9 Hearing: Physiology and Psychoacoustics

IT IS OFTEN SAID THAT HUMANS ARE VISUAL ANIMALS. We are reminded of the importance of vision in our lives 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. Your ears are always open, and you can hear perfectly well in the dark. When you enter a theater just before the movie starts, you don't need to wait half an hour to be able to hear the soft strains of the opening theme. You can hear things when your ears are not pointed at the source of the sound, and you can even hear around obstacles and corners. For better or worse, you can often hear through walls 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 cannot hear that train coming, and the blast of its air horn simply escapes into the night. At a concert, musicians move deftly as the audience stands or jumps or claps, but the enthusiasm of both performers and spectators is only a curiosity. A person cries out, and you don't help because you cannot hear the cry. You cannot talk on the phone. Deafness deprives you of the most fundamental of human abilities: communication through speech. When you imagine life without hearing, you imagine a more dangerous world in which you are relatively isolated by day, and nearly totally isolated by night. Your sense of vision, which may compensate for some of the hearing loss when it's light out, is now compromised by the fact that it's dark.

The Function of Hearing

For the next three chapters, you will be hearing all about hearing. Here 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 ameliorate these impairments. In Chapter 10 we will move on to discuss some of the ways we use auditory information to learn about our environment. Then in Chapter 11 we will cover the higher-level auditory functions of speech and music comprehension.

Many fundamental principles apply to vision, hearing, and all the other 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 have learned thus far will prove helpful for understanding hearing, you should also come to appreciate how biology has provided some very different (and very clever) solutions to the idiosyncratic problems involved in sensing and interpreting sound.

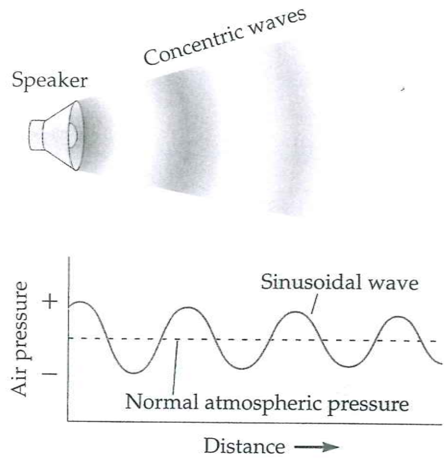


FIGURE 9.1 The pattern of pressure fluctuations of a sound stay the same as the sound wave moves away from the source, but the amount of pressure change decreases with increasing distance.

What Is Sound?

Sounds are created when objects vibrate. The vibrations of the object (the sound source) cause molecules in the object's surrounding medium (for humans, usually the Earth's atmosphere) to vibrate as well, and this in turn causes pressure changes in the medium (see Figure 9.1). These pressure changes are best described as waves, and they are similar to the waves on a pond that are caused by dropping a rock into the water. The water molecules that are 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 its 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 gets farther from its source.

Sound waves travel at a particular speed depending on the medium, traveling faster through denser substances. For example, the speed of sound through air is about 340 meters per second, depending on the humidity level (sounds travel a bit faster on muggy days), but the speed of sound through water is about 1500 meters per second. Light waves, whose speed is unaffected by the medium through which they travel, move through air almost a million times faster than sound waves, accounting for the lag between seeing lightning and hearing thunder (unless you are much too close to the lightning).

In the 1950s, there was a great race to fly an airplane faster than the speed of sound (over 760 miles per hour), and more recently a jet car has exceeded the speed of sound on land. When an object such as a jet plane travels faster than the speed of sound, the plane catches up to and passes the fronts of the sound waves it is creating. As a result, the sound waves combine into a shock wave, or a huge pressure fluctuation. When this shock wave reaches the ground and we hear it, it is called a "sonic boom."

Basic Qualities of Sound Waves: Frequency and Amplitude

As we've seen, the 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 area and the lowest pressure area—is called the **amplitude** or **intensity** of a wave (Figure 9.2). Pressure fluctuations can be very close together, or they can be spread apart over longer periods. For light waves, we usually describe the pattern of fluctuations by measuring the distance between peaks in the waves—the wavelength. Although sound waves also have wavelengths, we more typically describe their patterns by noting how quickly the pressure fluctuates; this rate of fluctuation is known as the **frequency** of the wave (see Figure 9.2). Sound wave frequencies are measured in **hertz (Hz)**, where 1 Hz equals one cycle per second. For example, the air 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.

Just as the amplitude and wavelength of light waves correspond to perceptual qualities (brightness and color, respectively) in vision, the amplitude and frequency of sound waves are highly correlated with auditory characteristics. Amplitude is associated with the perceptual quality of **loudness**: the more intense a sound wave, the louder it will sound. Frequency is associated with **pitch**: low-frequency sounds correspond to low pitches (e.g., low notes played by a tuba) and high-frequency sounds corresponding to high pitches

amplitude Magnitude of displacement (increase or decrease) of a sound pressure wave.

intensity Amount of sound energy falling on a unit area (such as cm^2).

frequency For sound, the number of times per second that a pattern of pressure change repeats.

hertz (Hz) A unit of measure for frequency, where 1 Hz equals one cycle per second.

loudness The psychological aspect of sound related to perceived intensity or magnitude.

pitch The psychological aspect of sound related mainly to the fundamental frequency.

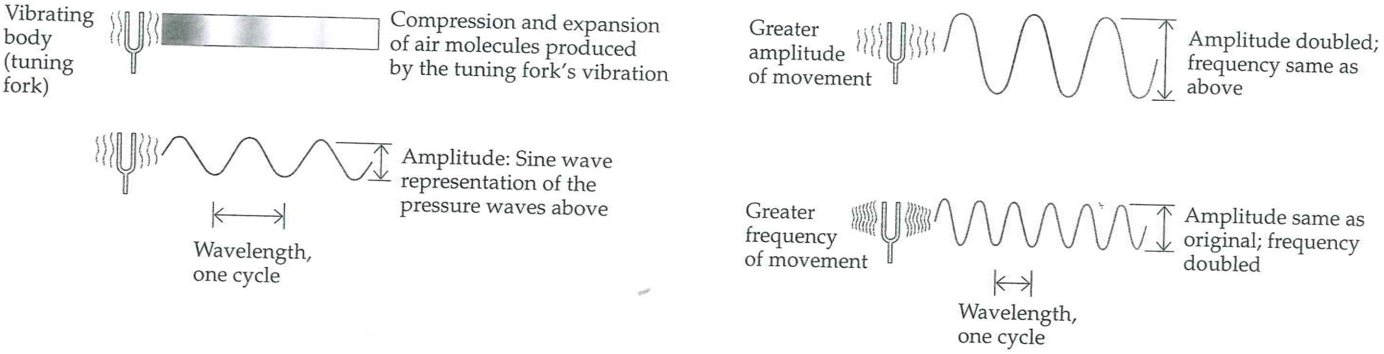


FIGURE 9.2 Sound waves are described by the frequency and amplitude of pressure fluctuations. Here, changes in frequency and amplitude are shown for sine waves, the simplest kind of sound wave.

(e.g., the high notes from a piccolo). We will have much more to say about the relationships between amplitude and loudness and between frequency and pitch later in this chapter.

In Chapter 1 you learned that 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 environmental sounds. If you are relatively young and you have 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. Elephants appear to 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 utilize sound frequencies above 60,000 Hz.

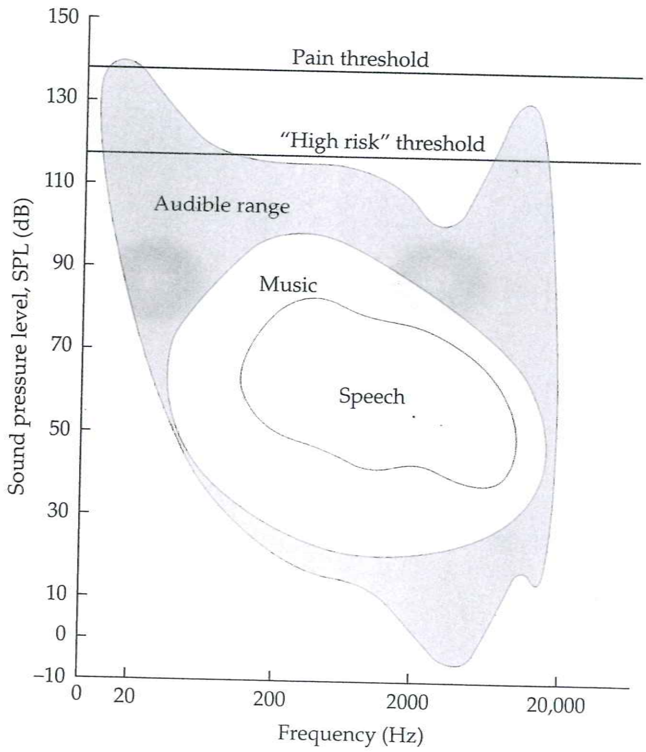


FIGURE 9.3 Humans can hear frequencies that range from about 20 to 20,000 Hz across a very wide range of intensities.

decibel (dB) A unit of measure for the physical intensity of sound. Decibels define the difference between two sounds as the ratio between two sound pressures. Each 10:1 sound pressure ratio equals 20 dB, and a 100:1 ratio equals 40 dB.

Humans hear across a very wide range of sound intensities: the ratio between the faintest sound humans can detect and the loudest sounds that do not cause serious damage to ears is more than one to one million. In order to describe differences in amplitude 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 10:1 sound pressure ratio is equal to 20 decibels (dB), so a 100:1 ratio is equal to 40 decibels. The equation for defining decibels is

$$\text{dB} = 20 \log(p/p_0)$$

The variable p corresponds to the pressure (intensity) of the sound being described. The constant term p_0 is a reference pressure, and is typically defined in auditory research contexts to be 0.0002 dyne/cm^2, and levels are defined as dB SPL (sound pressure level). This level (0.0002 dyne/cm^2) 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/cm^2.

Using a reference value such as p_0 is common to many measuring systems. For example, 0°C is defined as the temperature at which water freezes, and 100°C is the temperature at which water boils. If the pressure of the sound you are measuring (p) is equal to 0.0002 dyne/cm^2, then dB = 20 log(1). Since the log of 1 is zero, a sound pressure that low would be equal to 0 dB SPL. Sounds with amplitudes even smaller than p_0 will have negative decibel levels, just as substances colder than the freezing point of water have negative centigrade temperatures.

An important thing to remember about logarithmic scales such as decibels is that relatively small decibel changes can correspond to large physical changes. For example, an increase of 6 dB corresponds to a doubling of the amount of pressure. Figure 9.4 shows the decibel levels of some common sound sources.

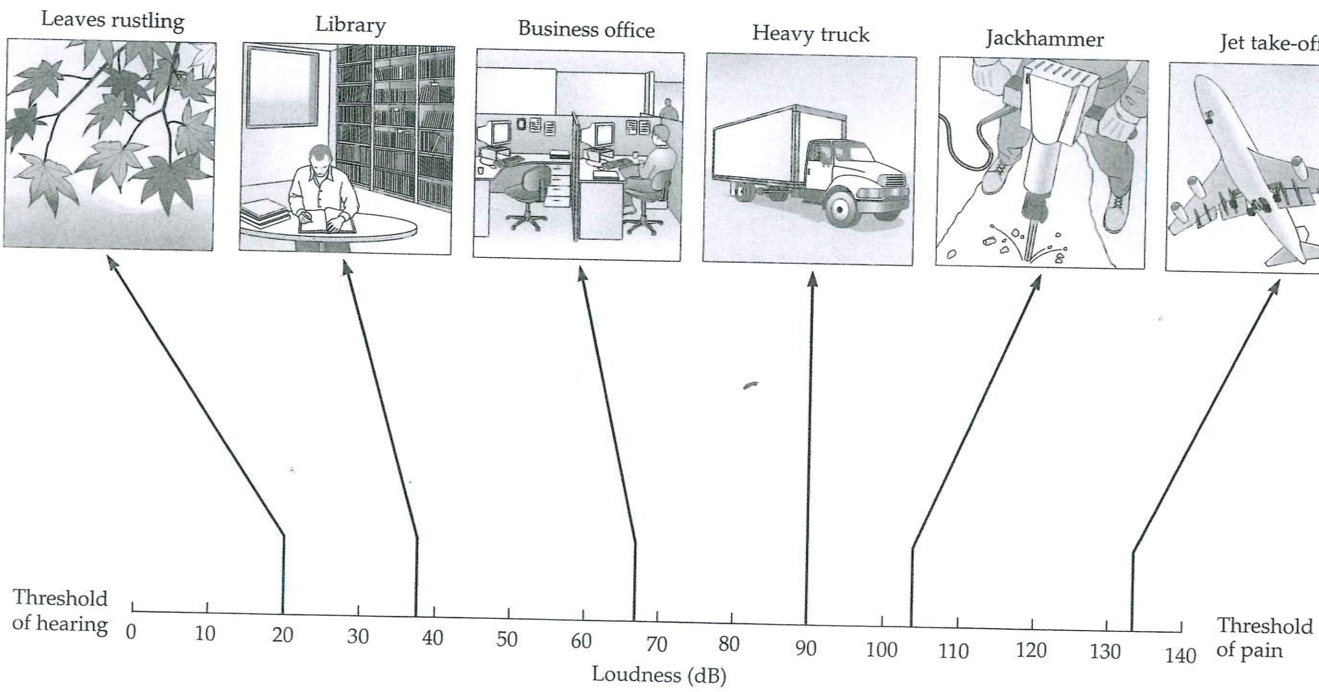


FIGURE 9.4 Sounds that we hear in our daily environments vary greatly in their intensities.

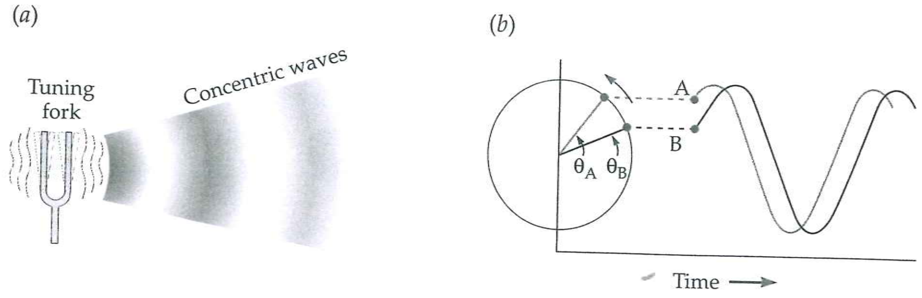


FIGURE 9.5 A sine wave is a circular motion extended over time. The two sine waves here have the same frequency, but different phases. At any point in time, A is at one phase angle in the cycle θ_A, and B is at another phase angle θ_B.

Sine Waves, Complex Tones, and Fourier Analysis

One of the simplest kinds of sounds is a **sine wave**, or **pure tone**. (See Web Activity 9.1 What We Hear.) The air pressure in a sine wave changes continuously (sinusoidally) at the same frequency (Figure 9.5). The time taken for one complete cycle of a sine wave is the **period** of the sine wave, and there are 360° of **phase** across one period. Thus, the undulation of the sine wave over time is described in degrees in the same way that rotations around a circle would be described.

Sine waves are not common everyday sounds, because few vibrations in the world are so pure. If you have taken a hearing test or used tuning forks, you may have heard sine waves. Flutes can produce musical notes that are close to pure tones, but other musical instruments, human voices, birds, cars, and almost all other sound sources in the world produce **complex tones**.

If pure tones are so uncommon, you may be wondering why we bother discussing them. It turns out that all sounds, no matter how complex, can be described as some combination of sine waves (Figure 9.6). Even the cacophony of a room full of people talking or the swelling sound of a full orchestra can be broken down into combinations of sine waves at many different frequencies with different amplitudes. The individual sine wave components of

sine wave Waveform for which variation as a function of time is a sine function. Also called *pure tone.*

period Time (or space) required for one cycle of a repeating waveform.

phase The relative position of two or more sine waves. For sounds, *phase* refers to relative position in time.

complex tone A sound wave consisting of more than one sinusoidal component of different frequencies.

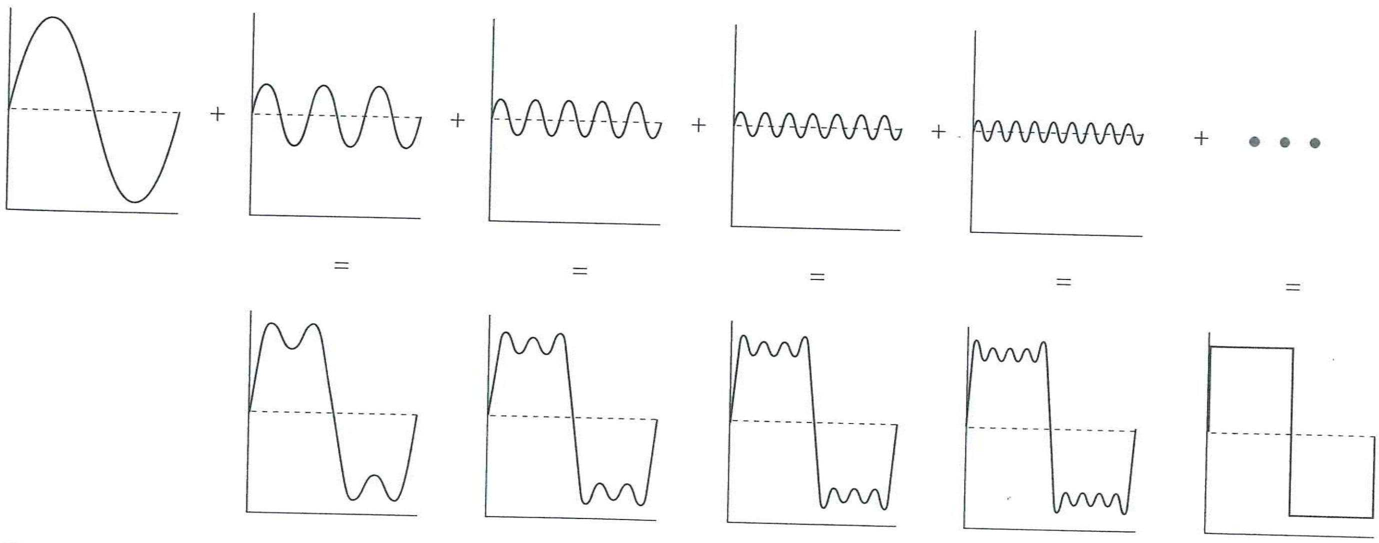


FIGURE 9.6 Every complex sound wave can be analyzed as a combination of sine waves, each with its own frequency, amplitude, and phase. Here, multiple sine waves are added together to form more complex waveforms. When infinitely more sine waves with even higher frequencies are added, a square wave (bottom right) can be constructed.

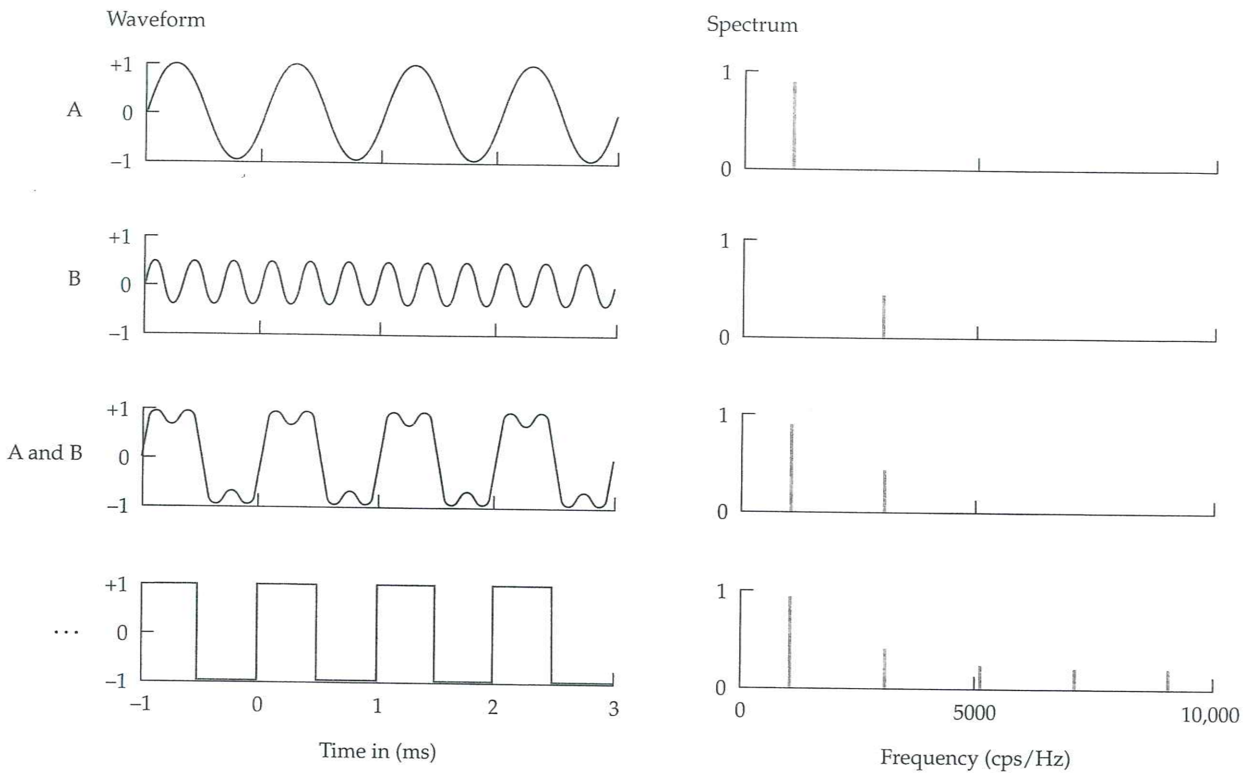


FIGURE 9.7 A spectrum displays the amplitude for each frequency present in a sound wave. Each signal is shown as a wave form (left) and as a spectrum (right).

a complex sound can be described by a process called **Fourier analysis**, and we can summarize the results of a Fourier analysis with a graph, called a **spectrum**, that shows the intensity of each sine wave frequency found in the complex tone (Figure 9.7). (See **Web Activity 9.2 Fourier Analysis**.)

Sounds with **harmonic spectra**, illustrated in Figure 9.8, are typically caused by a simple vibrating source, such as the string of a guitar or the reed of a saxophone. Each frequency component in such a sound is called a harmonic." The first harmonic, called 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 Fourier spectrum is one of the most important qualities that distinguish different sounds. Spectral shapes are related to properties of sound sources and help us identify sources. For example, Figure 9.8 illustrates spectra from three musical instruments. Each instrument is producing a tone with the same fundamental frequency (262 Hz; the note C_4 [middle C]) and the same harmonics (524 Hz, 786 Hz, 1048 Hz, and so on). However, the shapes of the spectra (the pattern of amplitudes for each harmonic) vary. **Timbre** (pronounced "tamber," like "amber") is a term used to describe the quality of a sound that depends, in part, upon the relative energy levels of harmonic components.

We will return to harmonics, timbre, and other aspects of complex sounds in Chapter 10. For this chapter, we will stick mainly to the story of how the auditory system perceives simple sounds such as sine wave tones.

Fourier analysis A mathematical theorem by which any sound can be divided into a set of sine waves. Combining these sine waves will reproduce the original sound.

spectrum (pl. spectra) A representation of the relative energy (intensity) present at each frequency.

harmonic spectrum The spectrum of a complex sound in which energy is at integer multiples of the fundamental frequency.

fundamental frequency The lowest-frequency component of a complex periodic sound.

timbre The psychological sensation by which a listener can judge that two sounds that have the same loudness and pitch are dissimilar. Timbre quality is conveyed by harmonics and other high frequencies.

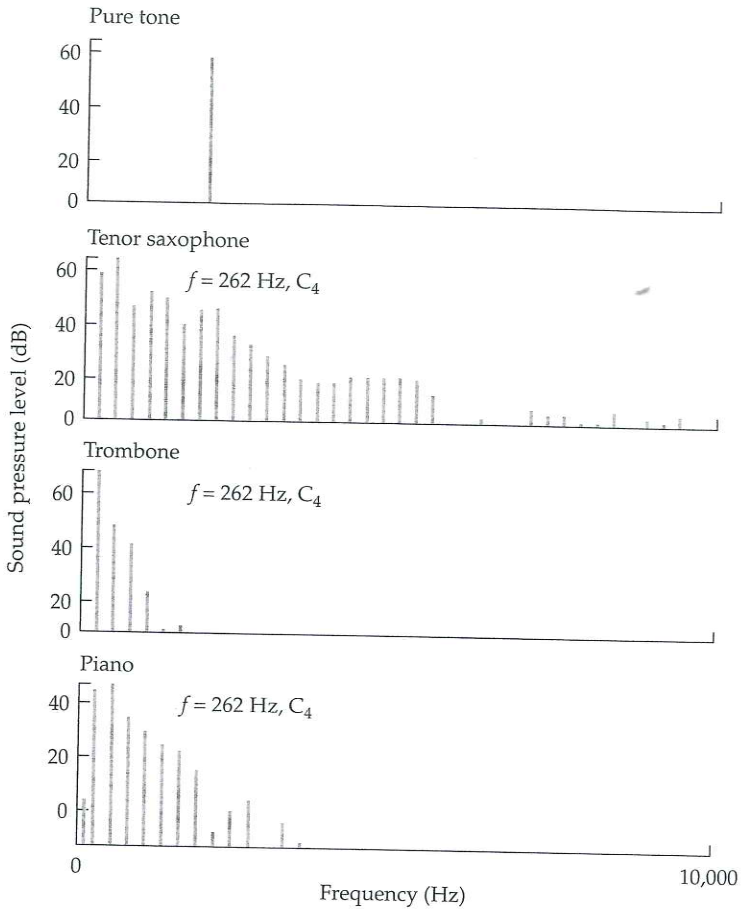


FIGURE 9.8 Harmonic sounds with the same fundamental frequency can sound different because amplitudes of individual frequency components are different, resulting in different spectral shapes. For example, different musical instruments playing the same note (the same fundamental frequency, abbreviated f) sound different.

Basic Structure of the Mammalian 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 pretty amazing things. You are about to learn about quite a few anatomical structures that are essential for understanding how sequences of tiny air pressure changes are turned into meaningful sound perception. Here and there you might become a bit confused. Consult the figures and Web Activity 9.3 Structure of the Auditory System often, and you will soon know the parts and how they fit together.

Outer Ear

Sounds are first collected from the environment by the **pinna** (plural *pinnae*), the curly structure on the side of the head that we typically call an ear. Only mammals have pinnae, which vary wildly in shape and size across species and vary somewhat less dramatically across individuals within species (Figure 9.9). As we will see in Chapter 10, the particular shapes of the pinnae play 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 mm into the head (Figure 9.10). The length and shape

pinna (pl. pinnae) 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.

FIGURE 9.9 Pinna size and shape vary greatly among mammals.

of the ear canal enhance sound frequencies between about 2000 and 6000 Hz, but the main purpose of the canal is to insulate 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. In fact, in most cases a damaged tympanic membrane will heal itself, just as other parts of your skin do. However, it is possible to damage the tympanic membrane beyond repair, so it's a good idea to follow your mother's advice and refrain from sticking things inside your ear.

Middle Ear

Together, the pinna and ear canal make up a division of the auditory system called the **outer ear** (see Figure 9.10). 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 **stapes**, which transmits the vibrations of sound waves to the **oval window**, another membrane which represents the border between the middle ear and the **inner ear**.

The ossicles, which are the smallest bones in the human body, amplify sound vibrations in two ways. First, the joints between the bones are hinged in a way that makes 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 33%.

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 9.10). Therefore,

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-gathering 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.

ossicles Three tiny bones of the middle ear: malleus, incus, and stapes.

malleus The first ossicle; the malleus receives vibration from the tympanic membrane and is attached to the incus.

incus The middle ossicle; the connection between malleus and stapes.

stapes The third ossicle; 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 vestibular canals.

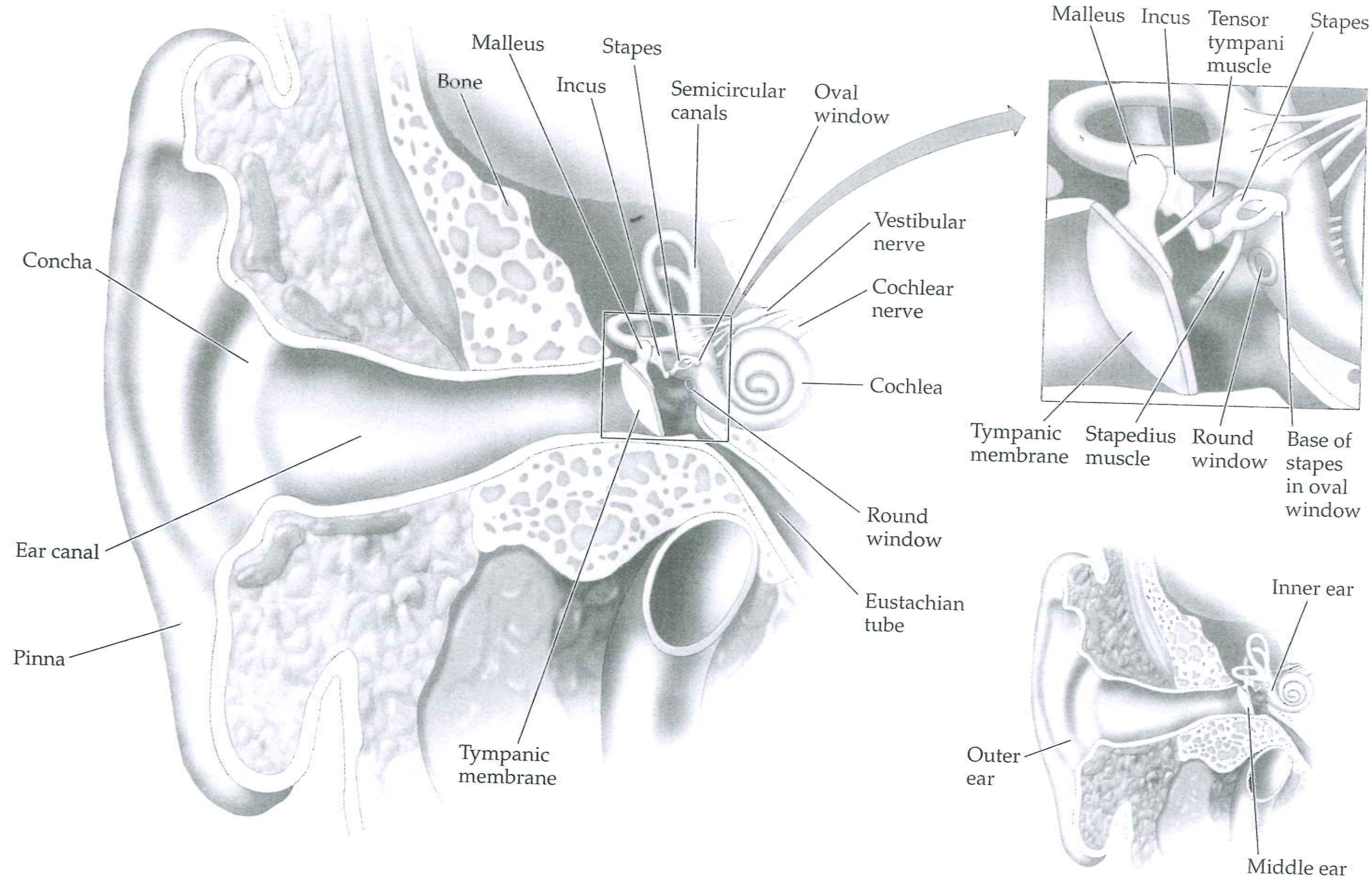


FIGURE 9.10 Structures of the human ear. Note that the tympanic membrane has about 18 times as much surface area as the oval window beneath the stapes.

the pressure on the oval window is magnified 18 times relative to the pressure on the tympanic membrane. This is the same principle that makes snowshoes effective for keeping feet from plunging through the snow and makes stiletto 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).

Amplification provided by the ossicles 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 an impedence mismatch: if sound waves were transmitted to the oval window directly, many would simply bounce back without moving the oval window at all.

The 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 9.10). As might be expected, given that the ossicles are the smallest bones in the body, these are the smallest muscles in the body. Their main purpose is to tense when sounds are very loud, restricting the movement of the ossicles and thus muffling pressure changes that might be so great as to damage the delicate structures in the inner ear. Unfortunately, this **acoustic reflex** follows the onset of loud sounds by about one-fifth of a second. So although it helps in environments that are loud for sustained periods, the acoustic reflex cannot protect against abrupt loud sounds, such as the firing of a gun. The muscles of the middle ear can

tensor tympani The muscle attached to the malleus; tensing the tensor tympani decreases vibration.

stapedius The muscle attached to the stapes in the cochlea; tensing the stapedius decreases vibration.

acoustic reflex A reflex that protects the ear from intense sounds, via contraction of the stapedius and tensor tympani muscles.

also be tensed 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 fine 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 is thus roughly analogous to that of the retina, where the visual system translates the information carried by light waves into neural signals.

COCHLEAR CANALS AND MEMBRANES The major structure of the inner ear is the **cochlea** (from the Greek *kochlos*, "snail"), a tiny coiled structure embedded in the temporal bone of the skull (see Figure 9.10). Rolled up, the cochlea is the size of a baby pea, about 4 mm in diameter in humans. Uncoiled, it would be a tube about 35 mm in length. The cochlea is filled with watery fluids in three parallel canals (Figure 9.11): 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 balloon (the kind clowns use to make hats and animals), blown up and folded back on itself. The middle canal is another long balloon that is squeezed, lengthwise, between the two halves of the first balloon.

The three canals of the cochlea are separated by two membranes (see Figure 9.11): **Reissner's membrane**, between the vestibular canal and the middle canal, and the **basilar membrane**, between the middle canal and the tympanic canal. 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. Rather, it is a plate made up of fibers that have some stiffness. 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 initiates sound waves in the air. However, because the cochlea is a closed system, the pressure changes cannot spread out in all directions, as they do in the atmosphere. Instead, a displacement, or "bulge," forms in the vestibular canal and travels from the base of the cochlea down to the apex. By the time the traveling wave reaches the apex, its displacement has mostly dissipated (see Figure 9.13). If sounds are extremely intense, any pressure that remains is transmitted through the helicotrema and back to the cochlear base through the tympanic canal, where it is absorbed by yet another membrane, called the **round window**.

Remember that the vestibular and tympanic canals are wrapped tightly around the middle canal. So when the vestibular canal bulges out, it puts pressure on the middle canal. This pressure has the effect of displacing the cochlear partition (which, recall, lies at the bottom of the middle canal), moving the partition down as the vestibular canal bulge comes through, and back up as the bulge passes by.

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

tympanic canal (or scala tympani) A fluid-filled passage that extends from the round window at the base of the cochlea to the helicotrema at the apex.

vestibular canal (or scala vestibuli) A fluid-filled passage that extends from the oval window at the base of the cochlea to the helicotrema at the apex.

middle canal (or scala media) Tympanic and vestibular canals, containing the cochlear partition.

helicotrema The connection between the tympanic and vestibular canals at the apex of the cochlea.

Reissner's membrane A thin sheath of tissue separating the vestibular and middle canals.

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

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 base of the tympanic canal that releases excess pressure remaining from extremely intense sounds.

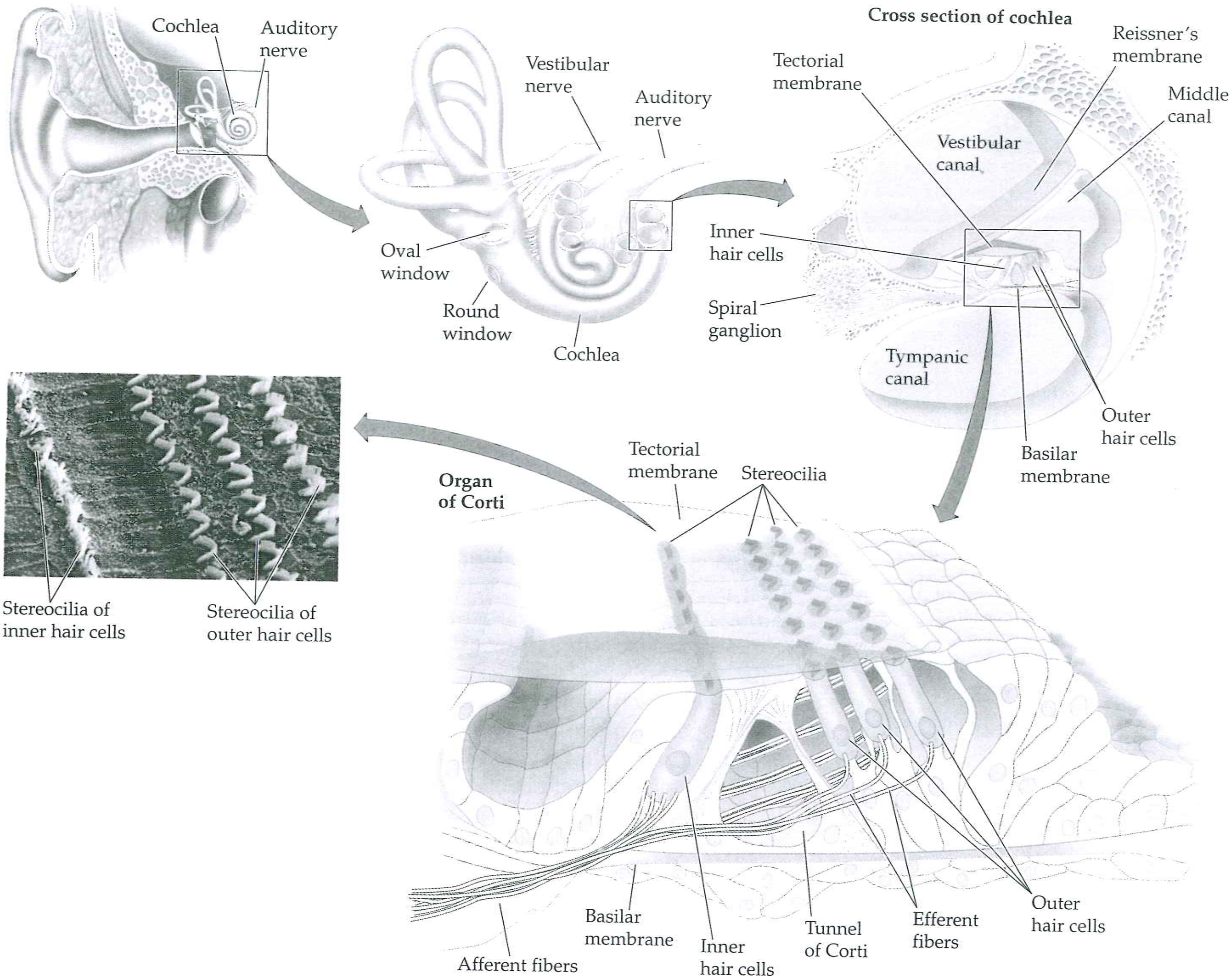


FIGURE 9.11 The cochlea. The illustration at the upper left is from the viewpoint of facing a person. The remaining illustrations show cross sections of the cochlea at successively greater levels of detail. Note the three canals of the cochlea: vestibular, middle, and tympanic (upper right). When vibrations enter the cochlea, the tectorial membrane shears across the organ of Corti. The photomicrograph shows real hair cells: the single row of inner hair cells (left) and three rows of outer hair cells (right). (Micrograph from Kessel and Kardon, 1979.)

THE ORGAN OF CORTI Movements of the cochlear partition are translated into neural signals by structures in the **organ of Corti**, which extends along the top of the basilar membrane (see Figure 9.11). The organ of Corti is made up of specialized neurons called **hair cells**, dendrites of **auditory nerve fibers** that terminate at the base of hair cells, and a scaffold of supporting cells. Hair cells in each human ear are arranged in four rows that run down the length of the basilar membrane: one row of about 3500 inner hair cells and three rows with a total of about 10,500 outer hair cells.

organ of Corti A structure on the basilar membrane composed of hair cells and dendrites of auditory nerve fibers.

hair cells Cells that support the stereocilia that transduce mechanical movement into neural activity sent to the brain stem; hair cells also receive inputs from the brain.

auditory nerve fibers A collection of neurons that convey information from hair cells to (afferent) and from (efferent) the brain stem. This collection also includes neurons for the vestibular system.

stereocilia (s. stereocilium) Hairlike extensions on tips of hair cells that initiate the release of neurotransmitters when they are flexed.

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

Inner and outer hair cells provide the foundation for minuscule hairlike bristles called **stereocilia** (singular *sterocilium*). On an inner hair cell, stereocilia are arranged as if posing for a group photo, in several nearly straight rows with the shorter stereocilia in front and the taller ones peering over their shoulders in the back. On an outer hair cell, stereocilia stand in rows that form the shape of a *V* or *W* (see Figure 9.11).

The **tectorial membrane** extends atop the organ of Corti. It isn't really a membrane either. Rather, it is a gelatinous structure that is attached on one end and floats above the outer hair cells on the other end. Taller stereocilia of outer hair cells are embedded in the tectorial membrane, and the cilia of inner hair cells are nestled against it. Because the tectorial membrane is attached on only one end, it shears across the width of the cochlear partition whenever the partition moves up and down. This shearing motion in turn causes the stereocilia of both inner and outer hair cells to bend back and forth (Figure 9.12).

Remember that, like photoreceptors in the retina, hair cells are specialized neurons. Deflection of a hair cell's stereocilia causes a change in voltage potential that initiates the release of neurotransmitters, which in turn encourages firing by auditory nerve fibers with dendritic synapses on hair cells.

The firing of the auditory nerve fibers finally completes the process of translating sound waves into patterns of neural activity. Here's a brief summary of the whole process: An air pressure wave is funneled by the pinna through the auditory canal to the tympanic membrane, which vibrates back and forth in time with the sound wave. The tympanic membrane moves the malleus, which moves the incus, which moves the stapes, which pushes and pulls on the oval window. The movement of the oval window causes pressure bulges to move down the length of the vestibular canal, and these bulges in the vestibular canal displace the middle canal up and down. This up-and-down motion forces the tectorial membrane to shear across the organ of Corti, moving the stereocilia atop hair cells back and forth. The flexing of the stereocilia starts a chain of biochemical reactions that results in the release of neurotransmitters into synapses between the hair cells and dendrites of auditory nerve fibers. These neurotransmitters initiate action potentials in

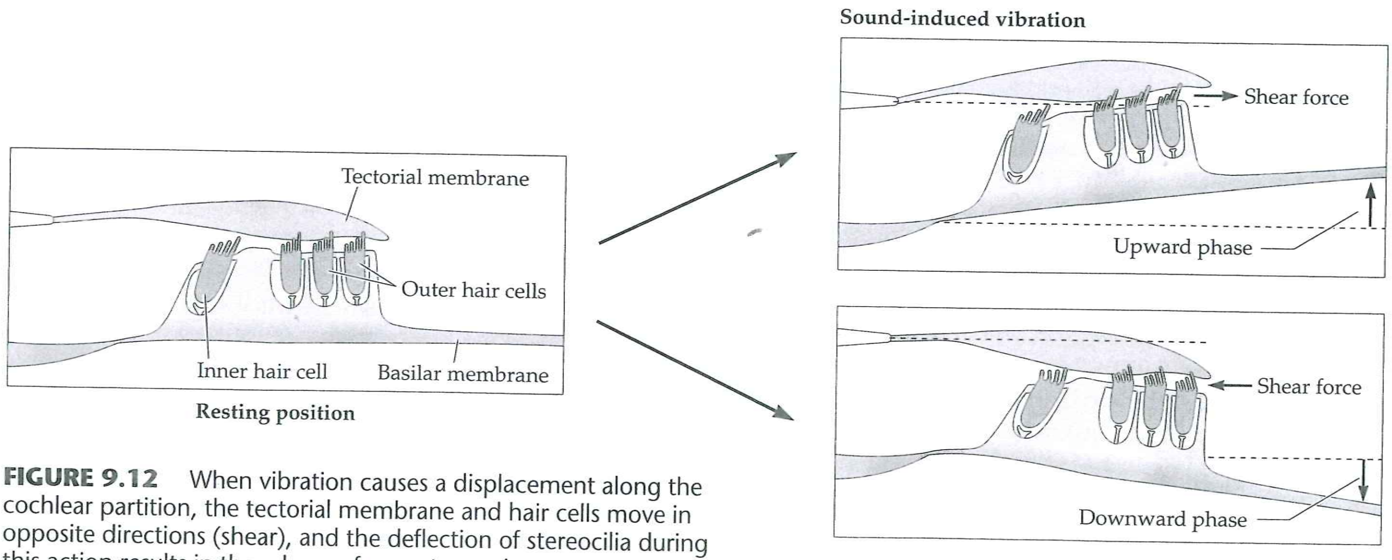


FIGURE 9.12 When vibration causes a displacement along the cochlear partition, the tectorial membrane and hair cells move in opposite directions (shear), and the deflection of stereocilia during this action results in the release of neurotransmitters.

the auditory nerve fibers that are carried back into the brain. And that's all there is to it!

CODING OF AMPLITUDE AND FREQUENCY IN THE COCHLEA Now let's return to the bulging vestibular canal and undulating middle canal, to see how the two fundamental characteristics of sound waves—amplitude and frequency—are encoded by the cochlea.

As the amplitude of a sound wave increases, 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 more forcefully, which causes the hair cells to bend farther back and forth, which causes more neurotransmitters to be released, which causes the auditory nerve fibers to fire action potentials more quickly. Thus, in the end, sound wave amplitude is basically conveyed in much the same way that light wave amplitude is: the larger the amplitude, the higher the firing rate of the neurons that communicate with the brain. (We will discuss some complications of this simple explanation later in the chapter.)

Coding for frequency is a bit trickier. Earlier, we said that the cochlear partition is displaced up and down in a pattern reflecting the pattern (frequency) of 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 displacements closer to the oval window, near the base of the cochlea, and lower frequencies cause displacements nearer to the apex. Thus, different parts of 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 the way the structure of the basilar membrane changes along the length of the cochlea (Figure 9.13). The cochlea as a whole becomes narrower from base to apex, but the basilar membrane actually becomes wider toward the apex. In addition, the basilar membrane begins thicker at the base and becomes thinner as it gets wider. Higher frequencies bend 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.

INNER AND OUTER HAIR CELLS The fact that there are two different types of hair cells, the inner hair cells and the outer hair cells, may remind you once again of the visual system and its two different types of photoreceptors, the rods and the cones. In this case, however, the analogy is not appropriate, because in the auditory system, over 90% of the **afferent fibers** in the auditory nerve—fibers that take information *to* the brain—synapse on the inner hair cells (10 to 30 auditory nerve fibers listen to each inner hair cell).

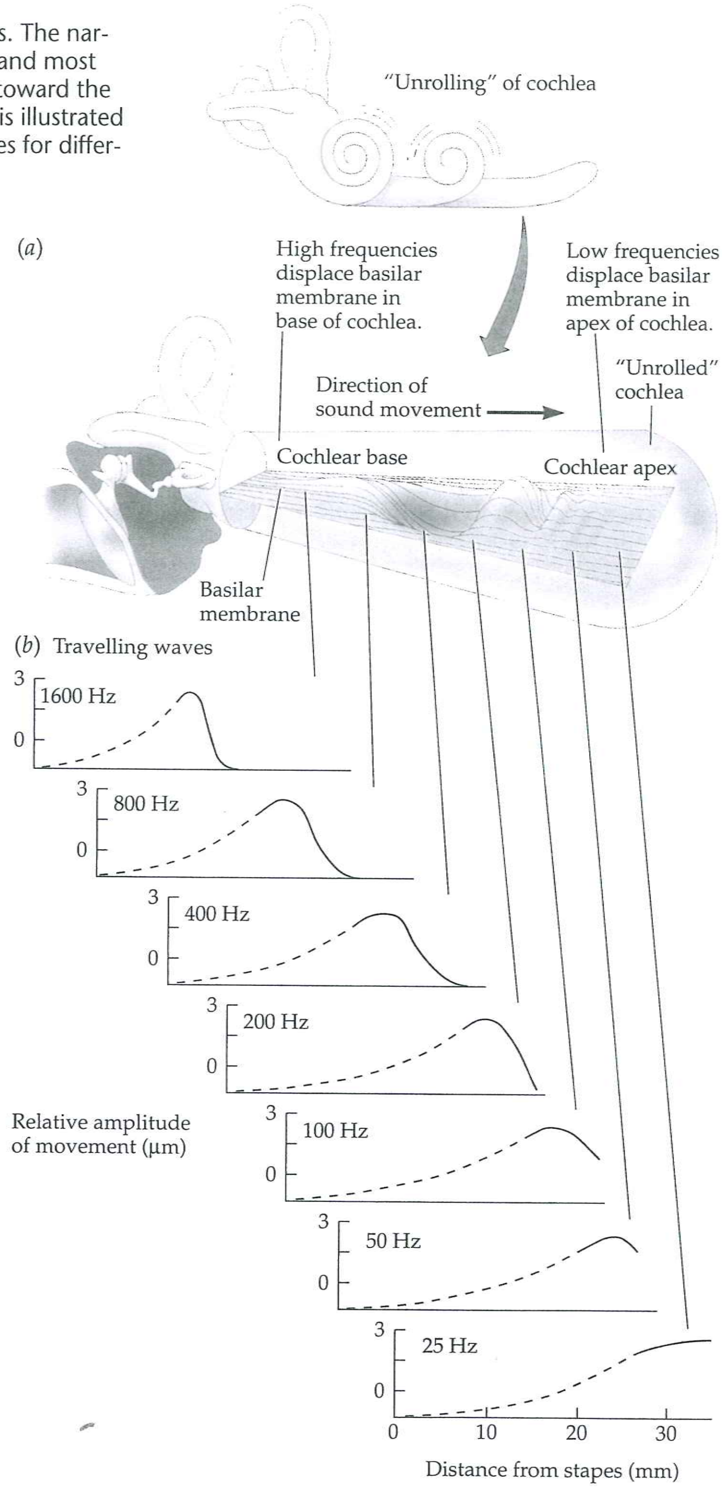
If the inner hair cells are conveying almost all the information about sound waves to the brain, then what are the outer hair cells for? It turns out that most of the nerve fibers that synapse with the outer hair cells are **efferent fibers**, conveying information *from* the brain. When these efferent fibers become active, outer hair cells with which they synapse become physically longer, making the nearby cochlear partition stiffer and less sensitive to pressure changes. Outer hair cells appear to be involved in an elaborate feedback system, and there are several hypotheses about how they alter our hearing. For example, changing the mechanical properties of the basilar membrane could change the frequency tuning of the cochlear partition or change sensitivity when there is background noise.

place code Tuning of different parts of the cochlea to different frequencies, in which information about the particular frequency of an incoming sound wave is coded by the place along the cochlear partition with the greatest mechanical displacement.

afferent fiber A neuron that carries sensory information to the central nervous system.

efferent fiber A neuron that carries information from the central nervous system to the periphery.

FIGURE 9.13 The cochlea is tuned to different frequencies. The narrower end of the basilar membrane toward the base is stiffer and most sensitive to higher frequencies. The wider, more flexible 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*).



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 the characteristics of these auditory nerve (AN) fibers. More specifically, we will consider the quality of the 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, and that inner hair cells, on which most afferent AN fibers synapse, extend along a line traveling the length of the cochlear partition. Put these two pieces of information together, and you 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 9.14 shows **threshold tuning curves** for several AN fibers. To graph these curves, researchers insert an electrode very close to a single AN fiber, then measure how intense 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 bottommost point on the threshold tuning curve) is called the neuron's **characteristic frequency (CF)**.

Up to this point, the way 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 characteristic frequencies, the brain 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 high-intensity, complex sounds are encoded in the auditory nerve. We will consider two of the specific complications that we have to deal with, then look at one additional mechanism that the auditory system uses to convey low-frequency components of sound waves.

TWO-TONE SUPPRESSION The rate of response for an AN fiber 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**. As Figure 9.15 shows, suppression effects are particularly pronounced when the second (suppressor) tone has a lower frequency than the first tone. In other words, if we are recording from an AN fiber whose CF is 8000 Hz and we use a 8000-Hz test tone, a 1000-Hz suppressor tone has a greater effect on the neuron's firing rate than an 15,000-Hz suppressor tone has.

The suppression effect appears to be caused by mechanical changes to the basilar membrane (Rhode and Cooper, 1993). The important point for our purposes is that 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.

threshold tuning curve A map plotting the thresholds of a neuron or fiber in response to sine waves with varying frequencies at the lowest intensity that will give rise to a response.

characteristic frequency (CF) The frequency to which a particular auditory nerve fiber is most sensitive.

two-tone suppression A decrease in the firing rate of one auditory nerve fiber due to one tone, when a second tone is presented at the same time.

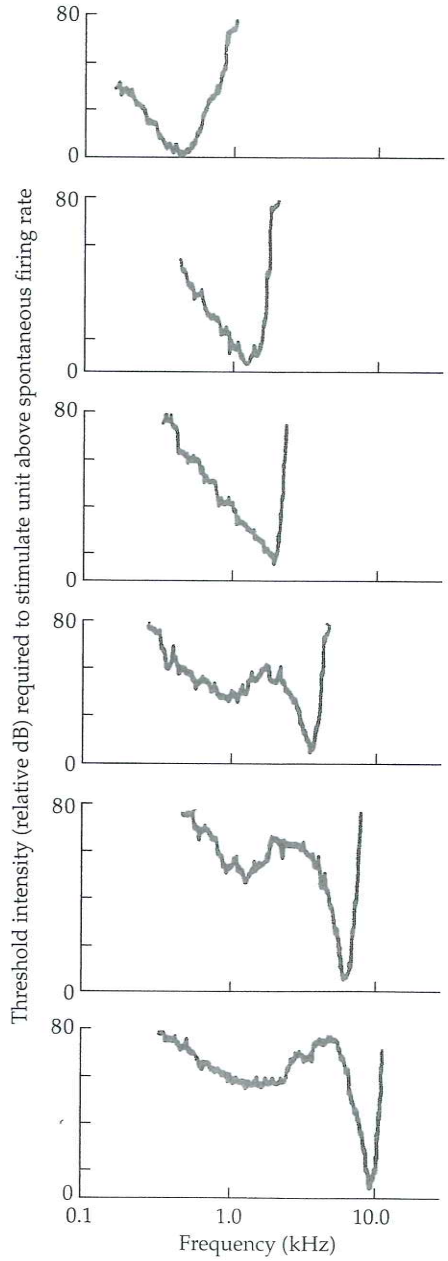
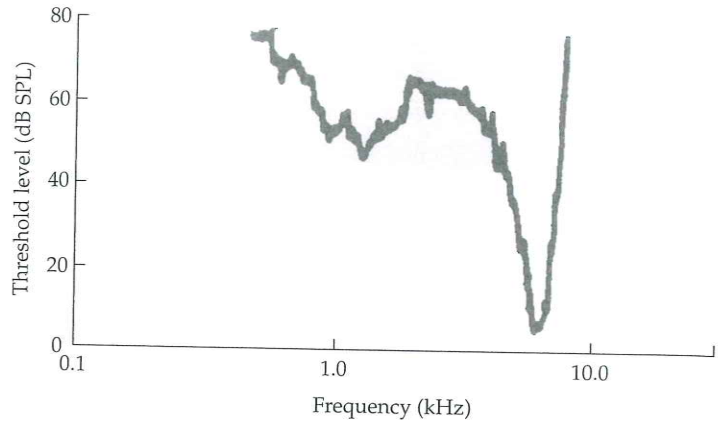


FIGURE 9.14 Threshold tuning curves for six auditory nerve fibers, each tuned to a different frequency. Curves define the lowest intensity necessary for the neuron to fire above its spontaneous rate at each frequency.

FIGURE 9.15 Two-tone suppression. The threshold tuning curve (dark red) is for one auditory nerve fiber with a characteristic frequency of 8000 Hz. Whenever a second tone is played at the frequencies and levels within the light red areas to each side, the response of this AN fiber to an 8000-Hz tone is reduced (suppressed).



isointensity curve A map plotting the firing rate of an auditory nerve fiber against varying frequencies at varying intensities.

RATE SATURATION The sounds that matter most to listeners—most conversational speech, for example—are usually heard at intensities that comfortably exceed the threshold for just detecting sounds. Are AN fibers as selective for their characteristic frequencies at levels well above threshold as they are for the barely audible sounds used to chart threshold tuning curves?

To answer this question, we can look at **isointensity curves**, which we chart by measuring an AN fiber's firing rate to a wide range of frequencies, all presented at the same intensity level. Figure 9.16 shows a family of isointensity curves for one AN fiber with a CF of 2000 Hz. The bottom curve shows the average firing rate (number of action potentials per second) of the neuron in response to 20-dB tones with frequencies between 50 and 3300 Hz. The other curves track firing rates for 40-dB, 60-dB, and 80-dB tones over the same frequency range (for higher frequencies, the neuron always fires at its spontaneous rate).

We learn from these curves that for relatively quiet, 20-dB sounds (this is about the sound level of autumn leaves rustling in the wind), the neuron is still quite selectively tuned, firing much faster in response to its CF (2000 Hz) than to neighboring frequencies. At 80 dB, however, the neuron appears to fire

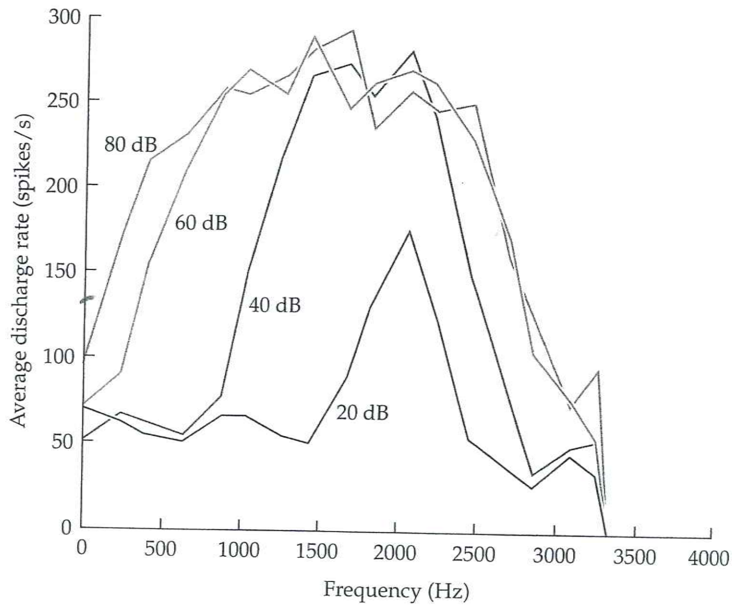


FIGURE 9.16 Isointensity functions for one auditory nerve fiber with a characteristic frequency of 2000 Hz. Tones of varying frequencies are presented at 20, 40, 60, and 80 dB. The neuron fires vigorously to a wider range of frequencies (mostly lower) when intensity is increased. Note that the 20-dB curve resembles an upside-down threshold tuning curve because 20 dB is almost as low as the intensities at which thresholds are measured.

at about the same rate for any frequency in the range of 800 to 2500 Hz. In other words, frequencies such as 1000 Hz, to which the AN fiber had almost no response at low intensity levels, evoke quite substantial responses when the intensity is increased.

The phenomenon behind this broadening of frequency selectivity is called **rate saturation**. Remember that AN fibers fire in response to the bending of stereocilia on hair cells, and that in general, the farther the cilia bend, the faster the firing rate. For a 20-dB tone at 1000 Hz, the cilia on the hair cell feeding the AN fiber featured in Figure 9.16 will not bend at all, so the fiber's firing rate remains at its resting level. The firing rate rises above this resting level when the frequency of the 20-dB tone is increased to 1700 Hz, and it reaches its highest level at the AN fiber's characteristic frequency, 2000 Hz.

When the intensity is increased to 40 dB, however, the bulge in the vestibular canal is so large that cilia start bending even to a 1000-Hz tone. If we increase the frequency to 1250 Hz, the firing rate increases, and it increases even higher at 1500 Hz. The problem is that at about 1500 Hz, the fiber's firing rate maxes out (saturates). The stereocilia are bending as far as they can at this point, so increasing the frequency of the tone has no additional effect on the AN fiber's firing rate until we increase the frequency above the fiber's CF and the firing rate starts dropping again. For 80-dB tones, the fiber maxes out at even lower and higher frequencies relative to the CF of 2000 Hz.

For moderately intense tones, then, 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 fire at its maximum rate to a 1000-Hz tone if the sound wave has a large enough amplitude.

One way in which 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 six fibers, all of which listen to the same hair cell (remember that dendrites from 10 to 30 auditory neurons synapse with each inner hair cell). To plot these curves, the intensity level of a tone at the AN fiber's CF is slowly raised from 0 dB up to 90 dB. The neurons whose functions are plotted in red are **low-spontaneous fibers**. As you can see, the resting rates of these fibers are less than 10 spikes per second. The blue lines plot

rate saturation The point at which a nerve fiber is firing as rapidly as possible and further stimulation is incapable of increasing the firing rate.

rate–intensity function A map plotting the firing rate of an auditory nerve fiber in response to a sound of constant frequency at increasing intensities.

low-spontaneous fibers Auditory nerve fibers with low rates (<10 per second) of spontaneous firing; low-spontaneous fibers require relatively intense sound before firing at higher rates.

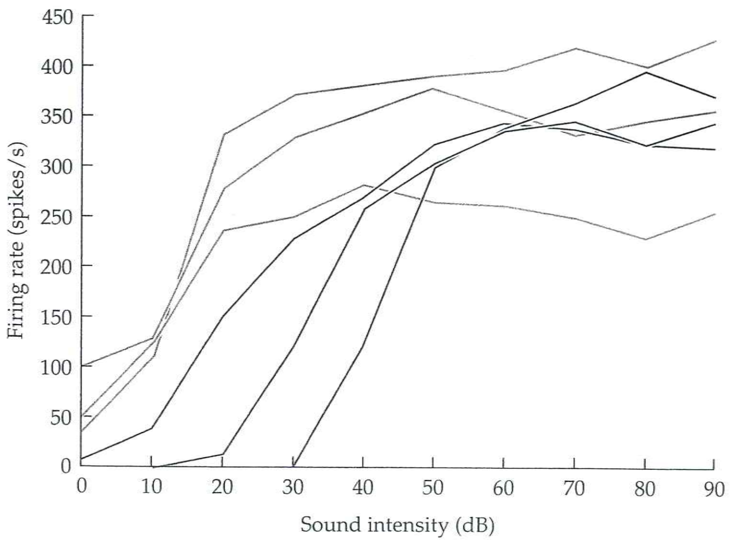


FIGURE 9.17 Firing rate plotted against sound intensity for six auditory nerve fibers: three low-spontaneous (red) and three high-spontaneous (blue). Firing rates for all six neurons increase with increasing sound level. Low-spontaneous neurons require higher-intensity sounds before they begin to fire, and they continue to increase firing rate to higher sound levels.

high-spontaneous fibers Auditory nerve fibers with high rates (>30 per second) of spontaneous firing; high-spontaneous fibers increase their firing rate in response to relatively low levels of sound.

mid-spontaneous fibers Auditory nerve fibers with medium rates (10–30 per second) of spontaneous firing. The characteristics of mid-spontaneous fibers are intermediate between *low-spontaneous fibers* and *high-spontaneous fibers*.

phase locking Firing of a single neuron at one distinct point in the period (cycle) of a sound wave at a given frequency. (The neuron need not fire on every cycle, but each firing would occur at the same point in the cycle).

temporal code Tuning of different parts of the cochlea to different frequencies, in which information about the particular frequency of an incoming sound wave is coded by the timing of neural firing as it relates to the period of the sound.

firing rates for **high-spontaneous fibers**, which fire 30 or more times per second even in silence. There are also **mid-spontaneous fibers** with resting rates between these levels.

High-spontaneous auditory nerve fibers are somewhat analogous to rods in the retina: they are especially sensitive to low levels of sound, responding with above-resting rates even when decibel levels are quite low. The tradeoff is that the firing rate of these fibers quickly reaches saturation, so their frequency selectivity is relatively poor when intensity is relatively high. Low-spontaneous 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 deploying different AN fibers with different spontaneous rates, the auditory system accurately determines the frequency of incoming sound waves by integrating the information across a broad range of AN fibers and considering the *pattern* of firing rates across all these fibers. Remember that in the visual system, we use the pattern of firing across our three types of cones to calculate the wavelength of a light ray. The auditory system uses the same principle, but it has some 14,000 auditory nerve fibers in each ear to work with. The calculations involved are obviously more complex, but the end result is even more impressive: the frequency sensitivity of the human auditory system as a whole is exquisite across a wide range of intensity levels, despite the coarse selectivity of individual auditory nerve fibers.

THE TEMPORAL CODE FOR SOUND FREQUENCY In addition to the cochlear place code, the auditory system has another way to encode frequency. As Figure 9.18 illustrates, many auditory nerve 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 they do not fire when the stereocilia move in the other direction.

The existence of phase locking means that the firing pattern of an AN fiber carries a **temporal code** for the sound wave frequency. For example, if the AN fiber fires an action potential every 0.01 (one one-hundredth) of a second, then downstream neurons listening to the AN fiber can infer that the sound wave includes a frequency component of $1/0.01 = 100$ Hz.

Temporal coding becomes inconsistent for frequencies higher than 1000 Hz and is virtually absent above 4000 or 5000 Hz. In large part, this is because of the refractory period of the AN fiber: for high frequencies, the fibers simply cannot produce action potentials quickly enough to fire on every cycle of the sound. However, multiple neurons could, in principle, encode higher fre-

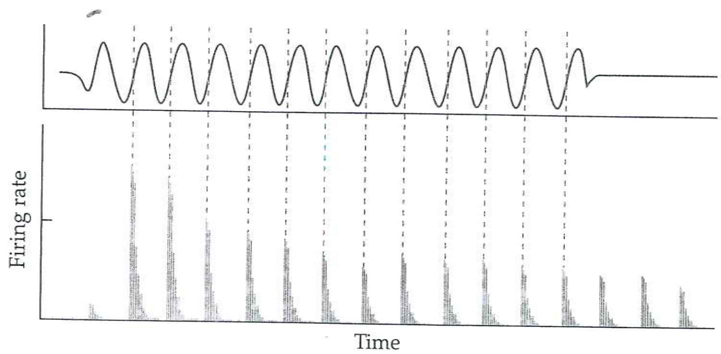


FIGURE 9.18 Histogram (bottom) showing neural spikes for an auditory nerve fiber in response to the same low-frequency sine wave (top) being played many times. Note that the neuron is most likely to fire at one particular phase of each cycle of the sine wave. This phase locking provides a temporal code to sound frequency.

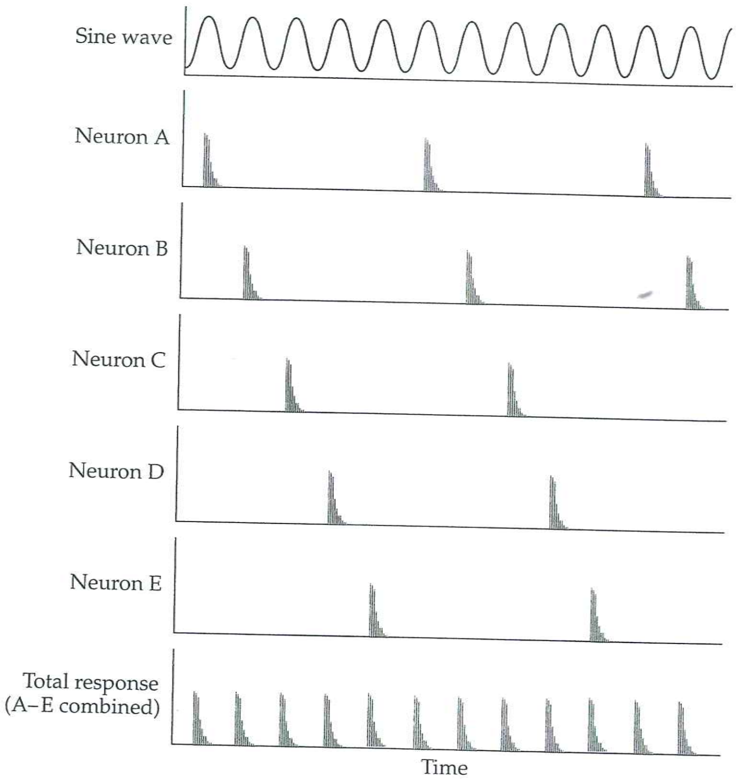


FIGURE 9.19 The volley principle. Even if one neuron cannot fire in response to every cycle of a higher-frequency tone, multiple auditory nerve fibers together can provide a temporal code for frequency if different neurons (A, B, C, D, E) each fire at different periods of the sine wave.

quencies 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 joint temporal pattern. This idea has a long history (Wever, 1949), and it is referred to as the **volley principle** (Figure 9.19). Neurons were hypothesized to 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.

Interestingly, 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 wildly 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 both about the high-frequency component (via place coding, because the brain knows the neuron's CF) and the low-frequency component (via temporal coding).

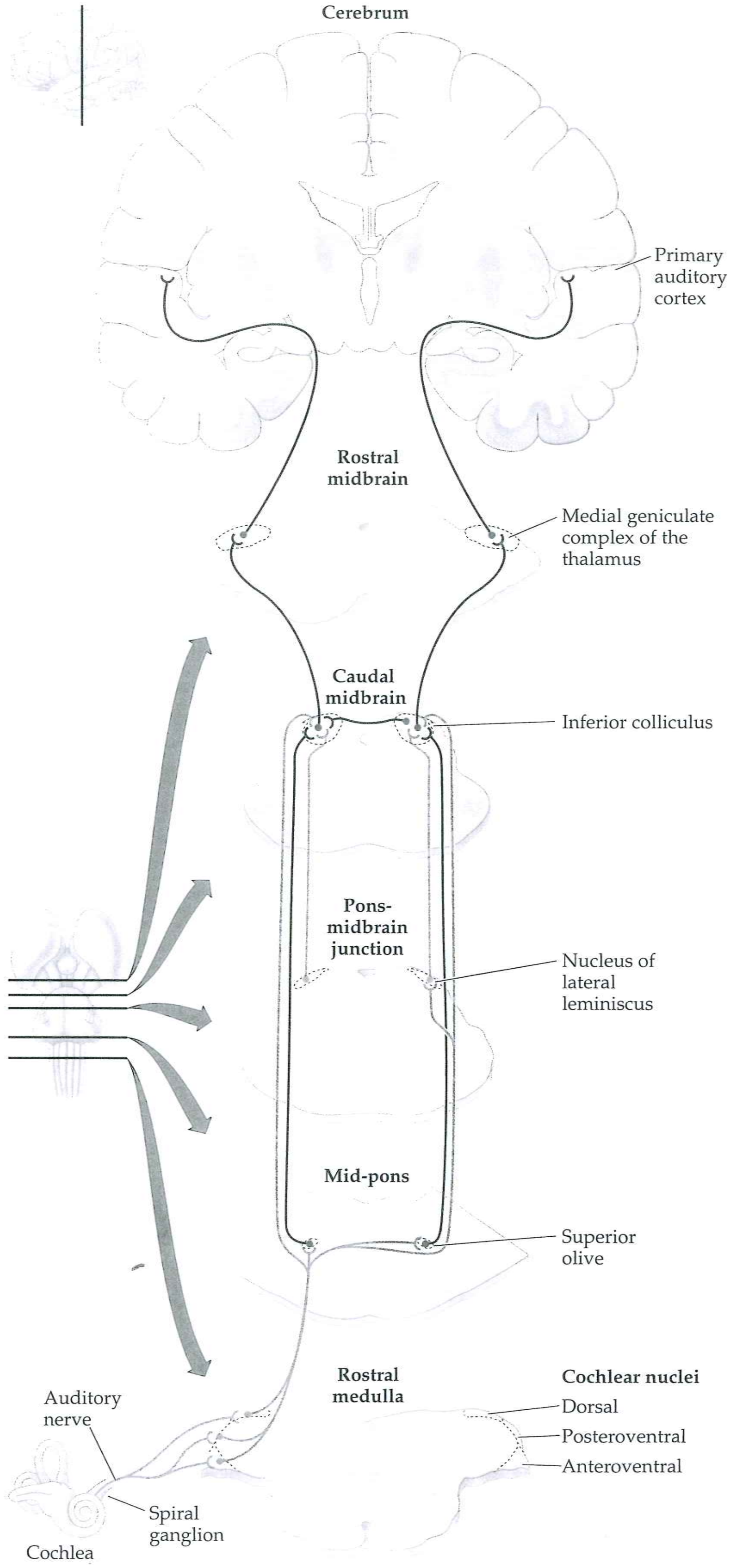
Auditory Brain Structures

The auditory nerve, also known as cranial nerve VIII, carries signals from the cochlea to the brain stem, where all AN fibers initially synapse in the **cochlear nucleus** (Figure 9.20). Some neurons in this nucleus are especially sensitive to just the onsets of sounds. Some are sensitive to the coincidence of onsets across many frequencies (that is, they fire when multiple frequencies are initially heard, but stop firing if the sound continues playing). Some cochlear nucleus neurons use lateral inhibition to sharpen the tuning to one frequency by sup-

volley principle An idea stating 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.

cochlear nucleus The first brain stem nucleus at which afferent auditory nerve fibers synapse.

FIGURE 9.20 Schematic of auditory system pathways. Not all pathways are shown. Although there are two parallel pathways, note that information from both ears comes together very early in the auditory system, at the superior olives.



pressing nearby frequencies—a mechanism reminiscent of that used by retinal ganglion cells to respond to spots of light instead of broad fields of light. Other cochlear nucleus neurons 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 brain stem nucleus.

As Figure 9.20 shows, some of the neurons that project from the cochlear nuclei 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 reach fairly far 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 from both ears is essential to using tiny timing differences between the two ears to detect the source of a sound.

Neurons from the cochlear nucleus and superior olive travel up the brain stem 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 primarily 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. As is the case in the lateral geniculate of the visual system, many more neurons project from the cortex to the medial geniculate than project from the medial geniculate to the cortex. 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 olive, inferior colliculus, and medial geniculate, show a consistent organizational pattern in which neurons are aligned respective to 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 frequency composition of sounds for auditory perception.

Tonotopic organization is maintained in the **primary auditory cortex**, which is referred to as **A1**, just as primary visual cortex is called V1. Neurons from A1 project to the surrounding **belt area** of cortex, and neurons from this belt synapse with neurons in the adjacent **parabelt area** (Figure 9.21). Just about any sound will cause activation in some part of A1. In the belt and parabelt areas, referred to as "secondary" or "associational" auditory areas, simple sounds such as sine waves and white noise elicit less activity, particularly if the stimuli have limited temporal structure. Thus we see that, as in other sensory systems, processing proceeds from simpler to more complex stimuli as we move farther along the auditory pathway. We also find greater evidence of cross-modal processing (e.g., combining acoustic and optic information), particularly in parabelt areas.

superior olive An early brain stem region in the auditory pathway where inputs from both ears converge.

inferior colliculus A midbrain nucleus in the auditory pathway.

medial geniculate nucleus Part of the thalamus that relays auditory signals to the temporal cortex and also receives input from the auditory cortex.

tonotopic organization An arrangement in which neurons that respond to different frequencies are organized anatomically in order of frequency.

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.

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.

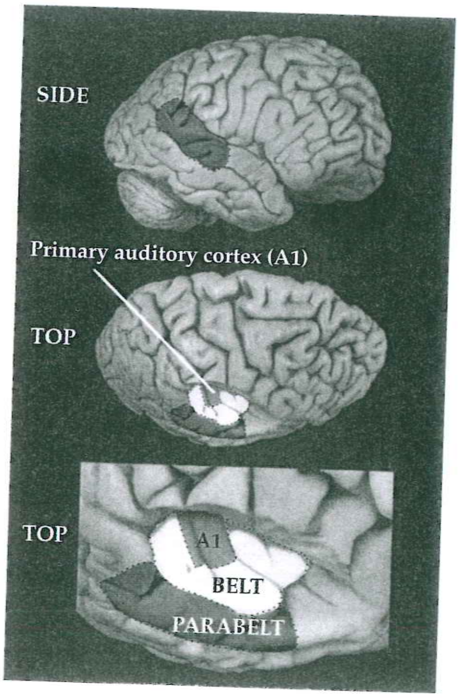


FIGURE 9.21 The first stages of auditory processing begin in the temporal lobe in areas within the Sylvian fissure. The top picture is from the side of the brain, and the lower two pictures are looking down at the brain with the parietal cortex cut away. Primary auditory cortex (A1) is in the center.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

Comparing the overall structure of the auditory and visual systems shows that a relatively large proportion of the processing in the auditory system is done before A1, whereas the majority of the most important visual processing occurs in cortical areas V1 and beyond. This difference may be a by-product of the evolutionary history of the two senses. The earliest mammals were small animals that were probably a favored prey for dinosaurs and early birds. They were also nocturnal, so hearing, smell, and touch were their primary senses. Survival may have required these mammals to become proficient listeners primarily on the basis of brain stem processes, because the stem was just about all there was to the brain in those days. The shift of some mammals to a diurnal lifestyle coincided evolutionarily with an enormous expansion of the cortex, so it makes sense that the new visual capabilities developed later by mammals developed largely in the cortex. Interestingly, one auditory capability that we know to be a very recent evolutionary advance, speech processing, is subserved almost entirely by cortical areas. We will return to the role of the cortex in auditory processing when we discuss speech and music perception in Chapter 11.

Basic Operating Characteristics of the Auditory System

Up to this point we have discussed the anatomy of the auditory system and the physiology of how the system encodes the two basic physical attributes of sound waves: amplitude (intensity) and frequency. As we've seen, these issues are investigated by direct observation of the anatomical structures: scientists examine basilar membranes, record from auditory nerve fibers, dissect various brain structures, and so on.

We turn now to the findings of researchers who have approached the auditory system from a different perspective. Instead of playing a sound and trying to determine how one particular neuron responds, we can instead play the sound and ask actual human listeners—each of whom is the sum total of a great many neurons—what they hear. When human listeners are asked to report their auditory sensations, their answers are due partly to the acoustic properties of the sound signal and partly psychological characteristics of the listeners. This method of investigation is thus called **psychoacoustics**.

Scientists who study psychoacoustics (psychoacousticians) are always careful to distinguish between 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 do, we could treat 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 mechanical listening devices (sound level meters and spectrum analyzers) and biological listening devices (human beings) provides great insight into how the human auditory system works.

psychoacoustics The study of the psychological correlates of the physical dimensions of acoustics; a branch of psychophysics.

Intensity and Loudness

The bottommost curved line 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). Note that the best (lowest) absolute thresholds for human hearing are between 2000 and 6000 Hz (remember that these frequencies are enhanced by the physical properties of the auditory 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**. We obtain these curves by asking listeners to equate the loudness of sounds with different frequencies. The starting point for each curve is always 1000 Hz, so the curve marked "40" tracks 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 levels necessary to match a 1000-Hz, 60-dB tone; and so on. As you can see, the same pattern of frequency-dependent sensitivity that we see in the audibility curve extends to sounds above threshold. (See Web Activity 9.4 Equal Loudness Curves.)

If you look at the orange tick marks in Figure 9.22, you will find that a 100-Hz tone presented at 60 dB sounds about as loud as a 1000-Hz tone presented at 50 dB (that is, both points fall on the equal-loudness curve marked "50"), whereas a 4000-Hz tone presented at 80 dB sounds louder than a 1000-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

audibility threshold A map of just barely audible tones of varying frequencies.

equal-loudness curve A map plotting sound pressure level (dB SPL) against the frequency for which a listener perceives constant loudness.

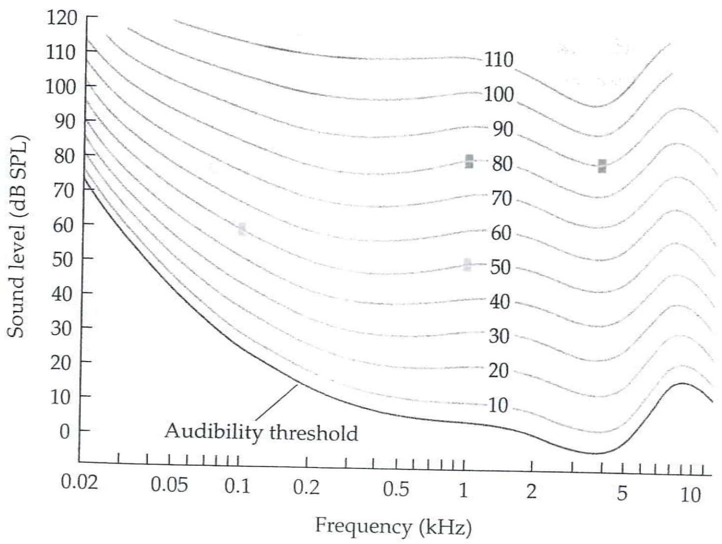


FIGURE 9.22 The lowest curve (red) illustrates the threshold for hearing sounds at varying frequencies. The curves labeled 10, 20, 30, and so on are equal-loudness curves. For a single contour, tones of different frequencies have different physical intensities, but they sound equally loud.

temporal integration The process by which a sound at a constant level is perceived as being louder when it is of greater duration. The term also applies to perceived brightness, which depends on the duration of light.

light. These power functions serve to extend the limited range of biological sensory systems across wider ranges of physical energy in the environment.

The loudness of a sound also depends on how long the sound is: within limits, longer sounds are heard as being louder. Again, the same thing happens in vision: flashes of light appear brighter when they are longer. The general phenomenon is due to the fact 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 to 200 ms. So if a sound is presented for less than 100 ms, it will be perceived as softer than 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 have been 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 about 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. As we noted earlier, 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. However, when researchers recorded from individual AN fibers, they found that the difference between the intensity at which a neuron just starts firing and the intensity at which the neuron's firing rate saturates is typically less than 30 dB—a much smaller window than the 100-dB range over which humans can detect loudness differences.

The auditory system has a number of ways to get around this limitation. Perhaps most importantly, intensity 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 15 to 40 dB range, a third might cover 38 to 65 dB, and so on (see Figure 9.17). A population of neurons with different thresholds can then 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. One result of this 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 was 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 (that is, an increase of 50 Hz), listeners will perceive a greater rise in pitch for lower frequencies than they do for higher frequencies. For example, listeners perceive a greater pitch difference when a tone shifts from 500 to 1000 Hz than when a tone shifts from 5000 to 5500 Hz.

Research done using pure tones (sounds 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 999-Hz and 1000-Hz tones—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. Part of the reason that discriminability decreases for high pitches appears to be that the temporal code for frequencies starts breaking down above 1000 Hz and is relatively nonfunctional above 5000 Hz. Performance thus suffers because the auditory system has to rely exclusively on place coding (the pattern of AN fiber responses from different parts of the cochlea).

Psychoacousticians have also used **masking** experiments to investigate frequency selectivity. In the research described in the previous paragraph, listeners never hear more than 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. That is, we look at how effective one sound—the masker—is at hiding another sound. We already discussed a phenomenon called "two-tone suppression," in which an auditory nerve's response to one frequency can be disrupted by a second frequency. If place coding underlies pitch perception, we should expect to see the same kind of phenomenon in masking experiments.

The classic approach to measuring frequency selectivity using masking involves placing a single sine wave tone 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 to 1500 Hz; an even smaller band could span 500 to 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 to 2025 Hz (Figure 9.23). 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 you 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 energy 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 9.23*a*). For the experiment whose data are plotted in Figure 9.23*b*, the critical bandwidth is 400 Hz. Thus, 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 to 200 to 400 Hz, but going from a 400-Hz noise band to an 800-Hz 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 to cement the role of place coding in pitch perception by revealing similarities between perceptual effects and physiological findings. First of all, the width of the critical band-

masking Using a second sound, frequently noise, to make the detection of another sound more difficult.

white noise Noise consisting of all audible frequencies in equal amounts. White noise in hearing is analogous to white light in vision, for which all wavelengths are present.

critical bandwidth The range of frequencies that are conveyed within a channel in the auditory system.

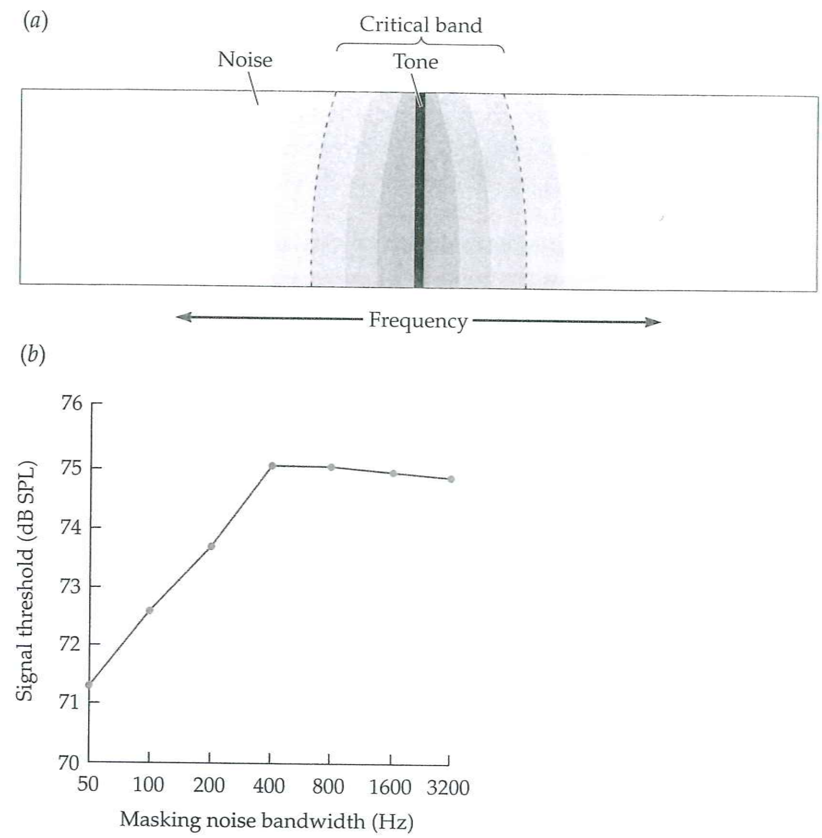


FIGURE 9.23 Critical bandwidth and masking. (*a*) To measure the width of a critical band, subjects listen for a tone in the center of a band of noise of constant intensity. It is harder to detect the tone as the band of noise widens up to a point, when further widening has no effect on detecting the tone. This point defines the width of a critical band. (*b*) A plot of bandwidth data when a 2000-Hz tone is used (Schoonevelt and Moore, 1989). The bandwidth of the noise has no effect on detecting the tone when it exceeds 400 Hz, so 400 Hz is the width of the critical band at 2000 Hz.

width changes depending on the frequency of the test tone, and these widths correspond to the physical spacing of frequencies along the basilar membrane. For example, we know that a greater proportion of the basilar membrane vibrates in response to low frequencies, and that 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.

Another important finding is that masking effects are asymmetrical. More specifically, for a mask whose bandwidth is below the critical bandwidth of a test tone, the mask is more effective if it is centered on a frequency *below* the test tone's frequency—a phenomenon called the "upward spread of masking." This may seem counterintuitive, but if you look back at Figure 9.13, you can see 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. In addition, two-tone suppression experiments indicate that suppression is greater if the suppressor tone is slightly below an AN fiber's characteristic frequency than if the suppressor tone is slightly above the fiber's characteristic frequency (see Figure 9.15).

Hearing Loss

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 going on in the environment. As we will see, common forms of hearing loss 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 at times, by wearing earplugs. A less intentional hearing loss can be created by the excessive buildup of ear wax (cerumen) in the ear canal. This problem is easy to remedy, as long as the effort to clear out the ear canal does not damage the tympanic membrane.

Another type of hearing impairment, called **conductive hearing loss**, occurs when the middle-ear bones lose (or are impaired in) their ability to freely convey (conduct) vibrations from the tympanic membrane to the oval window. This happens 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 **sensorineural hearing loss**, which occurs inside the cochlea, or sometimes as a result of damage to the auditory nerve. Most often, sensorineural hearing loss occurs when hair cells are injured. For example, certain antibiotics and cancer drugs are **ototoxic**, meaning that they kill hair cells directly. Physicians are well aware of these properties and typically avoid using such drugs, but sometimes a patient faces the decision of life with deafness versus no life at all.

A more mundane cause of hearing loss is damage to the hair cells by excessive exposure to noise. It is fairly well known that shooting a gun without ear protection can cause hearing loss. Extended exposure to loud sounds such as the noise of factory equipment also causes hearing loss. It is no coincidence that so many aging rock stars and race car drivers now wear hearing aids. (It is ironic, but wise, that many heavy-metal music fans now wear ear protection at concerts.) However, evidence suggests that cumulative exposure to everyday noises present in the environments of industrialized countries can also cause hearing loss. In one study (Goycoolea et al., 1986), middle-aged and elderly residents of Easter Island who stayed almost exclusively on their quiet island their whole lives were compared to other Easter Islanders, who made more frequent trips to the noisier outside world. As Figure 9.24*a* shows, the more time people spent off the island, the more hearing loss they experienced.

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, age-related hearing loss first affects the perception of high frequencies (Figure 9.24*b*). 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 the ability to hear frequencies above 15,000 Hz. The decrease

conductive hearing loss Hearing loss caused by problems with the bones of the middle ear.

otitis media Inflammation of the middle ear, commonly in children as a result of infection.

otosclerosis Abnormal growth of middle-ear bones that causes hearing loss.

sensorineural hearing loss Hearing loss due to defects in the cochlea or auditory nerve.

ototoxic Producing adverse effects on organs or nerves involved in hearing or balance.

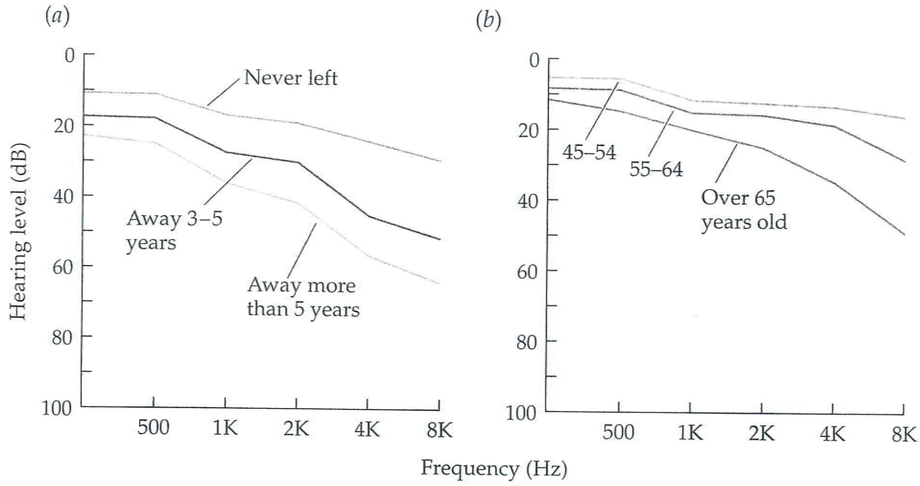


FIGURE 9.24 (*a*) The more people are exposed to noise in their environment, the more their hearing deteriorates. People who spent more years off of relatively quiet Easter Island had relatively worse hearing. (*b*) Hearing becomes less sensitive, particularly for high frequencies, as people get older. (After Goycoolea et al., 1986.)

in the ability to hear higher-frequency sounds continues throughout life, with the highest audible frequency becoming lower and lower as we get older. Fortunately, many of the auditory stimuli that people care most about, including speech and music, are composed predominantly of lower frequencies.

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 of the horn 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 they are at least 70 dB intense, compared to 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 9.25). Nearly all modern hearing aids use some means to amplify the signal while also compressing intensity differences in order to keep the highest intensities within a

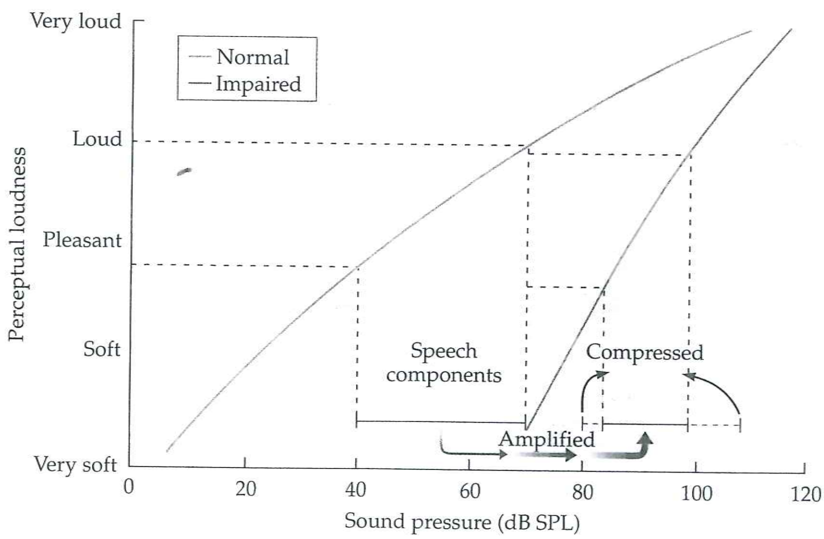


FIGURE 9.25 When thresholds are increased due to hearing impairment, it takes more energy to hear a sound, but loudness increases faster than it does with healthy ears. When a hearing aid is used to increase the intensity of sounds, all variation in sound levels must be compressed into a smaller range 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 Geisler, 1998.)

comfortable listening level. Newer 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).

One advantage of the old horns over electronic hearing aids was that they permitted listeners to "focus" 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. Thus 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 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 improving all the time, and they have provided relief to millions of Americans, including 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 (but see Web Essay 9.1 Cochlear Implants for the latest attempts by medical science to produce true bionic ears). The best advice is to never need a hearing aid: protect your hair cells by avoiding exposure to loud sounds and by using hearing protection such as earplugs or muffs when necessary.

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 perception 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 are flexed by 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. There are multiple places in the brain stem where different characteristics of sounds are processed 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. Hearing across such a wide range of intensities is accomplished by the use of many auditory neurons. Some neurons respond across certain levels of intensity; others span different levels of intensity. In addition, more neurons overall respond when sounds are more intense.
6. A series of channels (or filters) processes 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, they cannot restore hearing as well as glasses can improve vision.

Refer to the ***Sensation and Perception*** website (www.sinauer.com/wolfe) for activities, essays, study questions, and other study aids.

Robert ParkeHarrison, *The Lesson*, 1996

Chapter 10