These vibrations are transmitted to the cochlea by the bones in the middle ear. Vibrations of the oval window cause pressure changes in the cochlea that cause the basilar membrane to vibrate, with different places on the basilar membrane responding to different frequencies. Vibrations on the basilar membrane are detected by the inner hair cells, which cause electrical activity (spikes) in the auditory nerve. Now that the acoustic information has been represented in terms of neural activity, the hard part can begin. The task of analyzing the information and separating and identifying the different signals is performed by the brain.

4.5.1 Ascending Auditory Pathways

The auditory nerve carries the information about incoming sound from the cochlea to the *cochlear nucleus*, a collection of neurons in the *brainstem*. The brainstem is a "primitive" part of the brain on top of the spinal cord. The information is passed (via synapses) to a number of other brainstem nuclei: the *superior olivary complex* (or *superior olive*; neural connections from the opposite, or *contralateral*, ear come at this stage), the *lateral lemniscus*, and the *inferior colliculus* (see Fig. 4.19). At each stage in the pathway, information about the sound is processed by a network of neurons, with parallel projections carrying different types of information up the brainstem. Each nucleus may contain many different types of neurons, with varying properties. For example, in the cochlear nucleus, some neurons are similar to auditory nerve fibers (Fig. 4.16), whereas others fire at the onset of a continuous stimulus, and then produce little response. These onset neurons receive input from neurons with a wide range of characteristic frequencies (hence they have broad tuning curves). Some neurons have a response that builds up relatively slowly over

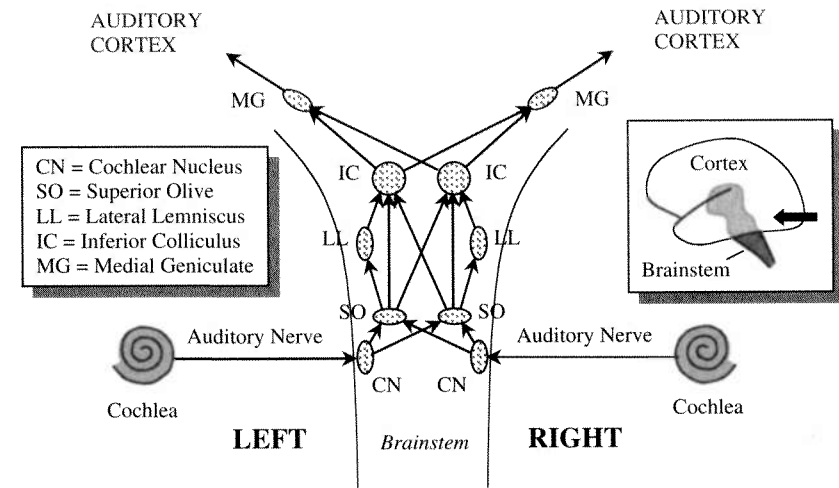


FIG. 4.19. A highly simplified map of the ascending auditory pathways, showing the main neural connections in the brainstem. The viewpoint is toward the back of the brain, as indicated by the arrow on the illustration to the right (cerebellum removed).

time: These neurons may receive inhibition from other neurons that suppresses their response at onset. Other neurons have the tendency to produce spikes at regular intervals, irrespective of the stimulation frequency. The properties of a given neuron depend on the complex excitatory and inhibitory connections from other neurons, and on the particular physiology of the neuron itself (for instance, if it is fast or slow to respond to inputs from other neurons).

Brainstem nuclei analyze and decode the auditory signal using these diverse neural populations. As we discuss later, it is possible that the superior olive is involved in sound localization (where a sound source is located in space), and the inferior colliculus may be involved in pitch perception, among other things. Although many of the properties of these neural populations are well documented, to understand their functions with respect to our perceptual abilities is a difficult task. It is difficult to relate the properties of individual neurons to an ability that may depend on many thousands or millions of neurons working together. Because many of the suggested functions are quite controversial (and depend upon detailed analyses that are beyond the scope of this chapter), I do not dwell on what particular neurons may or may not be doing. Instead, I discuss some of the less speculative proposals where they are relevant in later chapters.

Nerve fibers from the inferior colliculus synapse with the *medial geniculate body*, which is part of the *thalamus* in the *midbrain* (just about in the center of the head). The thalamus acts as a sort of relay station for sensory information. Nerve fibers from the medial geniculate body project to the *auditory cortex*, which is part

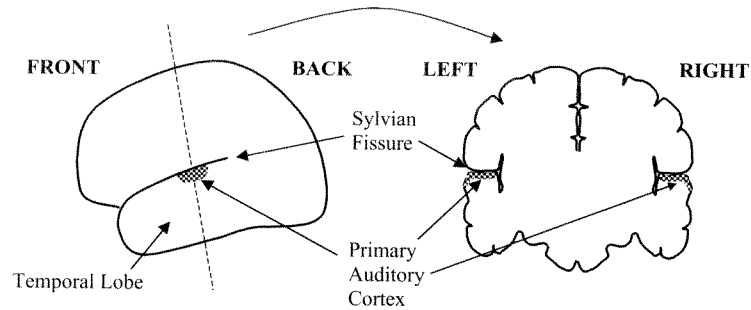


FIG. 4.20. The location of the primary auditory cortex on the cerebral cortex, shown from the side (left), and in a cross-section taken along the dashed line (right).

of the *cerebral cortex*. The cerebral cortex is the wrinkly bit you see when you look at a brain. It covers most of the surface of the brain, and is involved in high-level thought processes as well as basic sensory and motor functions. The wrinkles are present because the cortex is a relatively thin sheet of neurons (only 3 mm thick) that is greatly convoluted, so that the total area is large. Some regions of the cerebral cortex (the primary visual, auditory, and somatosensory areas) receive input from the sensory systems. The primary motor cortex projects (relatively) directly to the muscles. Regions adjacent to the primary areas (association cortex) carry out further processing on the sensory input, and integrate information between the senses. Broadly speaking, the farther is the region from the primary area, the more holistically it processes information (e.g., identifying a sentence as opposed to identifying a spectral feature).

The auditory cortex is located at the top of the temporal lobe, hidden in a crease (or *fissure*) in the cerebral cortex called the *Sylvian fissure* (see Fig. 4.20). The auditory cortex consists of a primary field (AI) and several adjacent fields. The primary field contains a *tonotopic* representation, in which neurons with similar characteristic frequencies are arranged in strips. The same may be true of the adjacent fields, so that there are multiple representations of the cochlea. However, the properties of cortical neurons are much more complex than this description suggests. Cortical neurons perform a detailed analysis of the individual *features* in the auditory signal. The response is often brief and synchronized with peaks in the sound waveform. Some cortical neurons are most sensitive to a particular range of sound levels, and their activity actually reduces as level is increased (or decreased) beyond this range. Many cortical neurons have complex binaural properties (reflecting input from the two ears that has been processed in the brainstem). Some cortical neurons have complex spectral properties with “multi-peaked” tuning curves. Some cortical neurons show a preference for particular *changes* in frequency over time (e.g.,

an increase in frequency produces a higher response than a decrease in frequency). The selectivity of cortical neurons for specific acoustic features may reflect stages in sound identification. Of course, most of our knowledge of the auditory cortex has been derived from neurophysiological experiments on non-human mammals. At this relatively high-level stage in processing, there may be considerable differences between species which reflect the priority of sounds that are important for different species.

Somewhere away from the auditory cortex, probably in the temporal and parietal lobes, the signal from the auditory system is finally identified as a specific word, melody, object, etcetera, and the information is linked to that from other sensory systems to provide a coherent impression of the environment.

4.5.2 Descending Auditory Pathways

Information flows through the auditory system in not just one way, from the ear to the brain. There are also *descending* auditory pathways, carrying information from higher auditory centers to lower auditory centers, even as far as the cochlea itself. The *olivocochlear bundle* contains fibers that originate in the ipsilateral (same side) and contralateral (opposite side) superior olivary complexes. These *efferent* (i.e., from the brain) fibers travel down the auditory nerve and synapse in the cochlea. Some synapse on the axons of the *afferent* (i.e., to the brain) fibers innervating the inner hair cells, and others synapse directly on the outer hair cells. Those that synapse on the outer hair cells can control the motion of the basilar membrane, to some extent. Stimulation of the olivocochlear bundle has been shown to suppress the motion of the membrane (see Guinan, 1996, for a discussion of the role of the olivocochlear efferents).

There are also descending pathways to the cochlear nucleus, which originate mainly from the superior olivary complex, and from the lateral lemniscus and inferior colliculus. In addition, there are descending pathways from the auditory cortex to the medial geniculate body, and to the inferior colliculus. A complete chain of connections may exist from the auditory cortex, through the brainstem nuclei, to the cochlea itself (Pickles, 1988, Chapter 8). It seems that the auditory system is designed so that higher auditory, and perhaps cognitive, centers can exert control on the activity of lower auditory centers, and thus influence the processing of sound.

4.6 SUMMARY

This chapter covers the main stages in the process by which pressure variations in the air (sound waves) are converted into electrical activity in neurons in the auditory system. On a basic level, this is how your ears work. The role of the

cochlea in performing a spectral analysis of the sound has been described, but such is the importance of this mechanism that we return to it in the next chapter.

1. Sound enters the ear through an opening in the pinna leading to the ear canal, which ends with the eardrum. Vibrations at the eardrum are transformed into pressure variations in the cochlear fluids by three tiny bones, the malleus, incus, and stapes.

2. The cochlea is a tube coiled up into a spiral, divided along its length by two membranes, Reissner's membrane and the *basilar membrane*. Pressure variations in the cochlear fluids cause the basilar membrane to vibrate in a wave-like motion traveling from base to apex (the *traveling wave*).

3. The mechanical properties of the basilar membrane vary along its length, so that the different frequency components of a sound cause different parts of the basilar membrane to vibrate, high frequencies toward the base, low frequencies toward the apex. Each place on the basilar membrane is tuned to a particular *characteristic frequency*. The basilar membrane as a whole behaves as a bank of overlapping band-pass filters (*auditory filters*). In this way, the basilar membrane extracts the *spectrum* of a sound.

4. Vibration of the basilar membrane causes a shearing force between the basilar membrane and the overlying tectorial membrane. This causes the stereocilia on the tops of the hair cells to sway from side to side at the same rate as the basilar membrane vibration. Motion of the stereocilia of the inner hair cells produces an electrical change in the cell (*depolarization*), leading to the release of a chemical neurotransmitter that induces electrical impulses (*spikes*) in adjacent auditory nerve fibers.

5. Each auditory nerve fiber is connected to a particular place in the cochlea, and represents the activity at that place. Because each place in the cochlea is most sensitive to a characteristic frequency, each neuron in the auditory nerve is also most sensitive to a characteristic frequency.

6. Neural firing rates increase as sound level increases, until they saturate at a firing rate of about 200 spikes per second. *High spontaneous rate* fibers start to increase their firing at low sound levels (they have low thresholds), but they also saturate at fairly low levels (up to about 60 dB SPL). *Low spontaneous rate* fibers have higher thresholds, and much higher saturation levels (maybe 100 dB SPL or more).

7. Because the inner hair cells depolarize when their stereocilia are bent in one direction (away from the center of the cochlea), nerve fibers tend to fire at a particular *phase* of the basilar membrane vibration for frequencies up to about 5000 Hz. This is called *phase locking*.

8. Information about a sound is represented in the auditory nerve in two ways; in terms of firing *rate* (nerve fibers represent the *magnitude* of vibration at different places on the basilar membrane) and in terms of phase locking or firing *synchrony*

(nerve fibers represent the *temporal pattern* of vibration at different places on the basilar membrane).

9. Information travels up the auditory nerve, through a chain of nuclei in the brainstem (the information receives processing at each stage), before being passed on to the auditory cortex. There also are descending neural pathways that allow higher centers to control lower centers, and even the basilar membrane itself through efferent neural connections to the outer hair cells.

4.7 READING

Pickles' book is beginning to look dated, but still provides an excellent introduction to auditory physiology:

Pickles, J. O. (1988). *An introduction to the physiology of hearing* (2nd ed.). London: Academic Press.

I also found the following more recent publication very useful:

Møller, A. R. (2000). *Hearing: Its physiology and pathophysiology*. New York: Academic Press.

Yost provides a good overview of the central auditory system:

Yost, W. A. (2000). *Fundamentals of hearing: An introduction*. New York: Academic Press. Chapter 15.

For detailed accounts of auditory neurophysiology:

Popper, A. N., and Fay, R. R. (Eds.). (1992). *The mammalian auditory pathway: Neurophysiology*. New York: Springer-Verlag.

Oertel, D., Fay, R. R., and Popper, A. N. (Eds.). (2002). *Integrative functions in the mammalian auditory pathway*. New York: Springer-Verlag.