Hearing: Physiology and Psychoacoustics

Questions to Contemplate

Think about the following questions as you read this chapter. By the chapter’s end, you should be able to answer and discuss them.

- What are the physical and psychological qualities of sound?
- How is sound energy turned into neural firing for the brain to interpret?
- How does the brain encode pitch and loudness?
- How is hearing loss caused, and what can be done about it?

Are humans really visual animals? We are reminded of the importance of vision whenever we close our eyes or awaken in the night, because so much of what we sense and know about our environment is suddenly gone. In contrast, most people never get such reminders of the importance of hearing. Ears are always open, and we can hear perfectly well in the dark. When we enter a dark place, we don’t need to wait half an hour to be able to hear the soft sounds that tell us we’re not alone. We can hear things when our nose is not pointed at the source of the sound, and we can even hear around obstacles and corners. For better or worse, we can often hear through barriers that light cannot penetrate. For all these reasons, it is easy to take hearing for granted.

Try to imagine a world, though, where nothing makes a sound. You watch a movie or online video, and the sound is always off. A person cries out, and you don’t help because you cannot hear the cry. Without hearing, the world is a more dangerous place. You are relatively isolated by day and nearly totally isolated by night when your sense of vision is compromised by darkness.

The Function of Hearing

The next three chapters are all about hearing. In this chapter we cover the basics: the nature of sound, the anatomy and physiology of the auditory system, and how we perceive the two fundamental sound qualities—loudness and pitch. We conclude this chapter by looking at some of the ways in which hearing can be impaired and what we can do to avoid and overcome these impairments. In Chapter 10, we will move on to discuss some of the ways we use acoustic information to learn about our surroundings. Then, in Chapter 11, we will cover the higher-level auditory functions that we use when we’re listening to speech and music.

Many fundamental principles in hearing apply to all of the senses. However, each sense developed at different periods in our evolutionary history and in response to different environmental challenges. So, although you should find that much of what you’ve learned thus far will help you to understand hearing, you will also be impressed by ways biology has provided some very different (and very clever) solutions to the challenges of sensing and interpreting sound.
What Is Sound?

Sounds are created when objects vibrate. The vibrations of an object (the sound source) cause molecules in the object’s surrounding medium (for humans, usually the Earth’s atmosphere) to vibrate as well. This vibration causes pressure changes in the medium (Figure 9.1). These pressure changes are best described as waves, and they are similar to the waves on a pond caused by dropping a rock into the water. Water molecules displaced by the rock do not themselves travel very far, but the pattern of displacement will move outward from the source until something (the shore, a boat, a swimming duck, or anything else) gets in the way. Although the patterns of pond and sound waves do not change as they spread out, the initial amount of pressure change is dispersed over a larger and larger area as the wave moves away, so the wave becomes less prominent as it moves farther from its source.

Sound waves travel through different media at different speeds, moving faster through denser substances. For example, the speed of sound through air is about 340 meters per second (somewhat less than a mile per second when moving over land, and a bit faster on muggy days), but the speed of sound through water is about 1500 meters per second. Light waves move through air almost a million times faster than sound waves do. This is why you see lightning before hearing thunder—the difference is almost 5 seconds (4¾) per mile.

Basic Qualities of Sound Waves: Frequency and Amplitude

As we’ve seen, sound waves that we hear are simply fluctuations in air pressure across time. The magnitude of the pressure change in a sound wave—the difference between the highest pressure and the lowest pressure of the wave—is called the amplitude or intensity (Figure 9.2). Pressure fluctuations may be very close together or spread apart over longer periods. For light waves, we usually describe the pattern of fluctuations by measuring the distance between peaks in the waves—that is, the wavelength. Sound waves also have wavelengths, but we describe their patterns by noting how quickly the pressure fluctuations rate is known as the frequency of the wave (see Figure 9.2). To see an example of frequency, dangle a thread in front of a stereo. As the speaker creates fluctuations in air pressure, the thread waves back and forth. The tempo of this waving is the thread’s frequency. Sound wave frequencies are measured by these back-and-forth cycles, and the unit of measure is called a hertz (Hz), where 1 cycle per second equals 1 Hz. For example, the pressure in a 500-Hz wave goes from its highest point down to its lowest point and back up to its highest point 500 times every second.

Table 9.1

<table>
<thead>
<tr>
<th>Decibel levels that correspond to different sound pressure ratios</th>
<th>Ratio relative to 0.0002 dyne/cm² (µPa)</th>
<th>dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>20</td>
<td>26</td>
<td>60</td>
</tr>
<tr>
<td>100</td>
<td>40</td>
<td>100</td>
</tr>
<tr>
<td>1000</td>
<td>60</td>
<td>10,000</td>
</tr>
<tr>
<td>10,000</td>
<td>80</td>
<td>100,000</td>
</tr>
<tr>
<td>100,000</td>
<td>100</td>
<td>1,000,000</td>
</tr>
</tbody>
</table>

Figure 9.9 illustrates the relationship between sound intensity and perceived loudness. The psychological aspect of sound related to perceived intensity (amplitude).

Figure 9.3 illustrates the relationship between sound frequency and pitch. The psychological aspect of sound related mainly to the fundamental frequency.

Table 9.2

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Hearing range</th>
<th>Frequency (Hz)</th>
<th>Hearing range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>20,000</td>
<td>20,000</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>40,000</td>
<td>40,000</td>
</tr>
<tr>
<td>100</td>
<td>40</td>
<td>100,000</td>
<td>100,000</td>
</tr>
<tr>
<td>1000</td>
<td>60</td>
<td>1,000,000</td>
<td>1,000,000</td>
</tr>
</tbody>
</table>

Just as the amplitude and wavelength of light waves correspond to perceptual qualities in vision (brightness and color, respectively), the amplitude and frequency of sound waves are closely related to auditory characteristics. Amplitude is associated with the perceptual quality of loudness; the more intense a sound wave is, the louder it will sound. Frequency is associated with pitch, low frequency sounds correspond to low pitches (e.g., notes played on a tuba), and high frequency sounds correspond to high pitches (e.g., the high notes from a piccolo). We will have much more to say about the relationships between frequency and pitch later in this chapter.

In Chapter 2, you learned how visible light makes up only a small portion of the much broader range of electromagnetic energy; similarly, human hearing uses a limited range of the frequencies present in electromagnetic energy. If you are relatively young and you’ve been careful about your exposure to loud sounds, you may be able to detect sounds that vary from about 20 to 20,000 Hz (Figure 9.3). Some animals hear sounds that have lower and higher frequencies than those heard by humans. In general, larger animals are better at hearing low frequencies, and smaller animals are better at using high frequencies. Elephants hear vibrations at very low frequencies that help them detect the presence of large animals such as other elephants. Dogs can be called with whistles that emit sounds at frequencies too high for humans to hear, and the sonar systems used by some bats use sound frequencies above 60,000 Hz.

Humans hear across a very wide range of sound intensities. The intensity ratio between the faintest sound humans can detect and the loudest sounds that do not cause serious damage to human ears is more than one million. To describe differences in intensity across such a broad range, sound levels are measured on a logarithmic scale using units called decibels (dB). Decibels define the difference between two sounds in terms of the ratio between sound pressures. Each 10 dB sound pressure ratio is equal to 20 dB, so a 100:1 ratio is equal to 40 dB. The equation for defining decibels is dB = 20 log(p0/p).

The variable p corresponds to the sound pressure level. The constant p0 is a reference pressure and is typically defined in auditory research contexts to be 0.0002 dyne per square centimeter (dyne/cm²), and levels are defined in terms of dB SPL (sound pressure level). The level 0.0002 dyne/cm² is close to the minimum pressure that can be detected at frequencies for which hearing is most sensitive, and decibel values greater than zero describe the ratio between a sound being measured and 0.0002 dyne/ cm². The range of human hearing extends from 0 to more than 120 dB SPL, and as shown in Table 9.1, this decibel range corresponds to a ratio of greater than 1,000,000:1.
The low-frequency component is called the fundamental frequency. All the other harmonics have frequencies that are integer multiples of the fundamental.

A complex sound is best described in a spectrum, which displays the energy levels of harmonic components. A complex sound can be described as a combination of sine waves. The spectrum shows how much energy, or amplitude, is present at each frequency. The fundamental frequency is the lowest-frequency component of the sound. All the other harmonics have frequencies that are integer multiples of the fundamental.

The shape of the spectrum (spectral shape) is one of the most important qualities that distinguish different sounds. The properties of sound sources determine the spectral shapes of sounds, and these shapes help us to identify sound sources. For example, Figure 9.6 shows that different musical instruments produce different spectral shapes. Different musical instruments playing the same note (the same fundamental frequency, abbreviated “f”) sound different. C4 = middle C.

Web Activity 9.2: Structure of the Auditory System
Now that you know what sound is, we can examine how sounds are detected and recognized by the auditory system. The sense of hearing has evolved over millions of years to be able to do some amazing things. We are about to describe quite a few anatomical structures that are essential to understanding how sequences of tiny air pressure changes are turned into meaningful sound perception. The discussion may occasionally be a bit confusing, but if you consult the figures and Web Activity 9.2: Structure of the Auditory System often, you will soon know the parts and how they fit together.

Basic Structure of the Mammalian Auditory System
The discussion may occasionally be a bit confusing, but if you consult the figures and Web Activity 9.2: Structure of the Auditory System often, you will soon know the parts and how they fit together.

Sine Waves and Complex Sounds
In Chapter 1, we introduced to one of the simplest kinds of sounds: a sine wave, which is often called a pure tone. Any sound, even those as complex as the sounds produced by musical instruments, human speech, and city traffic, can be described as a combination of sine waves. (See Web Activity 9.1: What We Hear)

A complex sound is best described in a spectrum (plural spectra), which displays the relative energy (intensity) present at each frequency. Many common sounds have harmonic spectra, illustrated in Figure 9.5. A spectrum displays the amplitude for each frequency present in a sound. Each signal is shown as a waveform (A) and as a spectrum (B).
pinna The outer, funnel-like part of the ear.

ear canal The canal that conducts sound vibrations from the pinna to the tympanic membrane and prevents damage to the tympanic membrane.

tympanic membrane The eardrum, a thin sheet of skin at the end of the outer ear canal. The tympanic membrane vibrates in response to sound.

outer ear The external sound-gatheror portion of the ear, consisting of the pinna and the ear canal.

middle ear An air-filled chamber containing the middle bones, or ossicles. The middle ear conveys and amplifies vibration from the tympanic membrane to the oval window.

ossicle Any of three tiny bones of the middle ear: malleus, incus, and stapes.

malleus The most exterior of the three ossicles. The malleus receives vibration from the tympanic membrane and is attached to the incus.

incus The middle of the three ossicles, connecting the malleus and the stapes.

stapes The most interior of the three ossicles. Connected to the incus on one end, the stapes presses against the oval window of the cochlea on the other end.

oval window The flexible opening to the cochlea through which the stapes transmits vibration to the fluid inside.

inner ear A hollow cavity in the temporal bone of the skull, and the structures within this cavity: the cochlea and the semicircular canals of the vestibular system.

tensor tympani The muscle attached to the malleus. Tensing the tensor tympani decreases vibration.

stapedius The muscle attached to the stapes. Tensing the stapedius decreases vibration.

Outer Ear

Sounds are first collected from the environment by the pinna (plural pinnae), the curvy structure on the side of the head that we typically call an ear. Only mammals have pinnae, and they vary wildly in shape and size across species and vary less dramatically across individuals within species (Figure 9.7). As we will see in Chapter 10, the shape of the pinnae plays an important role in our ability to localize sound sources.

Sound waves are funneled by the pinna into and through the ear canal, which extends about 25 millimeters (mm) into the head (Figure 9.8). The length and shape of the ear canal enhance sound frequencies between about 2000 and 6000 Hz, but the main purpose of the canal is to protect the structure at its end, the tympanic membrane (eardrum), from damage. The tympanic membrane is a thin sheet of skin that moves in and out in response to the pressure changes of sound waves.

It is a common myth that puncturing your eardrum will leave you deaf. While a ruptured eardrum can be excruciating, in most cases a damaged tympanic membrane will heal itself, just as other parts of the skin do. You probably know someone who has had tubes placed through tympanic membranes to remedy ear infections. Nevertheless, it is possible to damage the tympanic membrane beyond repair, so it’s a good idea to follow your mother’s advice to not stick things in your ear.

Middle Ear

Together, the pinna and ear canal make up a division of the auditory system called the outer ear (see Figure 9.8). The tympanic membrane is the border between the outer ear and the middle ear, which consists of three tiny bones, the ossicles, that amplify sound waves. The first ossicle, the malleus, is connected to the tympanic membrane on one side and to the second ossicle, the incus, on the other. The incus is connected in turn to the third ossicle, the stapes, which transmits the vibrations of sound to the oval window, another membrane, which forms the border between the middle ear and the inner ear.

The ossicles are the smallest bones in the human body, and they amplify sound vibrations in two ways. First, the joints between the bones are hinged in ways that make them work like levers: a modest amount of energy on one side of the fulcrum (joint) becomes larger on the other. This lever action increases the amount of pressure change by about a third. The second way the ossicles increase the energy transmitted to the inner ear is by concentrating energy from a larger to a smaller surface area. The tympanic membrane, which moves the malleus, is about 18 times as large as the oval window, which is moved by the stapes (see Figure 9.8). Therefore, pressure on the oval window is magnified 18 times relative to the pressure on the tympanic membrane. This is the same principle that makes stilletto heels a danger to wood floors (think of the tympanic membrane as the heel of the foot and the oval window as the tip of the stiletto), in contrast to the way snowshoes keep feet on top of the snow.

Amplification provided by these physical properties (leverage and different surface areas) is essential to our ability to hear faint sounds, because the inner ear, as we will see in a moment, is made up of a collection of fluid-filled chambers. Because it takes more energy to move liquid than it does to move air, this fluid creates a mismatch. If sound waves were transmitted to the oval window directly, many would simply bounce back without moving the oval window and the liquid behind it at all.

Ossicles play an important role for loud sounds, too. The middle ear has two muscles: the tensor tympani (attached to the malleus) and the stapedius (attached to the stapes) (see Figure 9.8). As might be expected because they are attached
acoustic reflex A reflex that protects the ear from intense sounds, via conduction of the stapedius and tensor tympani muscles.

cochlea A spiral structure of the inner ear containing the organ of Corti.

tympanic canal One of three fluid-filled passages in the cochlea. The tympanic canal extends from the round window at the base of the cochlea to the helicotrema at the apex. Also called scala tympani.

vestibular canal One of three fluid-filled passages in the cochlea. The vestibular canal extends from the oval window at the base of the cochlea to the helicotrema at the apex. Also called scala vestibuli.

middle canal One of three fluid-filled passages in the cochlea. The middle canal is sandwiched between the tympanic and vestibular canals and contains the cochlear partition. Also called scala media.

helicotrema The opening that connects the tympanic and vestibular canals at the apex of the cochlea.

Reissner’s membrane A thin sheet of tissue separating the vestibular and middle canals in the cochlea.

basilar membrane A plate of fibers that forms the base of the cochlear partition and separates the middle and tympanic canals in the cochlea.

cochlear partition The combined basilar membrane, tectorial membrane, and organ of Corti, which are together responsible for the transduction of sound waves into neural signals.

round window A soft area of tissue at the base of the tympanic canal that releases excess pressure remaining at the base of the tympanic canal that is relieved during swallowing, talking, and general body movement, helping to keep the auditory system from being overwhelmed by sounds generated by our own bodies.

inner ear The inner ear is an impressive feat of evolution. It is here that the minute changes in sound pressure available in the environment are translated into neural signals that inform the listener about the world. The function of the inner ear with respect to sound waves is roughly analogous to that of the retina with respect to light waves in vision: it translates the information carried by the waves into neural signals.

COCHLEAR CANALS AND MEMBRANES The major structure of the inner ear is the cochlea (from the Greek kochlea, “snail”), a tiny coiled structure embedded in the temporal bone of the skull (see Figure 9.8). Rolled up, the cochlea is the size of a baby pea, about 4 mm in diameter in humans. Uncoiled, it would be a tube almost ten times as long—about 35 mm or 1.4 inches. The cochlea is filled with watery fluids in three parallel canals (Figure 9.9): the tympanic canal (or scala tympani), the vestibular canal (or scala vestibuli), and the middle canal (or scala media). The tympanic and vestibular canals are connected by a small opening, the helicotrema, and these two canals are effectively wrapped around the middle canal. Think of the tympanic and vestibular canals as one long, skinny balloon unwound, and the middle canal as another long balloon that is sandwiched, lengthwise, between the two halves of the first balloon. The three canals of the cochlea are separated by two membranes (see Figure 9.9): Reissner’s membrane between the tympanic and middle canals and the basilar membrane between the middle and tympanic canals. Strictly speaking, the basilar membrane is not really a membrane, because it is not a thin, pliable sheet like the tympanic membrane, the oval window, or Reissner’s membrane. Instead, it is a plate made up of stiff fibers. The basilar membrane forms the base of the cochlear partition, a complex structure through which sound waves are transduced into neural signals. Vibrations transmitted through the tympanic membrane and middle-ear bones cause the stapes to push and pull the flexible oval window in and out of the vestibular canal at the base of the cochlea. This movement of the oval window causes waves of pressure changes, called traveling waves, to flow through the fluid in the vestibular canal, in much the same way that the membrane of a loudspeaker moves air to create sound waves. Because the cochlea is a closed system, changes in pressure cannot spread out in all directions. Instead, a displacement, or “bulge,” forms in the vestibular canal and extends from the base of the cochlea down to the apex (look ahead to Figure 9.12). If sounds are extremely intense, any pressure that remains at the apex passes through the helicotrema and back to the cochlear base through the tympanic canal, where it is relieved by stretching yet another membrane, called the round window (see Figure 9.9).

Because the vestibular and tympanic canals are wrapped tightly around the middle canal, when the vestibular canal bulges out, it puts pressure on the middle canal. This pressure has the effect of displacing the basilar membrane (which, recall, lies at the bottom of the middle canal), pushing down when the vestibular-canal bulge is created. The organ of Corti is made up of a scaffold of cells that support specialized neurons called hair cells. Axons and dendrites of these neurons synapse on the hair cells. Organ of Corti A structure on the basilar membrane of the cochlea that is composed of hair cells and dendrites of auditory nerve fibers.

hair cell Any cell that has stereocilia for transducing mechanical movement in the inner ear into neural activity sent to the brain. Some hair cells also receive inputs from the brain.
auditory nerve  A collection of neurons that convey information from hair cells in the cochlea to the brain stem (afferent neurons) and from the brain stem to the hair cells (afferent neurons).

tectorial membrane  A gelatinous structure, attached on one end, that extends into the middle canal of the cochlea, floating above inner hair cells and touching outer hair cells.

INNER AND OUTER HAIR CELLS  Like photoreceptors in the retina, hair cells are specialized neurons that transduce one kind of energy (in this case, sound pressure) into another form of energy (neural firing). Hair cells in the vestibular organs also report head movements to the brain, as you will learn in Chapter 12. Deflection of a hair cell’s stereocilium causes a change in voltage potential that initiates the release of neurotransmitters, which in turn encourages firing of auditory nerve fibers that have dendritic synapses on hair cells (Eggermont, 2017). However, it is the differences between photoreceptors and hair cells that are the most interesting.

While the retina has almost 100 million photoreceptors, the cochlea has only about 14,000 hair cells. Although outnumbered, stereocilia of hair cells blow away the competition when it comes to speed and sensitivity. Listeners can detect differences between onsets of two sounds as small as 1 millisecond (ms) (Zera and Green, 1993), and they can detect gaps between sounds as brief as 2–3 ms (Schneider and Hamstra, 1999). In contrast, when we watch a movie, pictures shown at 24 frames per second (over 40 ms apart) appear continuous to the visual system. Hair cells are not only extremely fast, but also extremely sensitive. It may take 30 minutes for our eyes to fully adjust to a dark theater, but our ears are always ready for the slightest sound.

Recall that the shortest stereocilia are in front of slightly taller stereocilia that are in front of still-taller stereocilia (Figure 9.11). Each stereocilium is connected to its neighbor by a tiny filament called a tip link, so the stereocilia connected by tip links bend together as a set when deflected by the shearing motion of the tectorial membrane. Because what happens next is very difficult to observe—tiny parts of tiny structures atop single hair cells—what follows is only a current hypothesis. When a stereocilium deflects, the tip link pulls on the taller stereocilium in a way that opens an ion pore somewhat like opening a gate for just a tiny fraction of a second. This action permits potassium ions (K+) to flow rapidly into the hair cell, causing rapid depolarization (see Figure 9.11B and C). In turn, depolarization leads to a rapid influx of calcium ions (Ca2+) and initiation of the release of neurotransmitters from the base of the hair cell to stimulate dendrites of the auditory nerve (Fettisplace and Hackney, 2006; Hudspeth, 1997).

Figure 9.9  When vibration causes a displacement along the cochlear partition (see Figure 9.9, lower right), the tectorial membrane and hair cells move in opposite directions (experience shear), and the deflection of stereocilia during this action results in the release of neurotransmitters.

Figure 9.10  Stereocilia regulate the flow of ions into and out of hair cells. (A) This photomicrograph shows the thread-like tip links that connect the tip of each shorter stereocilium to its taller neighbor. (B) C Bending the stereocilia atop a hair cell opens the ion pores, permitting a rapid influx of potassium ions (K+) into the hair cell, and this depolarization opens channels that allow calcium ions (Ca2+) to enter the base of the hair cell, causing the release of neurotransmitters into the synapse between the hair cell and an afferent auditory nerve fiber. (A courtesy of A. J. Hudspeth.)
The cochlea narrows from base to apex, but the basilar membrane inside actually spreads across different sound frequencies in the sound wave. The wider end toward the apex is most sensitive to lower frequencies. Here the cochlea is illustrated as if it were uncoiled (A), and the shapes of the traveling waves for different frequencies of vibration are shown (B).

**Coding of Amplitude and Frequency in the Cochlea**

Now that we know more about how ears work, we can return to the two fundamental characteristics of sound waves, amplitude and frequency, and learn how they are encoded by the cochlea.

If the amplitude of a sound wave is increased, the tympanic membrane and oval window move farther in and out with each pressure fluctuation. The result is that the bulge in the vestibular canal becomes bigger, which causes the cochlear partition to move farther up and down, which causes the tectorial membrane to shear across the organ of Corti, moving the stereocilia atop hair cells back and forth. The pivoting of the stereocilia initiates rapid depolarization (followed by equally fast hyperpolarization) that results in spurts of neurotransmitter released into synapses between the hair cells and dendrites of auditory nerve fibers. These neurotransmitters initiate action potentials in the auditory nerve fibers that are carried to the brain. And that’s all there is to it!

Coding for frequency is a bit trickier. Earlier we said that the cochlear partition responds to different frequencies in the sound wave. This statement is true as far as it goes, but it does not tell the whole story; because different parts of the cochlear partition are displaced to different degrees by different sound wave frequencies. High frequencies cause the largest displacements closer to the oval window, near the base of the cochlea. Lower frequencies cause displacements farther away and nearer the apex. In other words, different places on the cochlea are “tuned” to different frequencies. This tuning is known as the place code for sound frequency.

Cochlear tuning to frequency is caused, in large part, by differences in the structure of the basilar membrane along the length of the cochlea (Figure 9.12). The cochlea narrows from base to apex, but the basilar membrane inside actually widens toward the apex. In addition, the basilar membrane is thick at the base and becomes thinner as it widens. While the basilar membrane gets thinner and wider along its length, the cochlea separates frequencies like an acoustic prism. Higher frequencies affect the narrower, stiffer regions of the basilar membrane near the base more, and lower frequencies cause greater displacements in the wider, more flexible regions near the apex.

**Mechanical Displacement**

The opening of ion pores that results from the direct connection between stereocilia via tip links is known as mechanoelectrical transduction, which is responsible for both the extreme speed and the sensitivity of hair cells. Unlike the case in vision, depolarization in hearing does not await a cascade of biochemical processes such as those in photosactivation. Mechanoelectrical transduction is also extremely sensitive: ion pores open when deflection is as little as 1 nanometer (nm), roughly the diameter of a single atom.

Mechanoelectrical transduction is responsible for both the extreme speed and the sensitivity of hair cells. Unlike the case in vision, depolarization in hearing does not await a cascade of biochemical processes such as those in photosactivation. Mechanoelectrical transduction is also extremely sensitive: ion pores open when deflection is as little as 1 nanometer (nm), roughly the diameter of a single atom.

In the case of the basilar membrane, it is stiffer and most sensitive to higher frequencies. The wider end toward the apex is most sensitive to lower frequencies. Here the cochlea is illustrated as if it were uncoiled (A), and the shapes of the traveling waves for different frequencies of vibration are shown (B).

**Coding for Frequency**

Cochlear tuning to frequency is caused, in large part, by differences in the structure of the basilar membrane along the length of the cochlea (Figure 9.12). The cochlea narrows from base to apex, but the basilar membrane inside actually widens toward the apex. In addition, the basilar membrane is thick at the base and becomes thinner as it widens. While the basilar membrane gets thinner and wider along its length, the cochlea separates frequencies like an acoustic prism. Higher frequencies affect the narrower, stiffer regions of the basilar membrane near the base more, and lower frequencies cause greater displacements in the wider, more flexible regions near the apex.
When you look at displacements from sounds of different frequencies along the basilar membrane in Figure 9.12B, they are labeled as traveling waves. This is because the displacement really does travel from base to apex—faster near the base where the basilar membrane is narrower, and slower toward the wider apex. Because it takes time for the traveling wave to travel down the basilar membrane, high-frequency regions (base) are stimulated earlier than lower-frequency regions. This fact will become important when, in Chapter 10, you learn about how you know from what direction sounds are coming.

In addition to this passive, structural way of being tuned to frequency, the cochlea has processes that actively sharpen tuning. Remember that there are two different types of hair cells: inner and outer. Over 90% of the afferent fibers in the auditory nerve—fibers that take information to the brain—synthesize on the 3500 inner hair cells (5–30 auditory nerve fibers listen to each inner hair cell). If the inner hair cells are conveying almost all the information about sound waves to the brain, then what do the 10,500 outer hair cells do? It turns out that most of the nerve fibers that synapse with the outer hair cells are efferent fibers, conveying information from the brain. These efferent fibers play a special role in determining what kind of information is sent on to the brain by the afferent fibers (Fettiplace and Hackney, 2006).

The Auditory Nerve

Now that we’ve covered the mechanics of how the auditory system translates air pressure changes into auditory nerve firing, let’s discuss what we know about auditory nerve (AN) fibers. More specifically, we’ll consider the type of information conveyed by afferent AN fibers from the cochlea to the brain.

Remember that sounds with different frequencies displace different regions of the cochlear partition. Inner hair cells, which provide most of the information to the brain via AN fibers, are lined up single file along the length of the basilar membrane. Put these two pieces of information together, and we can infer that the responses of individual AN fibers to different frequencies should be related to their place along the cochlear partition. Sure enough, when scientists record from individual AN fibers in animals, they find that different fibers selectively respond to different sound frequencies.

This frequency selectivity is clearest when sounds are very faint: at very low intensity levels (even less than 0 dB), an AN fiber will increase firing to only a very restricted range of frequencies. Figure 9.13 shows threshold tuning curves for several AN fibers. To graph one of these curves, a researcher inserts an electrode very close to a single AN fiber and then measures how intense the sine waves of different frequencies must be for the neuron to fire faster than its normal, spontaneous firing rate. The frequency that increases the neuron’s firing rate at the lowest intensity (the lowest point on the threshold tuning curve) is called the neuron’s characteristic frequency (CF).

The sharp tuning measured from outputs of inner hair cells depends greatly upon the unsung heroes of the basilar membrane, the outer hair cells. From Figure 9.14, you can see how outer hair cells are required for threshold tuning curves to be focused on a narrow range of frequencies. Outer hair cells make parts of the cochlear partition stiffer in ways that make the responses of inner hair cells more sensitive (Dong and Olson, 2013) and much more sharply tuned to particular frequencies (Eggermont, 2017).

Up to this point, the ear transduces acoustic energy at different frequencies into a pattern of neural responses seems fairly straightforward. A low-intensity sine wave tone with a certain frequency will cause certain AN fibers to increase their firing rates, while other AN fibers continue to fire at their spontaneous rates. As long as the brain knows which AN fibers have which characteristics, it can interpret the pattern of firing rates across all the AN fibers to determine the frequency of any tone (as long as it is within the range of frequencies picked up by the human cochlea).

Unfortunately, it’s not quite this simple. Almost all sounds in the environment are more complex than single sine waves, and most sounds we hear are also much louder than the very quiet sound waves used to measure threshold tuning curves. So although the previous paragraph captures the gist of how AN fibers code for sound frequencies, we have to do a bit more work to understand how higher-intensity, complex sounds are encoded in the auditory nerve. We will consider two of the specific complications we have to deal with. Then, we’ll look at an additional mechanism, related to timing rather than place, that the auditory system uses to convey low-frequency components of sound waves.

### TWO-TONE SUPPRESSION

The rate at which an AN fiber responds changes when energy is introduced at nearby frequencies. In particular, when a second tone of a slightly different frequency is added, the rate of neural firing for the first tone actually decreases—a phenomenon called two-tone suppression (Figure 9.15). Suppression effects are particularly pronounced when the second supressor tone has a lower frequency than the first tone. In other words, if we’re recording from an AN fiber whose CF is 8000 Hz and we use an 8000-Hz test tone, a 1000-Hz suppressor tone has a greater effect on the neuron’s firing rate than a 15,000-Hz suppressor tone has. You can see how understanding the response of the whole auditory nerve to complex sounds (that is, frequency combinations) is more complicated than simply adding up the responses of individual AN fibers to individual pure tones.
Figure 9.16  Isointensity functions for one AN fiber with a characteristic frequency of 2000 Hz. Tones of varying frequencies are presented at 20, 40, 60, and 80 dB. The neuron fires vigorously against varying frequencies at varying intensities. A map plotting the firing rate of an auditory nerve fiber in response to a sound of constant frequency and intensity. The neuron will also fire at its maximum rate to a 1000-Hz tone if the sound wave is 2000 Hz” because, as Figure 9.16 illustrates, this phenomenon combines with the selectivity of individual AN fibers. Consequently, the frequency sensitivity of the human auditory system as a whole is exquisite across a wide range of intensity levels, despite the coarse rate selectivity of individual AN fibers.

The trade-off is that the firing rates of these fibers quickly reach saturation, so their frequency selectivity is relatively poor when intensity is relatively high. Low-slip AN fibers are more like cones, requiring more energy (higher-intensity sound waves) to start responding, but retaining their frequency selectivity over a broader range of intensity.

In addition to having different AN fibers with different spontaneous rates, the auditory system can accurately determine the frequency of incoming sound waves by integrating information across many AN fibers and using the patterns of firing rates across all of these fibers. Figure 9.17 shows a phase locking plot against sound intensity for six fibers, all of which listen to the same hair cell (remember that dendrites from 5–30 auditory neurons are synapsing with each inner hair cell). To plot these curves, we determine the intensity level of an AN fiber’s CF is slowly raised from 0 dB up to 90 dB. As you can see, the resting rates of some fibers (whose functions are plotted in red) are less than 10 spikes per second. These are low-slip AN fibers.

For moderately intense sounds, such as speech, the brain cannot rely on a single AN fiber to determine the frequency of the tone. For example, we can’t use the rule “if an AN fiber with a characteristic frequency of 2000 Hz is firing very fast, the sound must be 2000 Hz” because, as Figure 9.16 illustrates, this neuron will also fire at its maximum rate to a 1000-Hz tone if the sound wave has a large-enough amplitude. One way the auditory system gets around this problem is to use AN fibers with different spontaneous firing rates. Figure 9.17 shows rate-intensity functions for low-slip AN fibers that have a low rate (less than 30 spikes per second) of spontaneous firing. Low-slip AN fibers require relatively intense sound before they will fire at higher rates.

The auditory system has another way to encode frequency. As Figure 9.18 illustrates, many AN fibers tend to fire action potentials at one particular point in the phase of a sound wave—a phenomenon called phase locking. Phase locking may occur because AN fibers fire when the stereocilia of hair cells move in one direction (e.g., as the basilar membrane moves up toward the tectorial membrane) but do not fire when the stereocilia move in the other direction. Recall from our discussion of mechanoelectrical transduction in stereocilia that the encoding of time is extremely accurate.

The existence of phase locking means that the firing pattern of an AN fiber carries temporal code for the sound wave frequency. For example, if the AN fiber fires an action potential 100 times per second, then downstream neurons listening to the AN fiber can infer that the sound wave includes a frequency component of 100 Hz.

When we consider lower frequencies, temporal coding becomes inconsistent for frequencies higher than 1000 Hz and is virtually absent above 4000 or 5000 Hz. In large part, this inconsistency is due to the refractory period of the AN fiber. For high frequencies, fibers simply cannot produce action potentials quickly enough. The existence of phase locking means that the firing pattern of an AN fiber carries temporal code for the sound wave frequency. For example, if the AN fiber fires an action potential 100 times per second, then downstream neurons listening to the AN fiber can infer that the sound wave includes a frequency component of 100 Hz. When we consider lower frequencies, temporal coding becomes inconsistent for frequencies higher than 1000 Hz and is virtually absent above 4000 or 5000 Hz. In large part, this inconsistency is due to the refractory period of the AN fiber. For high frequencies, fibers simply cannot produce action potentials quickly enough.

The auditory system uses the same principle, the cochlea, to determine the frequency of light waves by integrating information across many AN fibers and using the patterns of firing rates across all of these fibers. Consequently, the frequency sensitivity of the human auditory system as a whole is exquisite across a wide range of intensity levels, despite the coarse rate selectivity of individual AN fibers.
**Phase locking.** The histogram (bottom) shows neural spikes for an AN fiber in response to the same low-frequency sine wave (top) being played many times. Note that the neuron is most likely to fire at one particular phase of each cycle of the sine wave. This phase locking provides a temporal code to sound frequency.

**cochlear nucleus.** The first brainstem nucleus at which all afferent auditory nerve fibers synapse.

**volley principle.** The idea that multiple neurons can provide a temporal code for frequency if each neuron fires at a distinct point in the period of a sound wave but does not fire on every period.

**Sine wave**

**Neuron A**

**Neuron B**

**Neuron C**

**Neuron D**

**Neuron E**

**Total response (A–E combined)**

**Time**

enough to fire on every cycle of the sound. However, multiple neurons could, in principle, encode higher frequencies as a group. For example, four neurons could each fire only once every fourth cycle of a 2000-Hz sound. If the four neurons “took turns,” each would have to fire only 500 times per second to fully encode the 2000-Hz sound in their combined temporal pattern. This idea has a long history (Wever, 1949), and it is referred to as the **volley principle** (Figure 9.19).

According to this hypothesis, neurons sustain a temporal pattern of firing much like the pattern of Revolutionary War-era soldiers firing guns from the front line of a formation while the second and third lines took time to reload. When it comes to temporal coding, neurons along the full length of the cochlea can participate. Even AN fibers with relatively high-frequency CFs encode lower-frequency energy in the temporal pattern of their responses. For example, if you are listening to a fairly loud sound that is a combination of 200- and 8000-Hz sine wave tones, a neuron near the base of the cochlea that becomes widely excited by the 8000-Hz component of the sound will also tend to be phase-locked to the 200-Hz component. This neuron thus carries information about both the high-frequency component (via place coding), because the brain knows the neuron’s CF) and the low-frequency component (via temporal coding).

**Auditory Brain Structures**

The nerve fibers that make up the auditory nerve share cranial nerve VIII with nerve fibers for the vestibular system (which is discussed in Chapter 12); this is why cranial nerve VIII is known not only as the auditory nerve, but also as the vestibulocochlear nerve. All AN fibers initially synapse in the **cochlear nucleus** (Figure 9.20). The cochlear nucleus contains many different types of specialized neurons. Some of these are especially sensitive to onsets of sound at particular frequencies. Some are sensitive to the coincidence of onsets across many frequencies (they fire when multiple frequencies initially begin but stop firing if the sound continues playing). Some cochlear nucleus neurons use lateral inhibition to sharpen the tuning to one frequency by suppressing nearby frequencies—a mechanism reminiscent of that used by retinal ganglion cells to respond to spots of light instead of broad fields of light. Others respond in exactly the same way as the AN fibers that feed them. Some neurons appear to serve as little more than quick relays from the cochlea to the **superior olive**, another brainstem nucleus.

As Figure 9.20 shows, some of the neurons that project from the cochlear nucleus to the superior olives cross over to the opposite side of the brain. Thus, unlike the visual system, where inputs from each visual field remain separate until they have extended a fair distance in the visual cortex, signals from both cochleas reach both sides of the brain after only a single synapse. As we will see in Chapter 10, this direct relay of information across both ears is essential to using tiny differences between the two ears to detect the direction of a sound. Neurons from the cochlear nucleus and superior olive travel up the brainstem to the **inferior colliculus**. Most (but not all) of the input to each inferior colliculus comes from the opposite (contralateral) ear, that is, the left inferior colliculus listens mostly to the right ear, and vice versa.

The **medial geniculate nucleus** of the thalamus is the last stop in the auditory pathway before the cerebral cortex. Like the lateral geniculate of the visual system, there are many more neurons that project from the cortex to the medial geniculate (efferent neurons) than project from the medial geniculate to the cortex (afferent neurons). These efferent connections, some of which presumably convey information back to lower stages of the auditory system, provide further anatomical evidence that sensory systems are two-way streets, in which feedback from the brain is tightly integrated with sensory information flowing up to the brain.
All structures of the auditory system, beginning with the basilar membrane and continuing through the cochlear nucleus, superior olivary, inferior colliculus, and medial geniculate nucleus, show a consistent organizational pattern in which neurons are aligned based on the frequencies to which they are most sensitive. That is, neurons most responsive to low-frequency energy lie on one edge of each structure, neurons responding to high frequencies lie on the other edge, and neurons responding to other frequencies are spread out in an orderly fashion in between. The pervasiveness of this tonotopic organization pattern reflects both the early mechanical properties of transduction and the importance of the frequency-constitutive sound for auditory perception.

**Primary auditory cortex (A1)**

The first area within the temporal lobes of the brain responsible for processing acoustic information.

**belt area**

A region of cortex, directly adjacent to the primary auditory cortex (A1), with inputs from A1, where neurons respond to more complex characteristics of sounds, as well as to input from other senses.

**parabelt area**

A region of cortex, lateral and adjacent to the belt area, where neurons respond to more complex characteristics of sounds, as well as to input from other senses.

**psychoacoustics**

The branch of psychology that deals with the physiological and psychological correlates of the physical dimensions of acoustics in order to understand how the auditory system operates.

![Figure 9.21](image)

**Figure 9.21** The first stages of auditory processing begin in the temporal lobe in areas within the Sylvian fissure. The top picture is from the side of the brain, and the lower two pictures are looking down at the brain with the parietal cortex cut away. Primary auditory cortex (A1) is in the center. It is surrounded by belt regions, and parabelt regions extend past the belt to the front and side. (From Brugge and Howard, 2002.)

Scientists who study psychoacoustics (psychoacousticians) are always careful to distinguish between the physical characteristics of sounds and the impressions of these sounds for listeners. As we noted earlier, whereas frequency, measured in hertz, is a physical description of the spectral composition of a sound, the subjective attribute of frequency for listeners is pitch. Sounds are measured with respect to frequency, but listeners hear pitch. Similarly, the intensity of sound is measured as sound pressure in decibels, but listeners hear loudness. If the auditory system operated exactly the same as electronic measuring devices, we could use the terms frequency and pitch and the terms intensity and loudness interchangeably. As we will see, however, biological auditory systems do not work exactly as electronic measuring devices work. For example, one sound wave may be heard as quite a bit louder than another, even though the two waves have exactly the same amplitude. Careful study of the differences between the responses of electronic devices (sound-level meters and spectrum analyzers) and biological listening devices (human beings) provides great insight into how the human auditory system works.

**Intensity and Loudness**

The bottom curve in Figure 9.22 shows the human audibility threshold, which graphs the lowest sound pressure level that can be reliably detected, across the frequency range of human hearing (20–20,000 Hz). That is the most absolute thresholds for human hearing are between 2000 and 6000 Hz. (2–6 kilohertz [kHz]). Remember that these frequencies are enhanced by the physical properties of the ear canal. Thresholds rise on both sides of this range, meaning that higher and lower-frequency sound waves must have larger amplitudes in order to be heard.

The other lines in Figure 9.22 are equal-loudness curves (Suzuki and Takeishi, 2004). We obtain these curves by asking listeners to equate the loudness of sounds with different frequencies. We also find greater evidence that this is true, showing that the equal-loudness curve is actually 1000 Hz, so the curve marked 40 shows the amplitude necessary to make tones at other frequencies sound exactly as loud as a 1000-Hz, 40-dB tone; the curve marked 60 represents the decibel level of the processing needed to match a 1000-Hz, 60-dB tone; and so on. As the figure shows, the same pattern of frequency-dependent sensitivity that we see in the audibility threshold curve extends to sounds above threshold. (See Web Activity 9.3: Equal-Loudness Curves.)
The two orange tick marks in Figure 9.22 indicate that a 200-Hz tone presented at 70 dB sounds about as loud as a 9000-Hz tone presented at 60 dB (that is, both points fall on the equal-loudness curve marked 60), whereas a 2000-Hz tone presented at 40 dB sounds much louder than a 100-Hz tone presented at the same level (purple tick marks). These observations demonstrate the inequality of sound pressure level and loudness: equal-amplitude sounds can be perceived as softer or louder than each other, depending on the frequencies of the sound waves.

Another way to track the relationship between amplitude and loudness is to pick one frequency, steadily raise the intensity of the sound, and ask a listener to judge how the loudness increases. Experiments of this type show, once again, that the relationship between intensity and loudness is far from perfect. Doubling the perceived loudness of a sound requires more than a doubling in the amount of acoustic energy present in a sound wave, especially above 40 dB. The same kind of relationship holds in vision: the number of photons must be more than doubled to double the perceived brightness of a light.

The loudness of a sound also depends on its duration. Within limits, longer sounds are heard as being louder. Again, the same thing happens in vision: flashes of light appear brighter when they last longer. The reason for this general phenomenon is that the perception of loudness or brightness depends on the summation of energy over a brief, but noticeable, period of time—a process called temporal integration. For hearing, temporal integration occurs over an interval of 100–200 ms. So if a sound is presented for less than 100 ms, it will be perceived as softer than a sound with the same amplitude and frequency presented for 300 ms. However, there will be little difference in loudness perception if the duration of the sound is increased from 300 to 1000 ms or longer.

In addition to studying absolute loudness judgments, psychoacousticians are interested in how proficient humans are at discriminating between the loudness levels of two sounds. There are several different ways to measure the smallest differences in intensity that can be detected, and many measures show sensitivity to changes of less than 1 dB. This ability is quite impressive, given the wide range of sound intensities (from 0 to over 100 dB) that humans can perceive and the fact that, unlike the visual system, the auditory system is always sensitive to this entire range (remember that, to achieve maximum visual sensitivity, we need time to adapt to lower or higher ambient light levels). Although the ability to discriminate between subtle loudness differences might not seem all that important to survival, we will see in Chapter 10 how the auditory system uses differences between the intensity levels of sounds reaching the left and right ears to determine where sound sources are located.

For a time, it was difficult to understand how listeners could be sensitive to such small differences in loudness over such a large range. Sound wave intensity is generally signaled by the firing rate of auditory nerve fibers: larger intensities (loud sounds) correspond to higher firing rates, and smaller intensities (quiet sounds) correspond to lower firing rates. You should recall that the intensities required to fire fibers (thresholds) vary from one AN fiber to the next. For example, one fiber might selectively respond to the range of amplitudes between 0 and 25 dB, another might span the range of 15–40 dB, a third might cover 38–65 dB, and so on (see Figure 9.17). A full population of neurons with different thresholds can therefore encode a much broader range of intensities than is possible with any single neuron. In addition, remember that neurons become responsive to a broader range of frequencies when intensity is higher (see Figure 9.16). One result is that, as sounds become more intense, many more AN fibers become excited.
Frequency and Pitch

The tonotopic organization of the auditory system, from basilar membrane to primary auditory cortex, is a very big hint that frequency composition is a fundamental determinant of how we hear sounds. More than anything else, psychoacousticians have studied how listeners perceive pitch, the psychological counterpart to frequency. As is the case with intensity and loudness, the frequency of a sound is related to, but not perfectly correlated with, the perceived pitch of the sound. For any given frequency increase (e.g., 50 Hz), listeners will perceive a greater rise in pitch for lower frequencies than they do for higher frequencies. Consequently, listeners perceive a greater pitch difference when a tone shifts from 580 to 1000 Hz than when a tone shifts from 5800 to 5500 Hz.

Research done using pure tones (each composed of a single sine wave) indicates that humans are remarkably good at detecting very small differences in frequency. For example, listeners can discriminate between tones of 1000 and 1001 Hz—a difference of only one-tenth of 1%! Pitch discrimination at the lower and higher ends of the auditory system’s frequency range is not quite as good, but it is still impressive.

Psychoacousticians also use masking experiments to investigate frequency selectivity. In the research described in the previous paragraph, listeners always hear only one sound frequency at a time. In a masking experiment, multiple frequencies are combined, and we see how well listeners can pick out certain components. We look at how effective one sound—the masker—is at hiding another sound.

In the classic approach to measuring frequency selectivity using masking, a single sine wave tone is placed in the middle of a band of acoustic noise (Fletcher, 1940). White noise is a signal that includes equal energy of every frequency in the human auditory range (20–20,000 Hz), just as white light includes light rays of all frequencies in the visible spectrum. A more limited band of noise might include all frequencies in the range of 500–1500 Hz, an even smaller band could span 500–510 Hz.

In a typical experiment, we might start with a 2000-Hz sine wave test tone presented along with a very narrow band of noise—say, 1975–2025 Hz. We would then adjust the intensity of the test tone until listeners could just hear it over the noise. Next we would increase the bandwidth of the noise, perhaps from 50 to 100 Hz, so that now the noise would include frequencies between 1950 and 2050 Hz. As we might expect, the listener would need to increase the intensity of the test tone to be able to hear it over this broader range of noise frequencies.

If we keep widening the bandwidth, however, we will eventually reach a point at which adding more frequencies to the noise stops affecting the detectability of the test tone. The size of the noise band at this point is called the critical bandwidth (Figure 9.24A). For the experimental data plotted in Figure 9.24B, the critical bandwidth is 2000 Hz. In this case, to pick out a 2000-Hz tone from the background noise, listeners must increase the intensity of the tone when the bandwidth is widened from 50 to 100 Hz and from 200 to 400 Hz, but going from a 400-Hz noise band does not require the listener to make the test tone any louder. In fact, the 400-Hz noise band is just as effective a masker as white noise covering the entire spectrum of human hearing. Results from the masking paradigm have helped cement the role of place selectivity. In response to low frequencies, and higher-frequency ranges vibrate smaller portions of the membrane. Correspondingly, masking studies show that the critical bandwidths for low frequencies are smaller than the critical bandwidths for high frequencies because the spacing between low frequencies is larger on the basilar membrane.

Another important finding is that masking effects are asymmetrical. Masking sounds at frequencies lower than that of the test tone are more effective—a phenomenon called the upward spread of masking. This phenomenon may seem counterintuitive, but a look back at Figure 9.12 shows how displacement of the basilar membrane (the traveling wave) extends from the high-frequency base to the low-frequency apex. Displacement for low-frequency energy sounds toward the apex leaves a trail of displacement across high-frequency regions toward the base. From two-tone suppression experiments, we also learned that suppression is greater if the suppressor tone is below an AN fiber’s characteristic frequency than if the suppressor tone is above the fiber’s characteristic frequency (see Figure 9.15).

Hearing Loss

Roughly 30 million Americans suffer some form of hearing impairment. When we talk about hearing loss, we typically do not mean the total loss of all hearing (deafness), but rather the elevation of sound thresholds. For example, frequencies that used to be audible at 20 dB may become inaudible unless they are presented at 40 or 60 dB. Of course, in the end we do not just need to detect sounds; the term hearing really refers to using spectral and temporal differences between sounds in order to learn something about events in the environment. As we will see, common forms of hearing loss can affect the ability to interpret sounds, even when the sounds are loud enough to be detectable.

Hearing can be impaired by damage to any of the structures along the chain of auditory processing from the outer ear all the way up to the auditory cortex. The simplest way to introduce some hearing loss is to obstruct the ear canal, thus inhibiting the ability of sound waves to exert pressure on the tympanic membrane. Many people do this on purpose by wearing earplugs. A less intentional
Hearing impairment occurs most often when the middle ear fills with mucus during ear infections—a condition known as otitis media. The oval window usually still vibrates under these conditions, but without the amplifying power of the ossicles, hearing thresholds can be raised by as much as 50 dB (that is, sounds need to be 50 dB louder in order to be heard). Thankfully, for the millions of young children who suffer ear infections, normal hearing returns after mucus is absorbed back into surrounding tissues; however, this reabsorption can take up to several months. A more serious type of conductive loss, otosclerosis, is caused by abnormal growth of the middle-ear bones, most typically around the oval window next to the stapes. Surgery can free the stapes from these bone growths and improve hearing.

By far the most common, and most serious, form of auditory impairment is sensory-neural hearing loss, which most commonly occurs inside the cochlea and sometimes as a result of damage to the auditory nerve. Disorders such as diabetes as well as bacterial and viral infections can cause hearing loss (Eggermont, 2017). In other patients, hearing loss can be present at birth, or it can appear during adolescence or early adulthood and progressively worsen over one or more decades. Mutations in over 150 different genes have already been linked to hereditary hearing loss in humans.

Sensory-neural loss is caused in two general ways: metabolic and sensory (Vaden Matthews, Eckert, and Dubno, 2017). Metabolic losses are caused by changes in the fluid environment of the cochlea that decrease the activity of hair cells. Sensory losses are caused by injury to hair cells. For example, certain antibiotics and cancer drugs are ototoxic, meaning that they kill hair cells directly. Physicians are well aware of these dangers and typically avoid using such drugs, but sometimes a patient faces the decision of life with deafness versus no life at all.

The major cause of sensory hearing loss is damage to the hair cells by excessive exposure to noise. Those exquisitely fast and sensitive hair cells are also very vulnerable to damage from excessive sound levels. It is widely accepted that damage to outer hair cells, the ones that make inner hair cells more sensitive and more sharply tuned (see Figure 9.14), has significant responsibility for noise-induced hearing loss. Damage to inner hair cells also occurs. With fewer inner hair cells, the neuronal firing pattern described by the volley principle for temporal coding of frequency would become more difficult to maintain, because there would be fewer neurons available to take turns firing (Sayles and Heinz, 2017).

It is fairly well known that shooting a gun without ear protection can cause hearing loss. At least since 1886, when Scottish surgeon Thomas Barr tested the hearing of Glasgow boilermakers, we have known that extended exposure to loud sounds such as the noise of factory equipment also causes hearing loss (Kratochwil, 2014). It is no coincidence that so many aging rock stars and race car drivers wear hearing aids. (It is ironic, but wise, that many heavy-metal music fans now wear ear protection at concerts.) And evidence suggests that even everyday noises present in the environments of industrialized countries can cause hearing loss. In one study (Groycooke et al., 1986), middle-aged and elderly residents of Easter Island who stayed almost exclusively on their quiet island their whole lives were compared with other Easter Islanders, who spent more years off of the relatively quiet island had relatively worse hearing. (After Groycooke et al., 1986.)

Unfortunately, it is not yet possible to state exactly how much exposure to sounds is safe. Readers of this textbook are especially likely to become the listeners from whom we learn what levels are unsafe. This is because of the use of personal listening devices (PLDs) delivering sounds directly to eardrums via earbuds. Right now, there is extensive debate concerning the contribution of PLDs to hearing loss, in part because consequences may be years down the road. At present, we know that use of PLDs and hearing loss are clearly correlated (Ivory, Kane, and Diaz, 2014).

Hearing loss is a natural consequence of aging for many people, and it is difficult to separate a person's age from the amount of exposure to noise. Typically, as people grow older, the ability to hear lower-frequency sounds is lost faster than the ability to hear higher-frequency sounds (Figure 9.25B). The 20- to 20,000-Hz frequency range for human hearing really applies only to young people; by the time most of us reach college age, we may have already lost some ability to hear frequencies above 15,000 Hz. The decrease in the ability to hear higher-frequency sounds continues throughout life, with the highest audible frequency becoming lower and lower as we grow older. Fortunately, many of the sounds that people care most about, including speech and music, are composed predominantly of lower frequencies.

**Treating Hearing Loss**

The earliest devices for helping people with hearing loss were simple horns. The small end of the horn would be held at the entry to the ear canal, and the wide end would be used to funnel more acoustic energy toward the listener's ear. Although effective, these horns were obviously somewhat cumbersome. Electronic hearing aids are much more convenient, but they must be designed to do more than simply amplify all sounds, because extremely loud sounds (above 100 dB) are just as annoying (or painful) for impaired listeners as they are for listeners with normal hearing. For a person who cannot hear sounds until their intensity is
at least 70 dB, compared with about 0 dB for healthy hearing, sounds that could normally vary between 0 and 100 dB must be squeezed between 70 and 100 dB (Figure 9.26). Nearly all modern hearing aids use some means to amplify the signal while also compressing intensity differences to keep the highest intensities within a comfortable listening level.

Hearing aids can also be tuned to provide the greatest amplification only for frequencies in the region of greatest loss (for most people, higher frequencies will need to be amplified more). An additional method is to move energy from frequency regions in which hearing is poor (usually high frequencies) into regions where hearing is normal (Alexander, 2016). Because the lower-frequency region is already being used for lower-frequency sounds, this strategy requires squeezing together some lower-frequency sounds to make room for the high-frequency sound that is being moved down.

One advantage of the old horns over electronic hearing aids was that they permitted listeners to direct their hearing toward the sound source they were most interested in. We may think about hearing aids as amplifying the voice to which one is listening, but they also amplify all the other sounds in the environment. The background noise in a car, or even the sound of a refrigerator, can become loud enough to compete with the sound of a person’s voice. When the entire range of hearing is compressed from a range of 100 dB to only 30 dB, a 10–dB difference between the rumbling of the car’s engine and the voice of the person in the passenger seat becomes compressed into only a 3–dB difference.

Hearing aids are gradually improving, and they have provided relief to millions of Americans, including former presidents Ronald Reagan and Bill Clinton. However, despite researchers’ many clever innovations for improving the signal that arrives at the tympanic membrane, damage to the mechanisms that transduce sound waves into neural signals is proving difficult or impossible to overcome completely. By analogy to vision, the best eyeglasses, contact lenses, or even laser surgery cannot change an image enough to overcome retinal degeneration. The best advice is to never need a hearing aid. Protect your ears by avoiding exposure to loud sounds and by using hearing protection such as earplugs or earmuffs when necessary. If someone else can sing along to the song you’re listening to on your personal media player, turn it down!

**Using versus Detecting Sound**

Even the best hearing aids serve mostly to amplify sounds for frequencies where thresholds are elevated. However, the ability to detect sounds is not the same as the ability to listen to and use sounds. The audiogram, our measure of the softest detectable tones at different frequencies, is not a perfect predictor of listeners’ abilities to use sounds. Understanding speech or enjoying music, especially when there is background noise, can be good with poor audiograms and weak with normal audiograms.

A great deal of recent research activity concerns what is called hidden hearing loss. Hickos and Liberman (2014) showed how exposure to moderately high levels of noise changes the ability of mice to use sound even when their ability to detect sounds remains normal. This hidden hearing loss has been shown to occur following only one week of exposure to 84-dB noise (Maisen, Ushbach, and Liberman, 2013). Even brief periods of exposure to loud noise result in a loss of synapses between AN fibers and hair cells (Figure 9.27). Unlike damage to hair cells, which results in decreased sensitivity, loss of synapses results in a loss of connectivity. It has been hypothesized that hidden hearing loss might explain some of the difficulties that human listeners have in noisy situations even when their audiograms are normal (Bharadwaj et al., 2014). Unfortunately, improving the ability of listeners to detect sounds is unlikely to improve the ability of people with hidden hearing loss to use sounds.

![Image](https://example.com/image.png)

*Figure 9.26* When hearing thresholds are increased by impairment, a sound must have more energy to be heard, but loudness increases faster than it does with healthy ears. When a hearing aid is used to increase the intensity of sounds, all variation in sound levels must be compressed into a smaller range of intensity because very loud sounds can be just as painful for listeners with hearing impairment as they are for listeners with healthy hearing. (After C. D. Geisler, 1998.)

**Figure 9.27** Following exposure to loud sounds, listeners can have difficulty using sounds even when hair cells appear to be undamaged. This is because some of the synapses between hair cells and neurons in the auditory nerve are lost. The top (purple) photos are of healthy hair cells before and after exposure to noise. The bottom photos show synapses highlighted as small yellow dots. There are fewer synapses following noise exposure (right) as compared to before noise exposure. (Courtesy of Vijaya Prakash Krishnan Muthaiah.)
CHAPTER 9 HEARING: PHYSIOLOGY AND PSYCHOACOUSTICS

Summary

1. Sounds are fluctuations of pressure. Sound waves are defined by the frequency, intensity (amplitude), and phase of fluctuations. Sound frequency and intensity correspond to our perceptions of pitch and loudness, respectively.

2. Sound is funneled into the ear by the outer ear, made more intense by the middle ear, and transformed into neural signals by the inner ear.

3. In the inner ear, cilia on the tops of inner hair cells pivot in response to pressure fluctuations in ways that provide information about frequency and intensity to the auditory nerve and the brain. Auditory nerve fibers convey information through both the rate and the timing patterns with which they fire.

4. Different characteristics of sounds are processed at multiple places in the brain stem before information reaches the cortex. Information from both ears is brought together very early in the chain of processing. At each stage of auditory processing, including primary auditory cortex, neurons are organized in relation to the frequencies of sounds (tonotopically).

5. Humans and other mammals can hear sounds across an enormous range of intensities. Not all sound frequencies are heard as being equally loud, however. Hearing across such a wide range of intensities is accomplished by the use of many auditory neurons. Different neurons respond to different levels of intensity. In addition, more neurons overall respond when sounds are more intense.

6. Series of channels (or filters) process sounds within bands of frequency. Depending on frequency, these channels vary in how wide (many frequencies) or narrow they are. Consequently, it is easier to detect differences between some frequencies than between others. When energy from multiple frequencies is present, lower-frequency energy makes it relatively more difficult to hear higher frequencies.

7. Hearing loss is caused by damage to the bones of the middle ear, to the hair cells in the cochlea, or to the neurons in the auditory nerve. Although hearing aids are helpful to listeners with hearing impairment, there is only so much that can be done to help after damage to hair cells that cannot be repaired.

Sensation & Perception in Everyday Life

Electronic Ears

Modern medical science and engineering are providing some degree of hearing to many people who are deaf. Cochlear prosthetics, more commonly known as cochlear implants (Figure 9.28A), are tiny flexible coils with about two dozen miniature electrode contacts along their length. Surgeons delicately thread these electrode arrays through the round window as far toward the apex of the cochlea as possible. The electrode array is connected to a tiny radio receiver under the scalp, and signals are transmitted from a small microphone device on the outside of the head behind the ear (Figure 9.28B). Signals coming in from the microphone activate the miniature electrodes at appropriate positions along the cochlear implant, which in turn stimulates associated auditory nerve fibers.

Although cochlear implants are a modern medical miracle, they cannot provide hearing that approaches what nature provides. We are not surprised, however, that a handful of electrodes cannot replace the function of 14,000 hair cells. Some people benefit more than others from implants, and many adults with electrical hearing converse flawlessly over the phone. Young children, receiving implants as young as age 1 or 2, do best of all because young brains are particularly plastic. Children’s brain circuitry develops to get the most information possible from their electronic ears.

Figure 9.28 Cochlear implants give some people who are deaf the ability to hear. (A) The cochlear implant electrode array. The flexible array of electrodes is inserted through the round window as far as possible toward the apex of the cochlea. (B) The electrode array is connected to a small receiver beneath the scalp, and a small microphone and transmitter are placed over the receiver on the outside of the head.