CHAPTER

9

How Do We Hear, Speak, and Make Music?

Barbara Haynor/Tony Stone
Micrograph: Dr. Dennis Kunkel/Phototake

In 1995, Ivan Turk, a paleontologist at the Slovenian Academy of Sciences in Ljubljana, excavated a cave in northern Slovenia that Neanderthals once used as a hunting camp. Neanderthals have long fascinated researchers, in part because they are considered to be the last truly "primitive" members of the genus *Homo*. Neanderthals (*Homo neanderthalensis*) originated approximately 230,000 years ago and disappeared some 200,000 years later. During that time, they coexisted with early modern humans (*Homo sapiens*), whom they resembled in many ways. In some locations, the two groups may have even shared resources and tools. But researchers had assumed that Neanderthal culture was significantly less developed than that of early modern humans. Although the Neanderthals did bury their dead with artifacts, which implies that they may have held spiritual beliefs, there is no evidence that they made cave paintings, as *Homo sapiens* did beginning near the end of the Neanderthal era. Perhaps even more important, some skeletal analyses suggest that, whatever language ability Neanderthals had, they were far less fluent than the *Homo sapiens* who lived at the same time.

This view of the Neanderthals as culturally "primitive" is what made one of Turk's discoveries so surprising and intriguing. Buried in the cave among a cache of stone tools was the leg bone of a young bear that looked as if it had been fashioned into a flute. As shown in Figure 9-1, the bone had holes aligned along one of its sides that could not have been made by gnawing animals. The holes were spaced unevenly, resembling the holes on a modern flute. But this flute was at least 43,000 years old—perhaps as old as 82,000 years. All the evidence suggested that Neanderthals, not modern humans, had made the instrument. Was music, then, a creation not just of *Homo sapiens* but of *Homo neanderthalensis,* too?

To help find out, Bob Fink, a musicologist, analyzed the flute for its musical capabilities. He found that a scale could be played on the flute, similar to our do-re-mi scale; but relative to the scale most familiar to us, one of the notes was slightly off. That note is called a "blue note" in jazz and is found in scales throughout Africa and India.

The similarity between the Neanderthal musical scale and one of our modern musical scales encourages us to speculate about the brain that made this ancient flute. The human brain has a specialized region in the right temporal lobe for analyzing music, and the Neanderthal flute suggests that its maker had this region, too. Furthermore, in the modern human brain, the specialization for analyzing music in the right temporal lobe is complemented by a specialization for analyzing speech

Figure 9-1
A piece of bear femur found in the cemented sediment on the ancient floor of a cave in Slovenia, apparently used by Neanderthals as a hunting camp. The alignment of the holes suggests that they were made not by gnawing animals but by ancient people. An analysis of the distances between the holes indicates that they match those of a modern-day flute, which suggests that Neanderthals made music.

Courtesy of Ivan Turk/Institut 2A Archeologijo, ZRC-Sazu, Slovenia. Photograph by Marko Zaplatil.

sounds in the left temporal lobe. We are therefore left to wonder if the evolutionary development of music and language may have been simultaneous. In other words, early in human evolution, the left temporal lobe may have become specialized for analyzing speech sounds while the right temporal lobe became specialized for analyzing musical sounds. If this is so, then perhaps the Neanderthals had the beginnings of a left-hemisphere specialization for language in addition to the start of a right-hemisphere specialization for music. Consequently, they may have had both more language skills and more cultural development than formerly assumed.

Spoken language and musical ability are not just linked as complementary systems in the brain. They are also linked conceptually because both are based on the use of sound. So, to understand how humans engage in each of these behaviors, we must first examine the nature of sound and how the human ear and nervous system are structured to detect it. We then ask how the human brain is designed to analyze both language and music. Finally, we look at two examples of how other species use auditory stimuli for specific functions: communication through song by birds and navigation through sonar by bats.

AN INTRODUCTION TO SOUND

Hertz. A measure of frequency; one hertz is equal to one cycle per second.

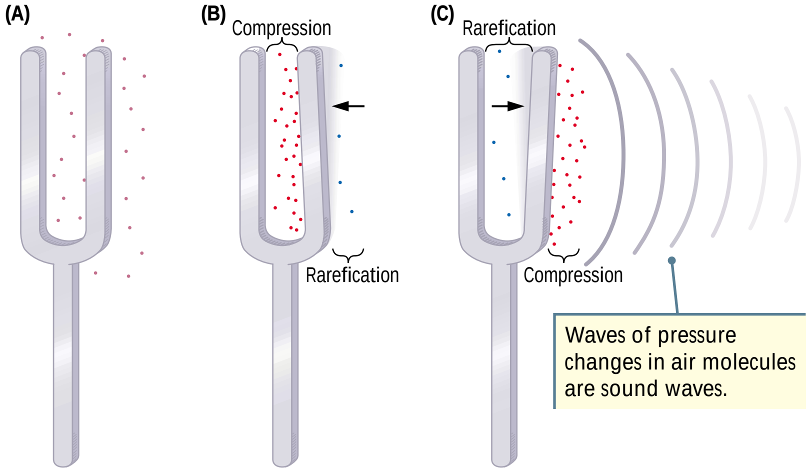
Sound is a creation of the brain and does not exist without it, as you learned in Chapter 8. When a tree falls in the forest, it makes no sound unless someone is there to hear it. What a falling tree makes are merely changes in air pressure. These pressure changes take the form of waves generated by vibrating air molecules.

One way to produce these so-called *sound waves* is to strike a tuning fork. The vibrating prongs of the fork displace the adjacent air molecules. Figure 9-2 shows that, as one prong moves to the left, it compresses (makes more dense) the air molecules to the left of it and rarefies (makes less dense) the air molecules to the right of it. The opposite happens when the prong moves to the right. This displacement of air molecules causes waves of changing air pressure to emanate from the fork. These waves are sound waves.

We can represent the waves of changing air pressure emanating from the tuning fork by plotting air-molecule density against time at a single point at the right-hand side of the fork. Such a graph is shown at the top of Figure 9-3. The bottom graph in

Figure 9-2

How a tuning fork produces sound. **(A)** When the fork is still, the air molecules are distributed randomly around it. **(B)** The right arm of the fork moves to the left, causing the air to be compressed on the leading edge and rarefied on the trailing edge. **(C)** The arm moves to the right, compressing the air to the right and rarefying the air to the left.



Waves of pressure changes in air molecules are sound waves.

Figure 9-3 shows how the right-hand prong of the fork moves to create the air-pressure changes associated with a single cycle. A *cycle* is one complete peak or valley on the graph—that is, the change from one maximum or minimum air-pressure level to the next.

The Basic Qualities of Sound Waves

The sound waves produced by the displacement of air have three qualities: frequency, amplitude, and complexity. These properties are summarized in Figure 9-4. The auditory system analyzes each of these properties separately, just as the visual system analyzes color and form separately. After examining these three properties of sound in some detail, we will consider how we perceive sounds and interpret them as language and music.

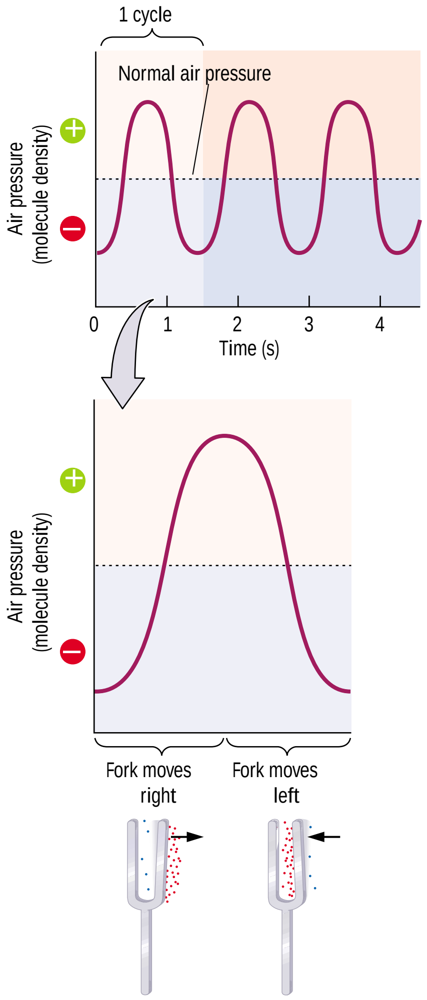
SOUND FREQUENCY

Although sound travels at a fixed speed of 1100 feet per second, sound waves vary in the rate at which they vibrate, called their frequency. More precisely, **frequency** refers to the number of cycles of a wave that are completed in a given amount of time. Sound-wave frequencies are measured in units of cycles per second called **hertz** (Hz), named after the German physicist Heinrich Rudolph Hertz. One hertz is one cycle per second; 50 hertz is 50 cycles per second; 6000 hertz is 6000 cycles per second; 20,000 hertz is 20,000 cycles per second; and so on. Figure 9-4 shows that sounds that we perceive as being low in *pitch* have low frequencies (few cycles per second), whereas sounds that we perceive as being high in pitch have high frequencies (many cycles per second).

Just as we can perceive light of only certain wavelengths, we can perceive sounds in only a limited range of frequencies. For humans, this range is from about 20 to 20,000 hertz. Like humans, many animals produce some form of sound to communicate, which means that they must have auditory systems designed to interpret their

Figure 9-3

Air molecule density plotted against time at a particular point to the right of the tuning fork. The resulting cyclical wave is referred to as a sine wave.



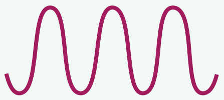
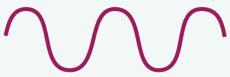
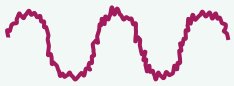
The Properties of Sound		
Frequency (pitch) The rate at which waves vibrate, measured as cycles per second, or hertz (Hz). Frequency roughly corresponds to our perception of pitch.	 Low frequency (low-pitched sound)	 High frequency (high-pitched sound)
Amplitude (loudness) The intensity of sound, usually measured in decibels (dB). Amplitude roughly corresponds to our perception of loudness.	 High amplitude (loud sound)	 Low amplitude (soft sound)
Complexity (timbre) Most sounds are a mixture of frequencies. The particular mixture determines the sound's timbre, or perceived uniqueness. Timbre provides information about the nature of a sound. For example, timbre allows us to distinguish the sound of a trombone from that of a violin playing the same note.	 Simple	 Complex

Figure 9-4

Sound has three physical dimensions: frequency, amplitude, and complexity. These dimensions correspond to the perceptual dimensions of pitch, loudness, and timbre.

species-typical sounds. After all, there is no point in making complicated songs or calls if other members of the species cannot hear them. The ranges of sound frequencies that different species use vary quite extensively. Figure 9-5 shows that some species (such as frogs and birds) have rather narrow ranges, whereas others (such as whales and humans) have broad ranges. Some species use extremely high frequencies, whereas others use low ones (compare bats and fish, for instance). It is quite an achievement for the auditory systems of whales and dolphins to be able to respond to sounds of such widely varying frequencies. The characteristics of these different frequencies allow them to be used in different ways. For example, very-low-frequency sound waves travel long distances in water, so whales produce them as a form of underwater communication over miles of distance. In contrast, high-frequency sounds create echoes and form the basis of sonar, so dolphins produce them in bursts, listening for the echoes that bounce back from objects and help the dolphins to locate fish.

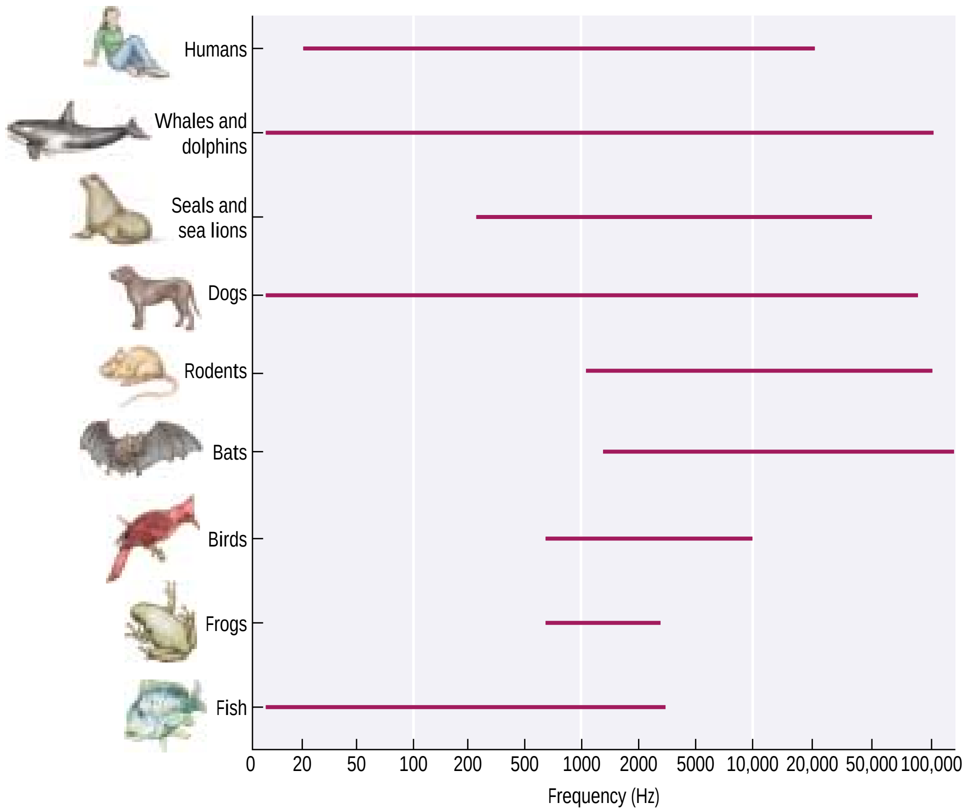


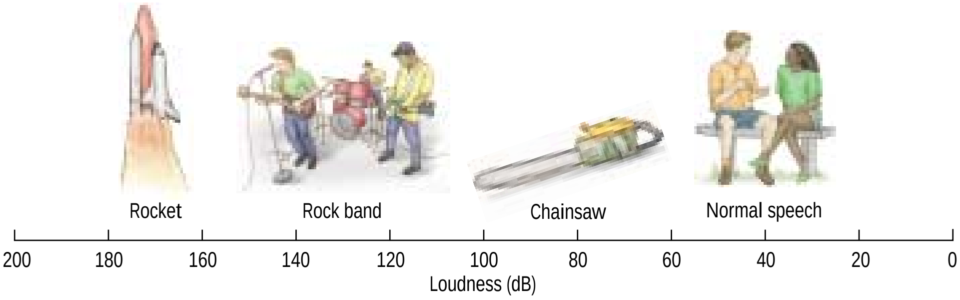
Figure 9-5

Hearing ranges of different animals. Frogs hear only a very narrow range of frequencies, whereas whales and dolphins have an extensive range. Although the human range of hearing is fairly broad, peaking at about 20,000 hertz, humans are unable to perceive many of the sounds that other animals can both make and hear.

As stated earlier, we hear differences in the frequency of sound waves as differences in pitch. Consequently, each note in the musical scale must have a different frequency because each has a different pitch. Middle C on the piano, for instance, has a frequency of 264 hertz. Although most people can discriminate between one musical note and another, some people are able to name any note (A, B flat, C sharp, and so forth) that they hear. This ability is referred to as *perfect pitch*, or *absolute pitch*. People with perfect pitch are more likely than not to have a family member who also has perfect pitch, suggesting that the ability is caused by some genetic influence on the auditory system. The difference in the auditory system may be analogous to differences in the ability to perceive the color red, discussed in Chapter 8. In addition, though, people with perfect pitch probably have early musical training that provides them with knowledge about which pitch goes with which note.

SOUND AMPLITUDE

Sound not only has variations in frequency, causing differences in perceived pitch, but also has variations in amplitude, causing differences in perceived *intensity*, or *loudness*. An example will help you understand the difference between amplitude and frequency. If you hit a tuning fork lightly, it produces a tone with a frequency of, say, 1000 hertz. If you hit it more forcefully, you still produce a frequency of 1000 hertz, but there is now more energy in the vibrating prong, so it moves farther left and right. This greater energy is due to an increased *quantity* of air molecules compacted in each wave, even though the same number of waves is created every second. The new dimension to the sound wave is its new **amplitude**, a term that refers to the magnitude of change in air-molecule density. An increased compaction of air molecules increases the amount of energy in a sound wave, which makes the sound seem louder—more amplified. Differences in sound amplitude can be graphically illustrated by increasing the height of a sound wave, as shown in Figure 9-4.



Sound amplitude is usually measured in *decibels* (dB), which describe the strength of a sound relative to a standard reference intensity. Sounds greater than about 70 decibels we perceive as loud, whereas those less than about 20 decibels we perceive as quiet. Normal speech sounds are about 40 decibels.

Because the human nervous system is sensitive to weak sounds, it is literally "blown away" by extremely strong ones. It is common for people to damage their hearing by exposure to very loud sounds (such as rifle fire at close range) or even by prolonged exposure to sounds that are only relatively loud (such as rock music in concert). As a rule of thumb, sounds louder than 100 decibels are likely to damage our hearing, especially if they are prolonged. Some bands, especially heavy-metal groups, routinely play music higher than 120 decibels and sometimes as high as 135 decibels. One researcher (Drake-Lee, 1992) found that rock musicians had a significant loss of sensitivity to sound, especially at about 6000 hertz. After a typical 90-minute concert, this loss was temporarily far worse—as much as a 40-fold increase in sound pressure was needed to reach a musician's hearing threshold.

SOUND COMPLEXITY

Sounds with a single frequency are **pure tones**, but most sounds are not single frequencies. Most sounds are made up of combinations of frequencies and so are called **complex tones**. A complex tone is a sound with a basic waveform that is repeated periodically, but a complex tone is more complicated than a pure tone. Complex tones contain at least two pure tones and often many more.

To better understand the nature of a complex tone, consider the sound of a clarinet playing a steady note. Figure 9-6 (top) represents the sound wave that the clarinet produces. Notice that this wave has a more complex pattern than those of the simple waves described earlier: even when a single note is played, the clarinet is making a complex tone, not a pure tone. Using a technique known as Fourier analysis, we can break this complex tone into its many component pure tones, as shown in Figure 9-6 (bottom). The *fundamental frequency* is the rate at which the basic pattern of the complex waveform is repeated. *Overtones* are higher-frequency waves that are multiples of the fundamental frequency. Different musical instruments produce overtones of different amplitudes. For the clarinet in Figure 9-6, wave 5 is of low amplitude, whereas wave 2 is of high amplitude.

Christian Ducasse/Gamma-Liaison

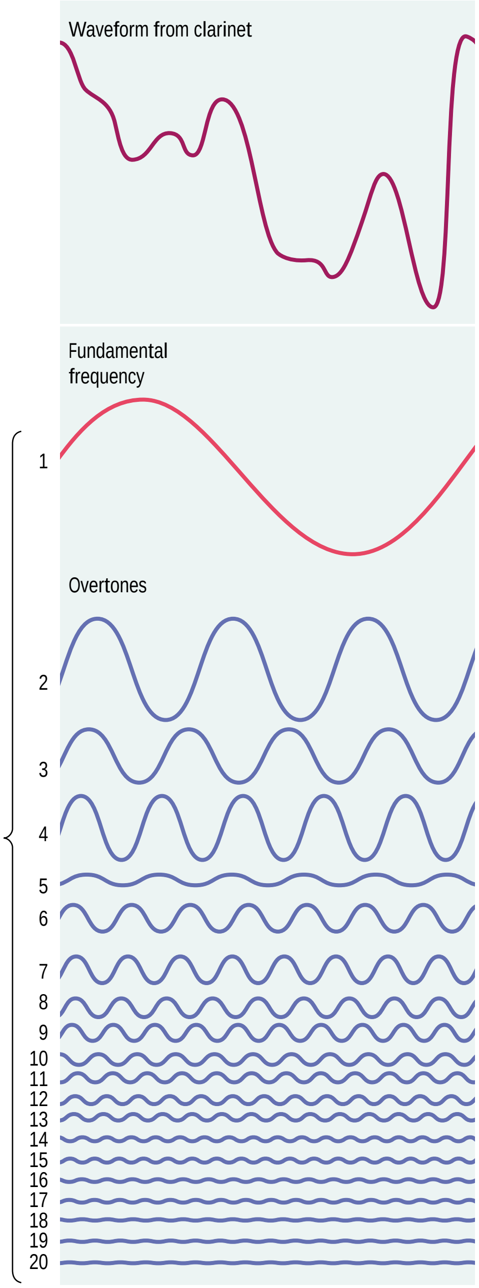


Figure 9-6

The shape of a complex sound wave from Don Byron's clarinet (*top*) and the component frequencies—the fundamental frequency (*middle*) and overtones (*bottom*)—that make up the complex sound wave.

Stereo Review, copyright 1977 by Diamandis Communications Inc.

Pure tones can be combined into complex tones in an almost infinite variety of ways. In addition to musical instruments, other examples of complex tone sources are the human voice, birdsong, and certain machines having rhythmic or repetitive mechanisms that give rise to buzzing or humming sounds. A key feature of complex tones, besides being made up of two or more pure tones, is some sort of periodicity, or repetition of the fundamental frequency. Sounds that are aperiodic, or random, are known as *noise*.

The Perception of Sound

The auditory system's task is to convert the air-pressure changes associated with sound waves into neural activity that travels to the brain. This task is accomplished through a series of steps to be described shortly. The critical point here is that waves of air-pressure changes are transformed into neural activity, which we then perceive as sounds. Remember that the waves themselves make no sounds. The sounds that we hear are only a product of the brain.

To better understand the relation between energy waves and sound perceptions, consider this analogy. When you toss a pebble into a pond, waves emanate from the point where the pebble enters the water. These waves produce no audible sound. But what would happen if your skin were able to convert the action of the waves in the water into neural activity that stimulated your auditory system? When you placed your hand into the rippling water, you would "hear" the waves, and, when you removed your hand, the "sound" would stop. The pebble hitting the water is much like a falling tree, and the waves that emanate from the pebble's point of entry are like the air-pressure waves that emanate from the place where the tree strikes the ground. The frequency of the waves determines the pitch of the sound heard by the brain, whereas the size (amplitude) of the waves determines the sound's loudness.

Our sensitivity to sound waves is extraordinary. At the threshold of hearing, we can detect the displacement of air molecules of about 10 picometers (10^{-11} meter). We are rarely in an environment where we can detect such a small air-pressure change, because there is usually too much background noise. A very quiet rural setting is probably as close as we ever get to an environment suited to testing the acuteness of our hearing. So, the next time you visit the countryside, take note of the sounds that you can hear. If there is no sound competition, you can often hear a single car engine literally miles away.

In addition to detecting very small changes in air pressure, the auditory system is also very adept at simultaneously perceiving different sounds. As you sit reading this chapter, you are able to differentiate all sorts of sounds around you, such as traffic on the street, people talking next door, your computer's fan humming, and footsteps in the hall. Similarly, as you listen to music, you are able to detect the sounds of different instruments and voices. You can perceive different sounds simultaneously because the different frequencies of air-pressure change associated with each sound stimulate different neurons in your auditory system.

The perception of sounds is only the beginning of your auditory experience. Your brain interprets sounds to obtain information about events in your environment, and it analyzes a sound's meaning. These processes are clearly illustrated in your use of sound to communicate with other people through both language and music.

Language and Music as Sounds

Language and music differ from other auditory inputs in several fundamental ways. First, both speech and musical sounds convey special meaning. The analysis of meaning in sound is considerably more complex than simply the detection of a sound's presence. To analyze speech and musical sounds for meaning, the brain has had to develop special systems. As you learned at the beginning of this chapter, these special systems lie in the left and right temporal lobes, respectively.

Another characteristic that distinguishes speech and musical sounds from other auditory inputs is the speed at which a series of these sounds arrives at our auditory receptors. Nonspeech and nonmusical noise is perceived as a buzz that occurs at a rate of about 5 segments per second. (A sound segment is a distinct unit of sound.) Normal speech is faster than this, on the order of 8 to 10 segments per second, and we are capable of understanding speech at rates of nearly 30 segments per second. Speech perception at these higher rates is truly amazing, because the speed of input far exceeds the auditory system's ability to transmit all the speech as separate pieces of information. Experience in listening to a particular language helps the brain to analyze rapid speech, which is one reason why unfamiliar languages often sound so incredibly fast when you hear them spoken. Your brain does not know where the foreign words end and begin, making them seem to run together in a rapid-fire stream.

A unique characteristic of our perception of speech sounds is our tendency to hear the variations of a sound as identical sounds, even though the sound varies considerably from one context to another. For instance, the English letter "d" is pronounced differently in the words "deep," "deck," and "duke," yet a listener perceives the pronunciations to be the same "d" sound. The auditory system must therefore have a mechanism for categorizing sounds as being the same despite small differences in pronunciation. This mechanism, moreover, must be affected by experience, because different languages categorize speech sounds differently. A major obstacle to mastering a foreign language is the difficulty of learning the categories of sound that are treated as equivalent.

Like other types of sounds, musical sounds differ from one another in the properties that we perceive them to have. One property is *loudness*, which refers to the magnitude of the sound as judged by a given person. Loudness, as you know, is related to the amplitude of a sound, which is measured in decibels, but it is also a subjective factor. What is "very loud" music for one person may be only "moderately loud" for another, whereas music that seems "soft" to one listener may not seem soft at all to someone else.

Another property of musical sounds is *pitch*, which refers to the position of each sound on a musical scale as judged by the particular listener. Although pitch is clearly related to the frequency of a sound, there is more to it than that. Consider the note middle C as played on a piano. This note can be described as a pattern of sound frequencies, as is the clarinet note in Figure 9-6. Like the note on the clarinet, any musical note is defined by its fundamental frequency, which is the lowest frequency of the sound pattern, or the rate at which the overall pattern is repeated. For middle C, the fundamental frequency is 264 hertz, as illustrated in Figure 9-7. An important feature of our brain's analysis of music is that middle C is perceived as being the same note regardless of whether it is played on a piano or a guitar, even though the sound is very different on different instruments. The right temporal lobe has a special function in extracting pitch from sound, whether the sound is in speech or music. In speech, pitch contributes to perceived "tone" of voice, which is known as prosody.

A final property of musical sound is its *quality*, which refers to the characteristics that distinguish a particular sound from all others of similar pitch and loudness. For example, we can easily distinguish the sound of a violin from that of a trombone even though the same note is being played on both instruments at the same loudness. We can do so because the two sounds differ in quality. The French word *timbre* [pronounced *tam brah*] is normally used to describe this characteristic of sound.

Prosody. Melody or tone of voice.

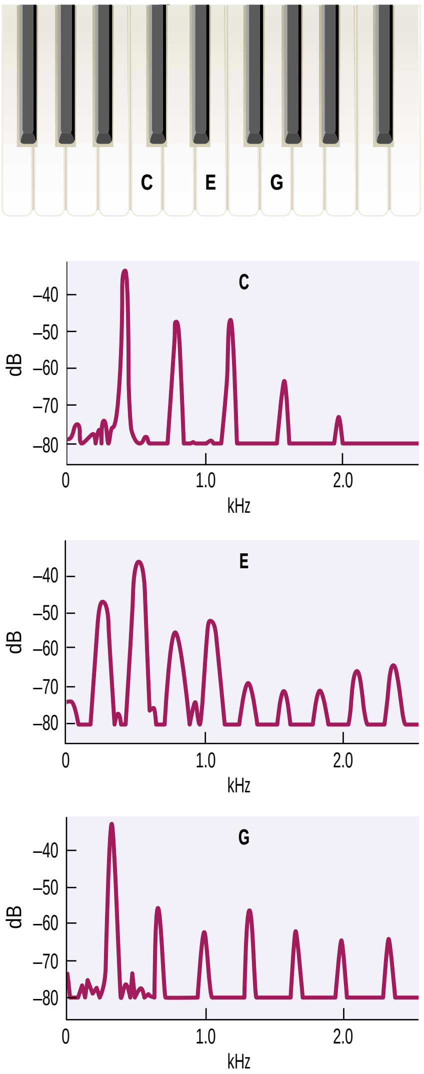


Figure 9-7

The shapes of the sound waves made by different musical notes.

Courtesy of D. Rendall.

In Review

Sound, the stimulus for the auditory system, is produced by changes in air pressure, which form complex waves that are converted into neural activity in the ear. Sound waves have three key qualities: frequency, amplitude, and complexity. Frequency is the rate at which the waves vibrate and roughly corresponds to perceived pitch. Amplitude is the magnitude of change in air-molecule density that the wave undergoes and roughly corresponds to perceived loudness. Complexity refers to the particular mixture of frequencies that create a sound's perceived uniqueness, or timbre. Combinations of these qualities allow the auditory system to comprehend sounds as complex as language and music. Our next task is to explain how this analysis of sound is accomplished. We begin by tracing the pathway taken by sound through the brain.

THE ANATOMY OF THE AUDITORY SYSTEM

The ear collects sound waves from the surrounding world and converts them into neural activity, which then begins a long route through the brainstem to the auditory cortex. Before we can trace the journey from the ear to the cortex, we need to ask what the auditory system is designed to do. Because sound has the properties of frequency, amplitude, and complexity, we can assume that the auditory system is structured to code these properties. In addition, most animals can tell where a sound comes from, so there must be some mechanism for locating sounds in space. Finally, many animals, including humans, not only analyze sounds for their meanings but also make sounds themselves. Because the sounds that they produce are often the same as the ones that they hear, we can infer that the systems for sound production and analysis must be closely related. In humans, the development of sound-processing systems for both language and music led to the development of specialized cortical regions, especially in the temporal lobes. In fact, a major difference between the human and the monkey cortex is a marked expansion of auditory areas in humans.

The Ear

The ear is a masterpiece of engineering that consists of three sections: the outer, middle, and inner ear, all illustrated in Figure 9-8. The outer ear consists of both the *pinna*, the funnel-like external structure made of cartilage and flesh, and the *external ear canal,* which extends a short distance from the pinna inside the head. The pinna is designed to catch the waves of air pressure in the surrounding environment and deflect them into the external ear canal. This canal amplifies the waves somewhat and directs them to the *eardrum* at its inner end. When sound waves strike the eardrum, it vibrates, the rate of vibration varying with the frequency of the waves.

On the inner side of the eardrum is the middle ear, an air-filled chamber that contains three small bones, or **ossicles,** connected to one another in a series. These three small bones (the smallest bones in the human body) are called the *hammer,* the *anvil,* and the *stirrup* because of their distinctive shapes. The ossicles attach the eardrum to the *oval window,* an opening in the bony casing of the **cochlea,** which is the inner-ear structure that contains the auditory receptors. When sound waves cause the eardrum to vibrate, those vibrations are transmitted to the ossicles. The ossicles, in turn, produce a leverlike action that conveys and amplifies the vibrations onto the membrane that covers the cochlea's oval window.

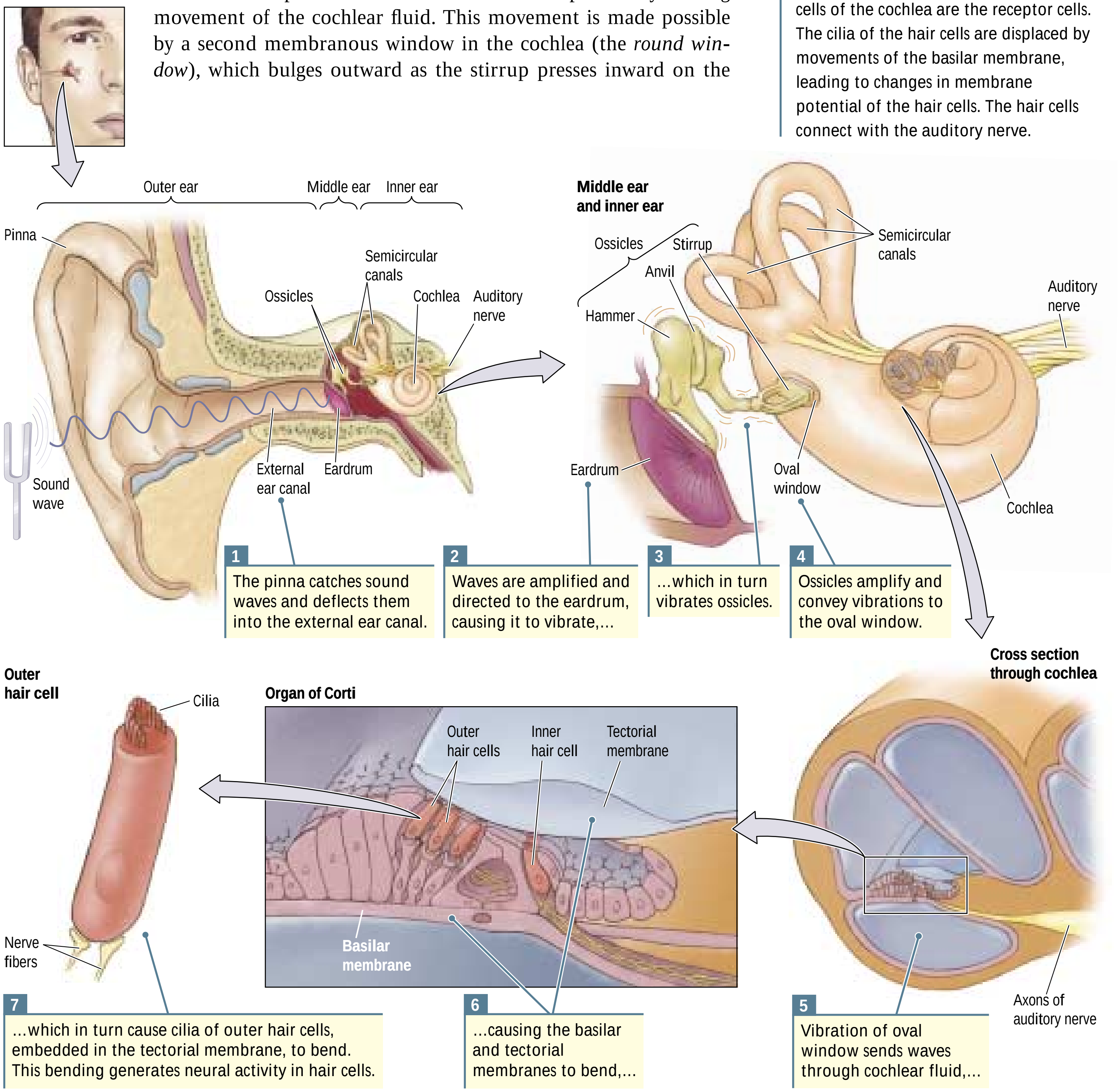
Basilar membrane. The receptor surface in the cochlea that transduces sound waves into neural activity.

As Figure 9-8 shows, the cochlea coils around itself and looks a bit like the shell of a snail. (The name *cochlea* derives from the Latin word for "snail.") Inside its bony exterior, the cochlea is hollow, as the cross-sectional drawing reveals. The compartments of this hollow interior are filled with fluid, and floating in the middle of the fluid is a thin membrane called the **basilar membrane.** Embedded in a part of the basilar membrane are **hair cells,** which are the auditory receptors. These hair cells and their supporting cells are collectively called the **organ of Corti.** The tips of the hair cells have little filaments, known as *cilia,* and the cilia of the outer hair cells are embedded in an overlying membrane, called the **tectorial membrane.**

The movement of the cilia converts sound waves into neural activity. Pressure from the stirrup on the oval window starts the process by causing movement of the cochlear fluid. This movement is made possible by a second membranous window in the cochlea (the *round window*), which bulges outward as the stirrup presses inward on the

Figure 9-8

The ear comprises three major anatomical parts: the outer, middle, and inner ears. Sound is gathered by the outer ear, transduced from air pressure into mechanical energy in the middle ear, and transduced into neural activity by the cochlea in the inner ear. The hair cells of the cochlea are the receptor cells. The cilia of the hair cells are displaced by movements of the basilar membrane, leading to changes in membrane potential of the hair cells. The hair cells connect with the auditory nerve.



1 The pinna catches sound waves and deflects them into the external ear canal.

2 Waves are amplified and directed to the eardrum, causing it to vibrate,...

3 ...which in turn vibrates ossicles.

4 Ossicles amplify and convey vibrations to the oval window.

5 Vibration of oval window sends waves through cochlear fluid,...

6 ...causing the basilar and tectorial membranes to bend,...

7 ...which in turn cause cilia of outer hair cells, embedded in the tectorial membrane, to bend. This bending generates neural activity in hair cells.

oval window. In a chain reaction, the movement of the cochlear fluid causes movements of the basilar and tectorial membranes, which in turn bend the cilia in one direction or another.

The structure and function of the basilar membrane are easier to visualize if the cochlea is uncoiled and laid flat, as in Figure 9-9. When the oval window vibrates in response to the vibrations of the ossicles, it generates waves that travel through the fluid of the cochlea. The waves cause the basilar and tectorial membranes to bend, and, when the membranes bend, the cilia of the hair cells are stimulated. This stimulation generates action potentials in the hair cells.

The key question is how this arrangement can code for the various properties of sound. In the late 1800s, German physiologist Hermann von Helmholtz proposed that sounds of different frequencies cause different parts of the basilar membrane to resonate. Von Helmholtz was not precisely correct. Actually, *all* parts of the basilar membrane bend in response to incoming waves of any frequency. The key is where on the basilar membrane the *peak* displacement takes place.

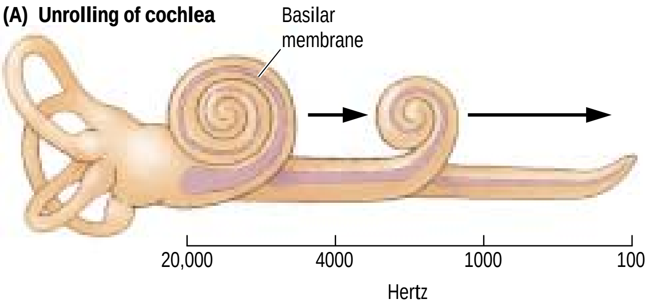
George von Békésy
(1899–1972)

This solution to the coding puzzle was not determined until 1960, when George von Békésy was able to observe the basilar membrane directly and see that a traveling wave moves along it all the way from the oval window to the membrane's apex. Békésy placed little grains along the basilar membrane and watched them jump in different places with different frequencies of incoming waves. Faster wave frequencies caused maximum peaks of displacement near the base of the basilar membrane, whereas slower wave frequencies caused maximum displacement peaks near the membrane's apex.

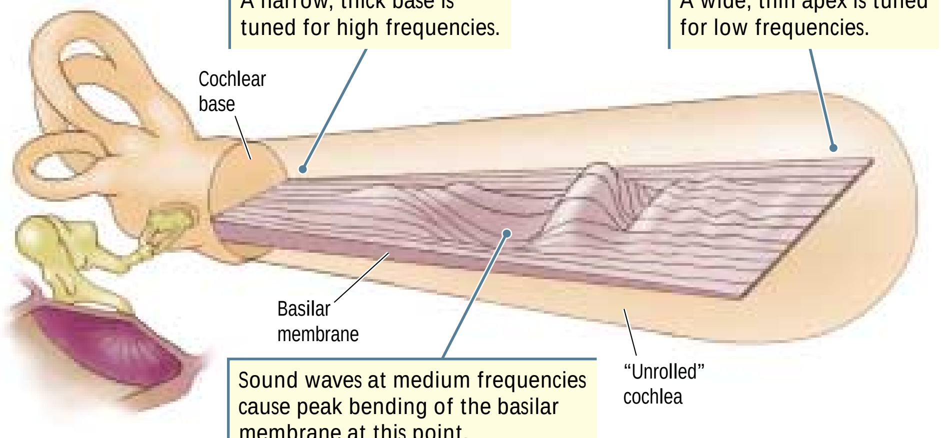
As a rough analogy, consider what happens when you shake a rope. If you shake it very quickly, the waves are very small and remain close to the part of the rope that you are holding. But, if you shake the rope slowly with a larger movement, the waves reach their peak farther along the rope. The key point is that, although both rapid and slow shakes of the rope produce movement along the rope's entire length, the maximum displacement of the rope is found at one end or the other, depending on whether the wave movements are rapid or slow.

This same response pattern is true of the basilar membrane. All sounds cause some displacement along the entire length of the basilar membrane, but the amount of displacement at any point varies with the frequency of the sound. In the human cochlea, the basilar membrane near the oval window

(A) Unrolling of cochlea



(B)



A narrow, thick base is tuned for high frequencies.

A wide, thin apex is tuned for low frequencies.

Sound waves at medium frequencies cause peak bending of the basilar membrane at this point.

Figure 9-9

(A) Anatomy of the cochlea, which is being unrolled. The numbers are the frequencies to which the basilar membrane is maximally responsive. **(B)** The basilar membrane has been unrolled. A traveling wave moves along the basilar membrane, producing its maximal displacement of the membrane as it approaches the apex.

is maximally affected by frequencies as high as about 20,000 hertz, whereas the most effective frequencies at the membrane's apex are less than 100 hertz. Intermediate frequencies maximally displace points on the basilar membrane between its two ends, as shown in Figure 9-9. When a wave of a certain frequency travels down the basilar membrane, hair cells at the point of peak displacement are stimulated, resulting in a maximal neural response in those cells. An incoming signal composed of many frequencies causes several different points along the basilar membrane to vibrate and excites hair cells at all these points.

Not surprisingly, the basilar membrane is much more sensitive to changes in frequency than our rope is. This greater sensitivity is achieved because the basilar membrane is not the same thickness along its entire length. Instead, it is narrow and thick near the oval window and wider and thinner at its opposite end. The combination of varying width and thickness enhances the effect of small differences in frequency on the basilar membrane. As a result, the cochlea can code small differences in frequency.

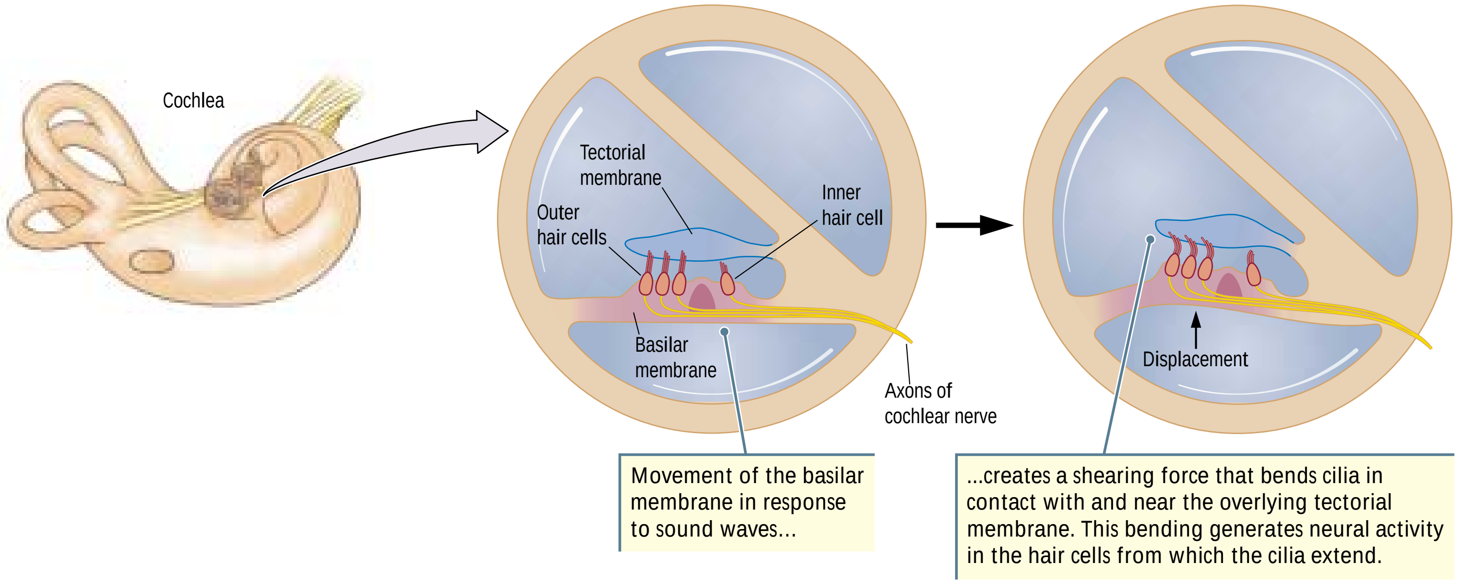
The Auditory Receptors

As you know, hair cells are what ultimately transform sound waves into neural activity. Figure 9-8 shows the structure of a hair cell, and Figure 9-10 illustrates how it functions to generate an action potential. There are two sets of hair cells: *inner hair cells* and *outer hair cells*. The human cochlea has 3500 inner hair cells and 12,000 outer hair cells, which is a small total number considering how many different sounds we can hear. As is also shown in Figure 9-10, the tips of the cilia of outer hair cells are attached to the overlying tectorial membrane, but the cilia of the inner hair cells do not touch the tectorial membrane. Nevertheless, the movement of the basilar and tectorial membranes causes the cochlear fluid to flow past the cilia of the inner hair cells, bending them back and forth. The inner hair cells are the auditory receptors. Animals with intact outer hair cells but no inner hair cells are effectively deaf. The outer hair cells function simply to sharpen the resolving power of the cochlea by contracting or relaxing and thereby changing the stiffness of the tectorial membrane.

One puzzle is how this function of the outer hair cells is controlled. How do these cells know when they need to contract or relax? The answer seems to be that the outer hair cells, through connections with axons in the auditory nerve, send some type of

Figure 9-10

Transduction of movement into neural activity is mediated by the hair cells. Movement of the basilar membrane creates a shearing force that bends the cilia, leading to the opening or closing of calcium channels in the hair cell. The influx of calcium ions leads to the release of transmitter by the hair cell, which stimulates an action potential in the cochlear (or auditory) nerve.



Movement of the basilar membrane in response to sound waves...

...creates a shearing force that bends cilia in contact with and near the overlying tectorial membrane. This bending generates neural activity in the hair cells from which the cilia extend.

message to the brainstem auditory areas. These areas then send a message back to the outer hairs, causing appropriate alterations of tension on the tectorial membrane. In this way, the brain helps the receptors create an auditory world.

A final question remains about the workings of hair cells: How does movement of their cilia generate neural activity? The answer is that movement of the cilia causes a change in polarization of the hair cell. Look at Figure 9-8 again and notice that the cilia of a hair cell differ in height. Movement of the cilia in the direction of the tallest cilia results in depolarization, which in turn opens calcium channels and leads to the release of transmitter onto the dendrites of the cells that form the auditory nerve. Movement in the direction of the shortest cilia results in hyperpolarization and a corresponding decrease in transmitter release. Hair cells are amazingly sensitive to the movement of their cilia. A movement sufficient to allow sound detection is only about 0.3 nanometer, or about the diameter of a large atom.

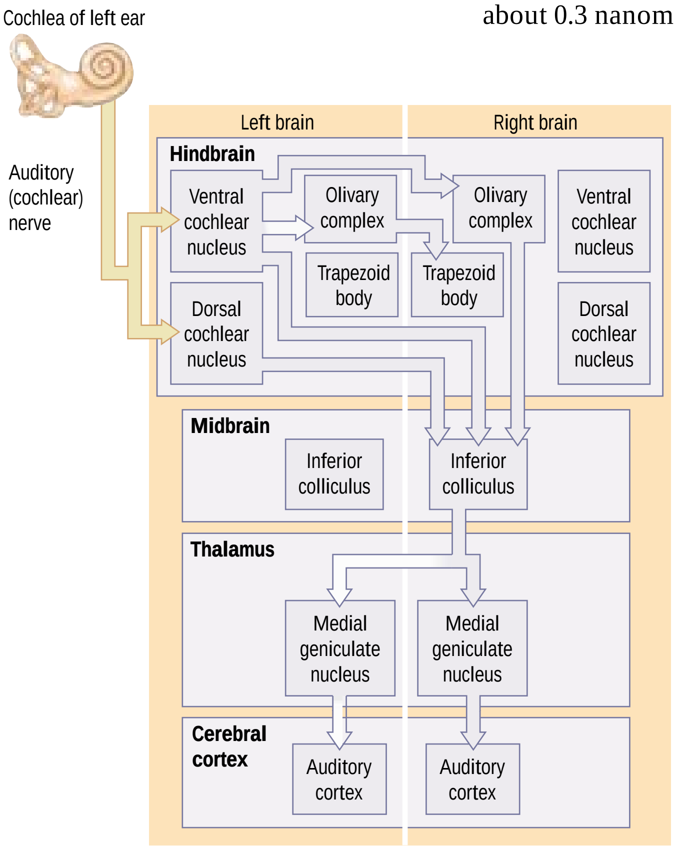


Figure 9-11

The primary route of auditory information from the cochlear nucleus through the hindbrain to the midbrain to the cortex. Three major points are evident: (1) the auditory inputs cross to the opposite side in the hindbrain; (2) there is recrossing of information, so that information from the ear reaches both hemispheres; and (3) there are multiple nuclei processing inputs en route to the cortex.

The Auditory Pathways

Hair cells in the organ of Corti synapse with neighboring bipolar cells, the axons of which form the **cochlear (auditory) nerve,** which in turn forms part of the eighth cranial nerve. Each bipolar cell receives information from only one inner hair cell. This arrangement contrasts with inputs to retinal ganglion cells, which you learned about in Chapter 8. Whereas ganglion cells in the eye receive inputs from many receptor cells, bipolar cells in the ear receive input from but a single receptor.

The cochlear-nerve axons enter the brainstem at the level of the medulla and synapse in the *cochlear nucleus*, which has two subdivisions (the ventral and dorsal). Two other nearby structures in the brainstem, the **superior olivary complex** (superior olive) and the **trapezoid body,** each receive connections from the cochlear nucleus, as shown in Figure 9-11. The projections from the cochlear nucleus connect with cells on the same side of the brain as well as with cells on the opposite side. This arrangement mixes the inputs from the two ears to form the perception of a single sound. The cochlear nucleus and the superior olive both send projections to the **inferior colliculus** at the top of the midbrain. Two distinct pathways emerge from the inferior colliculus, coursing to the **medial geniculate nucleus,** which lies in the thalamus. The ventral region of the medial geniculate nucleus projects to the primary auditory cortex (known as A1), whereas the dorsal region projects to the auditory cortical regions adjacent to A1.

The Auditory Cortex

In the human cortex, area A1 is within **Heschl's gyrus** and is surrounded by other secondary cortical areas, as shown in Figure 9-12. The secondary cortex lying behind Heschl's gyrus is called the **planum temporale** (meaning "temporal plane"). In right-handed people, the planum temporale is larger on the left side of the brain than it is on the right, whereas Heschl's gyrus is larger on the right side than on the left. The cortex of the left planum forms a speech zone, known as **Wernicke's area (the posterior speech zone),** whereas the cortex of the right Heschl's gyrus has a special role in

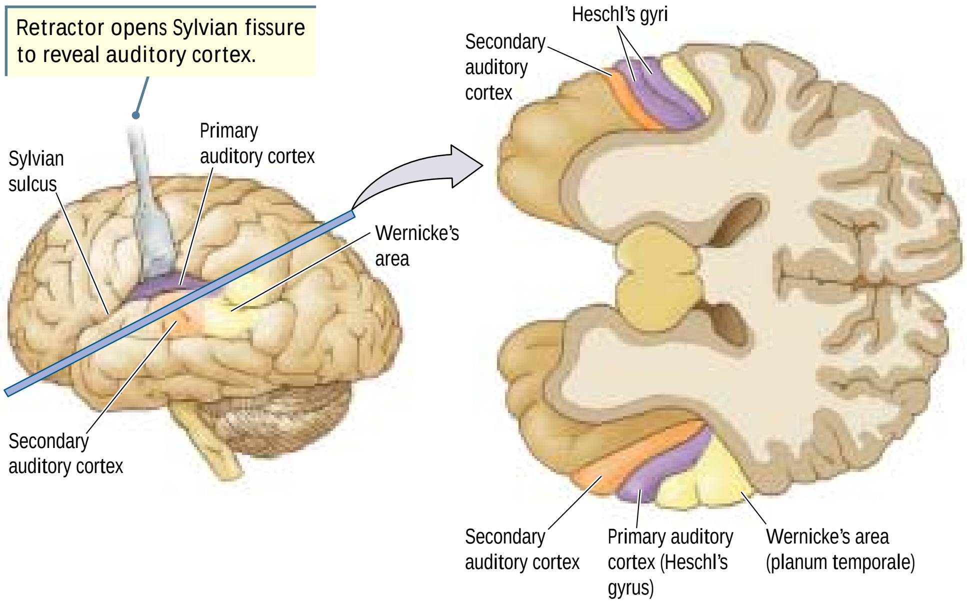


Figure 9-12

The human auditory cortex areas. (*Left*) Diagram of the left hemisphere of the brain showing the primary auditory cortex and the adjacent secondary regions. (*Right*) The posterior auditory cortex forms the posterior speech zone (Wernicke's area), which is larger on the left. Heschl's gyrus is larger on the right.

the analysis of music. These differences mean that the auditory cortex is anatomically and functionally asymmetrical, a property called *cerebral asymmetry*. Although cerebral asymmetry is not unique to the auditory system, it is most obvious here because the auditory analysis of language takes place only in the left hemisphere. Most left-handed people (about 70 percent) have the same anatomical asymmetries as those of right-handers, an indication that language organization is not related to hand preference. The remaining 30 percent of left-handers fall into two distinct groups: about half have an organization opposite that of right-handers; the other half have some type of idiosyncratic bilateral representation of language in which some language functions are in one hemisphere and others are in the other hemisphere.

The localization of language on the left side of the brain is often referred to as the lateralization of language. As a rule of thumb, if one hemisphere is specialized for one type of analysis, such as language, the other hemisphere has some type of complementary function, which in regard to audition appears to be music. We will return to the lateralization of language later in this chapter and again in Chapter 14.

In Chapter 8, you learned that there are two distinct visual pathways through the cortex: the temporal stream for object recognition and the dorsal stream for the visual control of movement. A similar distinction exists in the auditory cortex (Romanski et al., 1999). Just as we can identify objects by their sound characteristics, we can direct our movements by sound. The role of sound in guiding movement is less familiar to sight-dominated people than it is to people who are blind. Nevertheless, it exists in us all. Imagine waking up in the dark and reaching to pick up a ringing telephone or to turn off an alarm clock. Your hand will automatically form the appropriate shape needed to carry out these movements just on the basis of the sound that you have heard. That sound is guiding your movements much as a visual image can guide them. Although relatively little is known about the auditory pathways in the cortex, one pathway appears to continue through the temporal lobe, much like the ventral visual pathway, and plays a role in identifying auditory stimuli. A second auditory pathway apparently goes to the posterior parietal region, where it forms a type of dorsal pathway for the auditory control of movement.

◉ In the overview of the brain area in the Central Nervous System module of your CD, investigate cortical anatomy and the four lobes.

Heschl's gyrus. The primary auditory cortex found in the temporal lobes.

Wernicke's area. A region in the posterior part of the left temporal lobe that regulates the comprehension of language; sometimes referred to as the posterior speech zone.

Tonotopic representation. A representation of the auditory world in which sounds are located in a systematic fashion in a procession from lower to higher frequencies.

Tuning curve. A curve representing the maximum sensitivity of a neuron to a range of auditory frequencies. Each hair cell is maximally responsive to a particular frequency, but it also responds to nearby frequencies, although the sound must be louder for the cell to respond.

In Review

Changes in air pressure are converted into neural activity by hair cells in the inner-ear structure called the cochlea. Incoming sound waves vibrate the eardrum, which in turn vibrates the tiny bones of the middle ear, the innermost one of which presses on the cochlea's oval window and sets in motion the cochlear fluid. The motion of this fluid displaces the basilar membrane, causing cilia on the hair cells to bend and generate action potentials. The frequencies of incoming sound waves are largely coded by the places on the basilar membrane that are most displaced by the incoming sound waves. The axons of bipolar cells of the cochlea form the auditory (cochlear) nerve, which enters the brain at the medulla and synapses on cells in the cochlear nucleus. The neurons of the cochlear nucleus and associated regions in the medulla then begin a pathway on each side of the brain that courses to the midbrain (inferior colliculus), the thalamus (medial geniculate nucleus), and the cortex. In the brains of right-handed people, the auditory cortex on the left and right are asymmetrical, with the planum temporale being larger on the left and Heschl's gyrus being larger on the right. This anatomical asymmetry is correlated to a functional asymmetry: the left temporal cortex analyzes language-related sounds, whereas the right temporal cortex analyzes music-related ones. Most left-handed people have a similar arrangement, although about 30 percent of left-handers have a different pattern.

NEURAL ACTIVITY AND HEARING

We now turn to the ways in which the activities of neurons in the auditory system create our perception of sounds. Neurons at different levels in this system serve different functions. To get an idea of what the individual cells do, we will consider how the auditory system codes for pitch, loudness, location, and pattern.

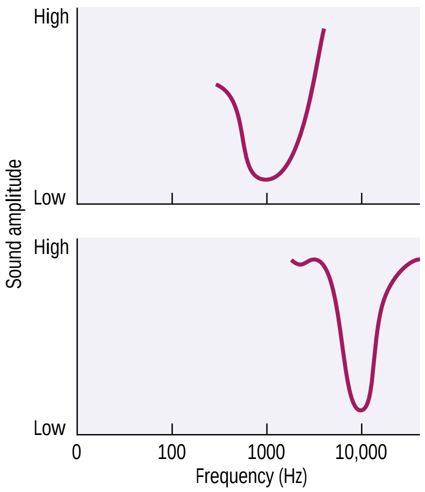


Figure 9-13

The frequency tuning curves of two different axons in the cochlear nerve. Each graph plots the frequency against the amplitude of the sound required to increase the firing rate of the neuron. The lowest point on each graph is the frequency to which that cell is most sensitive. The upper tuning curve is centered on 1000 hertz, whereas the lower tuning curve is centered on 10,000 hertz.

Hearing Pitch

Recall that our perception of pitch corresponds to the property of sound called frequency, which is measured in cycles per second, or hertz. Cells in the cochlea code frequency by their location on the basilar membrane. The cilia of hair cells at the base of the cochlea are maximally displaced by high-frequency sounds, and those at the apex are displaced the most by sounds of low frequency. This arrangement is referred to as a **tonotopic representation** (literally meaning "tone place"). Because axons of the bipolar cells that form the cochlear nerve are each connected to only one hair cell, they contain information about the place on the basilar membrane being stimulated.

If we record from single fibers in the cochlear nerve, we find that, although each axon transmits information about only a small part of the auditory spectrum, the cells do respond to a range of frequencies. In other words, each hair cell is maximally responsive to a particular frequency, but it also responds to nearby frequencies, even though the sound must be louder for these nearby frequencies to generate a response. This range of responses to different frequencies at different amplitudes can be plotted to form a **tuning curve**, like those in Figure 9-13. Such a curve is reminiscent of the curves in Figure 8-7, which show the responsiveness of cones in the retina to different wavelengths of light. Each cone is maximally sensitive to a particular wavelength of light, but it still responds somewhat to nearby wavelengths.

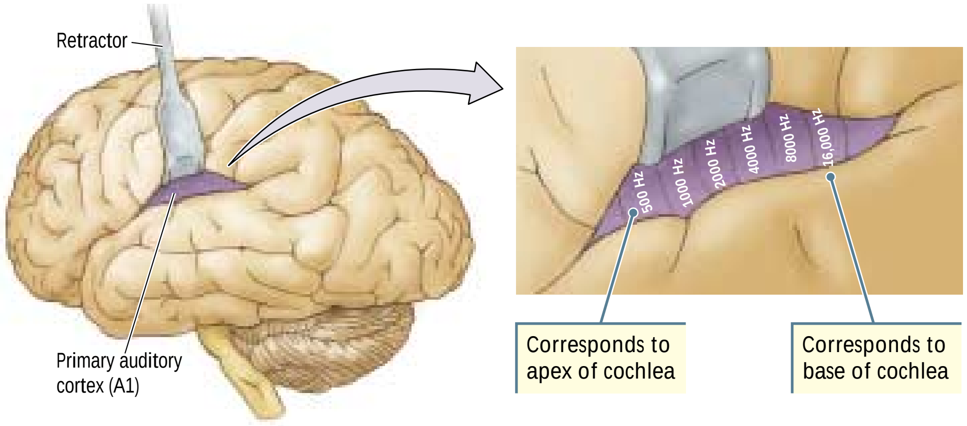


Figure 9-14

The tonotopic representation of the primary auditory cortex, or region A1. The anterior end of A1 represents the apex of the cochlea and hence low-frequency sounds, whereas the posterior end represents the base of the cochlea and hence high-frequency sounds. Because the auditory cortex is buried in the Sylvian fissure, a retractor is used to open the fissure to reveal the auditory cortex.

The axons of the bipolar cells in the cochlea project to the cochlear nucleus in an orderly manner. The axons entering from the base of the cochlea connect with one location, those entering from the middle connect to another location, and those entering from the apex connect to yet another place. As a result, the tonotopic representation of the basilar membrane is reproduced in the cochlear nucleus. This systematic representation is maintained throughout the auditory pathways and can be found in cortical region A1—the primary auditory cortex. Figure 9-14 shows the distribution of projections from the base and apex of the cochlea across A1. Similar tonotopic maps can be constructed for each level of the auditory system.

The systematic organization of tonotopic maps has enabled the development of **cochlear implants**—electronic devices surgically inserted in the inner ear to allow deaf people to hear (see Loeb, 1990). A miniature microphone-like processor detects the component frequencies of incoming sounds and sends them to the appropriate place on the basilar membrane through tiny wires. The nervous system does not distinguish between stimulation coming from this artificial device and stimulation coming through the ear. As long as appropriate signals go to the correct locations on the basilar membrane, the brain will "hear." Cochlear implants work very well, even allowing the deaf to detect the fluctuating pitches of speech. Their success provides corroborating evidence for the tonotopic representation of pitch in the basilar membrane.

Cochlear implants bypass the normal route for hearing by processing incoming sounds and sending them directly to the correct locations on the basilar membrane.

Michael Newman/PhotoEdit

One minor difficulty with the tonotopic theory of frequency detection is that the cochlea does not use this mechanism at the very apex of the basilar membrane, where hair cells, as well as the bipolar cells to which they are connected, respond to frequencies below about 200 hertz. At this location, *all* the cells respond to movement of the basilar membrane, but they do so in proportion to the frequency of the incoming wave. Higher rates of firing signal a relatively higher frequency, whereas lower rates of firing signal a lower frequency. Why the cochlea uses a different system to differentiate pitches within this range of very-low-frequency sounds is not clear. The reason probably has to do with the physical limitations of the basilar membrane. Although discriminating among low-frequency sounds is not important to humans, animals such as elephants and whales depend on these frequencies for their communication. Most likely they have more neurons at this end of the basilar membrane than do people.

Learn more about the development of the cochlear implant on the Web site at **www.worthpublishers.com/kolb/chapter9.**

Detecting Loudness

The simplest way for cochlear cells to indicate sound intensity is to fire at a higher rate when sounds are louder, which is exactly what happens. More-intense air-pressure changes produce more-intense vibrations of the basilar membrane and therefore

Cochlear implant. An electronic device that is implanted surgically into the inner ear to transduce sound waves into neural activity and allow deaf people to hear.

greater shearing of the cilia. This increased shearing leads to a greater amount of transmitter released onto bipolar cells. As a result, the bipolar axons fire more frequently, telling the auditory system that the sound is louder.

Detecting Location

The fact that each cochlear nerve synapses on both sides of the brain provides mechanisms for determining the source of a sound. One such mechanism is for brain cells to compute the difference in a sound's arrival time at the two ears. Figure 9-15 shows how a sound originating on the left reaches the left ear slightly before it reaches the right ear. Such differences in arrival time need not be large to be detected. If two sounds presented through earphones are separated by as little as 10 microseconds, the listener will perceive that a single sound came from the leading ear. This computation of left-ear–right-ear arrival times is carried out in the medial part of the superior olivary complex. Because these cells receive inputs from each ear, they are able to compare exactly when the signal from each ear reaches them.

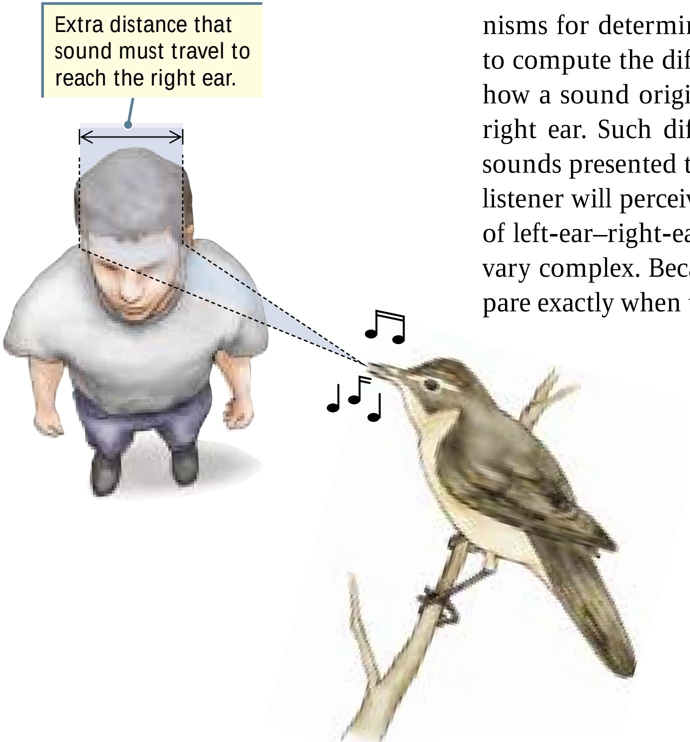


Figure 9-15

Sounds that originate on the left side of the body reach the left ear slightly before they reach the right ear, providing information about the location of the sound source. The difference in arrival time is subtle and, although the auditory system can use this information to localize sound, it is able to fuse the two sounds perceptually so that we hear but a single, clear sound.

As the source of a sound moves from the side of the head toward the middle, a person has greater and greater difficulty locating the source of the sound. The reason is that the difference in arrival time becomes smaller and smaller until there is no difference at all. When no difference exists, we can infer that the sound is either directly in front of us or directly behind us. To tell which it is, we move our heads, making the sound strike one of the ears sooner. We have a similar problem distinguishing between sounds directly above and below us. Again, we solve the problem by tilting our heads, thus causing the sound to strike one ear before the other.

Another way in which the auditory system can tell the location of the source of a sound has to do not with the difference in arrival times of a sound at the two ears, but instead with the sound's relative loudness on the left or the right. The basis of this mechanism is that higher-frequency sounds do not easily bend around the head, so the head acts as an obstacle to them. As a result, higher-frequency sounds on one side of the head are louder on that side than on the other. This difference is detected in the lateral part of the superior olivary complex and the trapezoid body. For sounds coming from directly in front or behind or from directly above or below, the same problem of differentiation exists, requiring the same solution of tilting or turning the head.

Although head tilting and turning are effective aids in localizing sounds, doing so takes time. The time needed is not usually important for humans, but it is important for certain other animals, such as owls, that hunt by using sound. Owls need to know the location of a sound simultaneously in at least two directions (left and below, for example, or right and above). The owl's solution is to have ears that are slightly displaced in the vertical direction so that sounds from above strike one ear sooner, whereas sounds from below strike the other ear sooner. This solution, shown in Figure 9-16, allows owls to hunt entirely by sound in the dark, which is bad news for mice.

Detecting Sound Patterns

Music and language are good examples of sound patterns that we can recognize. Because the right and left temporal lobes have roles in music and language, respectively, we can guess that neurons in the right and left temporal cortex take part in analyzing

Art Wolfe/Tony Stone

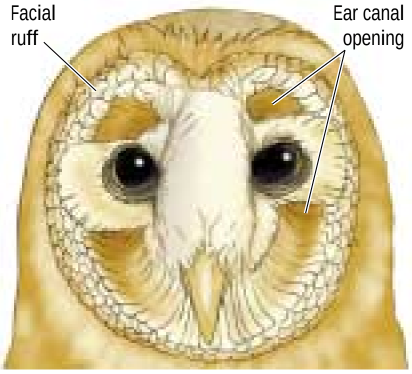


Figure 9-16

In this photograph of a barn owl catching a mouse in the dark, the owl has aligned its talons with the body axis of the mouse. The drawing above depicts the facial structure of the barn owl. The face is formed by rows of tightly packed feathers, called the facial ruff, that extend from the relatively narrow skull. The external ears are troughs formed by the part of the ruff that runs down the length of the face to join below the beak. The ruff collects sounds and funnels them into ear-canal openings. The owl's left ear is more sensitive to low-frequency sounds from the left, and the right ear is more sensitive to low-frequency sounds from the right. With sounds of high frequency, however, the right ear is more sensitive to those coming from above, because the ear canal is lower on the right and the trough is tilted up. The ear canal is higher on the left side and the trough is tilted down. Therefore, the left ear is more sensitive to sounds coming from below. For an owl, then, differences in perceived loudness can yield clues to the elevation of a source of sounds as well as to its horizontal direction.

Drawing adapted from "The Hearing of the Barn Owl," by E. I. Knudsen, 1981, *Scientific American, 245*(6), p. 115.

patterns in these two kinds of auditory experience. Studying the activities of these neurons in humans is not easy, however. Most of the knowledge that we have comes from studies of how individual neurons respond in nonhuman primates. For instance, Peter Winter and Hans Funkenstein (1971) found that neurons in the auditory cortex of the squirrel monkey are specifically responsive to squirrel monkey vocalizations. More recently, Joseph Rauschecker and his colleagues (1995) discovered that neurons in the secondary auditory areas of rhesus monkeys are more responsive to mixtures of sounds than to pure tones. Other researchers also have shown that the removal of the temporal auditory cortex abolishes the ability to discriminate the vocalizations made by other members of the species (Heffner and Heffner, 1990). Interestingly, discrimination of species-typical vocalizations in monkeys seems more severely disrupted by injury to the left temporal cortex than to the right. This finding implies that there may be a functional asymmetry for the analysis of complex auditory material in nonhuman primates, too.

In Review

Neurons in the cochlea form a tonotopic map that codes for a sound's frequencies. Such tonotopic maps of sound stimuli are maintained throughout the auditory system. The same cells in the cochlea that code for different frequencies of sound can also code for differences in sound amplitude by varying their firing rate, depending on a sound's loudness. Detecting the location of a sound is a function of neurons in the superior olive and trapezoid body of the brainstem. These neurons perform this role by computing differences in sound arrival time and loudness in the two ears. Understanding the meaning of sounds such as music and language requires pattern recognition, which is performed by cortical neurons.

THE ANATOMY OF LANGUAGE AND MUSIC

This chapter began with the story of the Neanderthal flute. The fact that Neanderthals made this instrument implies that they not only could process musical sounds but also could produce music. In the modern human brain, musical ability in general is largely a right-hemisphere specialization that forms a complementary system with language ability, which is largely localized in the left hemisphere. No one knows whether these two complementary systems evolved together in the brain, but it is certainly very possible that they did. If so, the Neanderthals, in addition to having the beginnings of music, must have had some form of language.

In the modern human brain, both language and music abilities are highly developed. Although little is known about how language and music are processed at the cellular level, electrical-stimulation and blood-flow studies have been sources of important insights into the regions of the cortex in which they are processed. We will investigate such studies next, starting with those that focus on how the brain processes language.

Processing Language

There are more than 4000 human languages in the world today, and probably many others have gone extinct in past millennia. Researchers have wondered whether the brain has a single system for understanding and producing *any* language, regardless of its structure, or whether very different languages, such as English and Japanese, are processed in different ways. To answer this question, it helps to analyze languages to determine just how similar they are, despite their various differences in vocabulary and grammar.

THE UNIFORMITY OF LANGUAGE

Foreign languages often seem impossibly complex to those who are unfamiliar with them. Their sounds alone may seem odd and difficult to make. If you are a native speaker of English, for instance, Asian languages, such as Japanese, probably sound peculiarly tonal to you, whereas European languages, such as German or Dutch, may sound heavily guttural. Even within related languages, such as English and French, there are marked differences in grammatical rules that can make the foreign language challenging to learn. Yet, as real as these linguistic differences may be, it turns out that they are superficial. The similarities in human languages, although not immediately apparent, are actually far more fundamental than their differences.

Noam Chomsky is usually credited with being the first linguist to stress the similarities over the differences in how human languages are structured. In a series of books and papers written in the past 40 years, he made a very sweeping claim, as have researchers such as Steven Pinker (1997) more recently. They argue that all languages have common structural characteristics because of a genetically determined constraint on the nature of human language. When this idea was first proposed in the 1960s, it was greeted with some skepticism. Since then, however, it has become clear that it is probably correct.

An obvious piece of evidence in favor of a genetic basis of human language is the fact that language is universal in human populations. All people everywhere use language. Furthermore, the complexity of language is not related to the complexity of a group's culture. The languages of technologically primitive peoples are every bit as complex and elegant as the languages of industrialized cultures. Old English is not inferior to modern English; it is just different. Humans, apparently, have a built-in capacity for creating and using language.

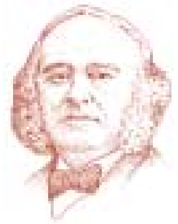
Another piece of evidence in favor of a genetic basis of human language is the fact that language is learned early in life and seemingly without effort. At about 12 months of age, children everywhere start to speak words. By 18 months, they are combining words, and, by age 3 years, they have a rich language capability. Perhaps the most amazing thing about language development is that children are not specifically taught the structure of their language. As toddlers, they are not painstakingly instructed in the rules of grammar. In fact, their early grammatical errors—for example, sentences such as "I goed to the zoo"—are seldom even corrected by adults. Yet children master language rapidly just the same. They also go through a series of stages of language acquisition that are remarkably similar across cultures. Indeed, the process of language acquisition plays an important role in Chomsky's theory of the innateness of language.

This is not to say that language development is not also influenced by experience. At the most basic level, children learn the language that they hear spoken. In an English household, they learn English; in a Japanese home, they learn Japanese. They also pick up the vocabulary of the people around them, which can vary from one speaker of a language to another. Furthermore, children go through a sensitive period of language acquisition, probably from about 1 to 6 years of age. If they are not exposed to language during this period, their language skills are severely compromised (see Chapter 7).

A third piece of evidence in favor of a genetic basis of language is the fact that all languages have many basic structural elements in common. Granted, every language has its own particular rules of grammar that specify exactly how the various parts of speech are to be positioned in a sentence, how words are to be inflected to convey different meanings, and so forth. But there are also overarching rules of grammar that apply to all human languages. For instance, all languages have the parts of speech that we call subjects, verbs, and direct objects. Consider the sentence "Jane ate the apple." "Jane" is the subject, "ate" is the verb, and "apple" is the direct object. The location of these three words in the sentence is not specified by any universal rule of grammar. Their positioning is a matter of the particular language spoken. In English, the order is subject, verb, object; in Japanese, the order is subject, object, verb; in Gaelic, the order is verb, subject, object. Nonetheless, all languages have these three classes of words.

The existence of basic structural elements in all human languages can be seen in the phenomenon of *creolization*—the development of a new language from what was formerly a very rudimentary pidgin language. This process took place in the seventeenth-century Americas, when slave traders and the owners of colonial plantations brought together people from various African villages who lacked a common language. Because the new slaves needed to communicate, they quickly created a simplified pidgin language that was based on whatever language the plantation owners spoke, be it English, French, Spanish, or Portuguese. The pidgin language had a crude syntax (word order), but it lacked a real grammar. The children of the slaves who spoke this pidgin language were brought up by caretakers who spoke only pidgin to them. Yet, surprisingly, these children did not learn the pidgin too. Rather, within a generation, they had created their own language, complete with a genuine grammar. Clearly, the pidgin was not a learnable language for children. The innate biology of their language-control systems shaped the development of a new language that was similar in basic structure to all other human languages. All creolized languages seem to evolve in a similar way, even though they are unrelated. This phenomenon can happen only if there is a significant innate component to language development.

THE LOCALIZATION OF LANGUAGE IN THE BRAIN



Paul Broca
(1824–1880)

The universality of language's basic structure set researchers on a search for a set of innate brain regions that underlie language use. By the late nineteenth century, it had become clear that language functions were at least partly localized—not just within the left hemisphere, but to specific areas there. Clues to this localization began in the early part of the nineteenth century, when neurologists observed patients with frontal-lobe injuries who suffered language difficulties. It was not until 1861, however, that Paul Broca examined a patient who had entirely lost his ability to speak except to say "tan" and to utter an oath. The man died shortly thereafter and Broca examined his brain, finding a fresh injury of the left frontal lobe. On the basis of this case and several subsequent cases, Broca concluded that language functions are localized in the left frontal lobe in a region just in front of the central fissure. This region, which is shown in Figure 9-17, soon became known as **Broca's area**, and the syndrome that results from damage to it became known as **Broca's aphasia**. A person with Broca's aphasia is unable to speak despite normal language comprehension and an intact vocal apparatus. The discovery of Broca's area was significant because it initiated the idea that the left and right hemispheres might have different functions.

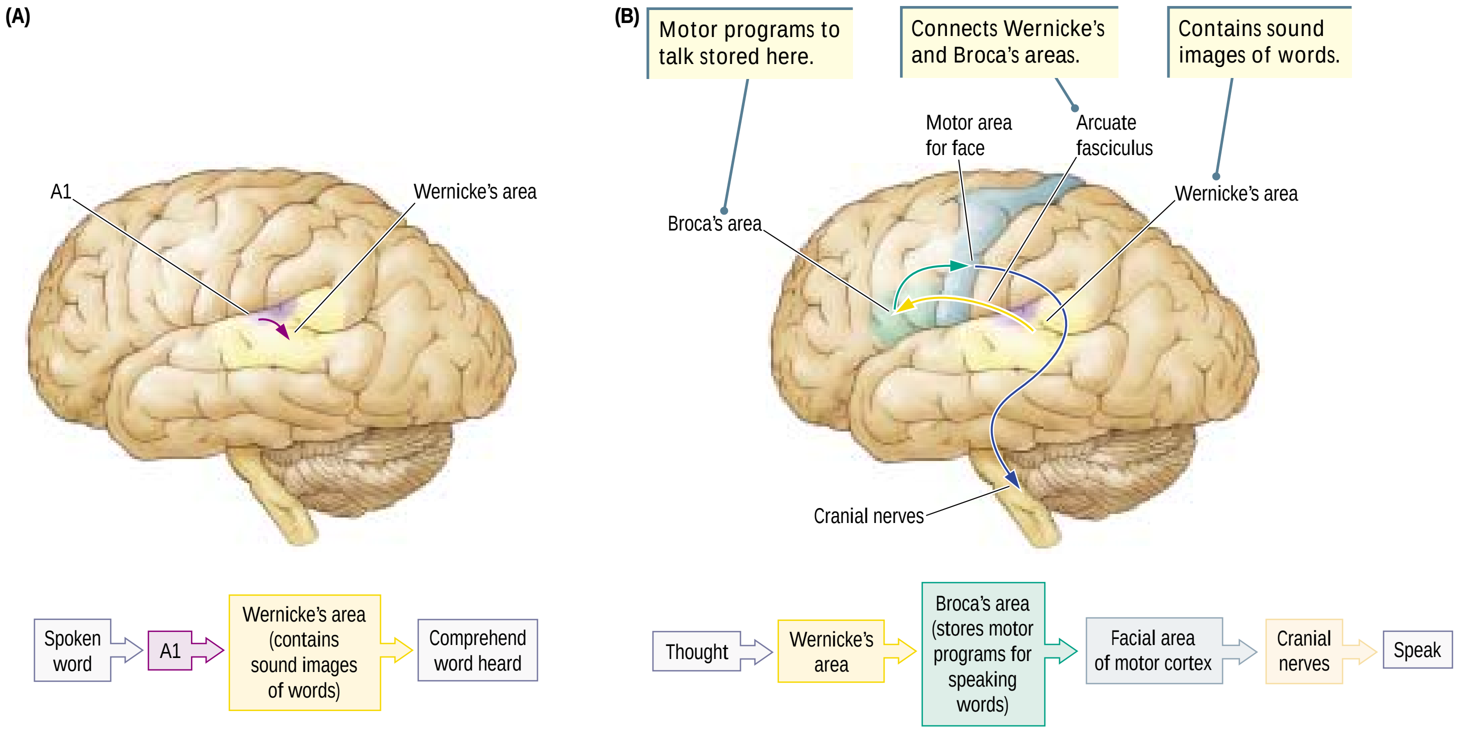
Other neurologists at the time believed that Broca's area might be only one left-hemisphere region that controls language. In particular, neurologists suspected a relation between hearing and speech. Proving this suspicion to be correct, Karl Wernicke later described patients who had difficulty comprehending language after injury to the posterior region of the left temporal lobe. This region subsequently became known as Wernicke's area (the speech zone referred to earlier; see Figure 9-12), and the syndrome associated with damage to it became known as **Wernicke's aphasia**. People who suffer Wernicke's aphasia can speak fluently, but their language is confused and makes little sense, as if they have no idea of what they are saying.

Wernicke went on to propose a model for how the two language areas of the left hemisphere interact to produce speech. He theorized that images of words are encoded by their sound and stored in Wernicke's area. When we hear a word that

● Look at the area of the brain involved in language on the three-dimensional brain model on your CD located in the brain overview section in the module on the Central Nervous System.

Figure 9-17

Wernicke's model of the neurology of language, showing the regions of the cortex involved. In this model, words are believed to be understood in Wernicke's area as summarized in **(A).** Words are produced through the connection that the arcuate fasciculus makes between Wernicke's area and Broca's area as summarized in **(B).**



matches one of those sound images, we recognize the word, which is how Wernicke's area contributes to speech comprehension. To *speak* words, Broca's area must come into play, because the motor program to produce each word is stored in this area. Messages are sent to Broca's area from Wernicke's area through a pathway known as the arcuate fasciculus, which connects the two regions. Broca's area in turn controls the articulation of the words by the vocal apparatus.

Wernicke's model provided a simple explanation for the existence of two major language areas in the brain and for the contribution of each area to the control of language. One difficulty, however, was that the model was based on postmortem examinations of patients with brain lesions that were often extensive. Not until the pioneering studies of Wilder Penfield, begun in the 1930s, were the language areas of the left hemisphere clearly and accurately mapped.

Broca's area. Just in front of the motor representation in the left hemisphere, the region that functions to produce the movements needed for language; sometimes referred to as the anterior speech area.

Broca's aphasia. The inability to speak fluently despite the presence of normal comprehension and intact vocal mechanisms.

Wernicke's aphasia. An inability to understand or to produce meaningful language even though the production of words is still intact.

AUDITORY AND SPEECH ZONES MAPPED BY BRAIN STIMULATION

Neurosurgeon Wilder Penfield took advantage of the chance to map auditory and language areas of the brain when operating on patients undergoing elective surgery to treat intractable epilepsy (see "Epilepsy" on page 144). The goal of this surgery is to remove abnormal tissues that cause the epileptic discharges. A major problem for the surgeon is to ensure that critical regions that serve important functions such as language are spared from injury. To determine the location of these regions, Penfield used a tiny electrical current to stimulate the surface of the brain. By monitoring the response of the patient to stimulation in different locations, Penfield could map brain functions.

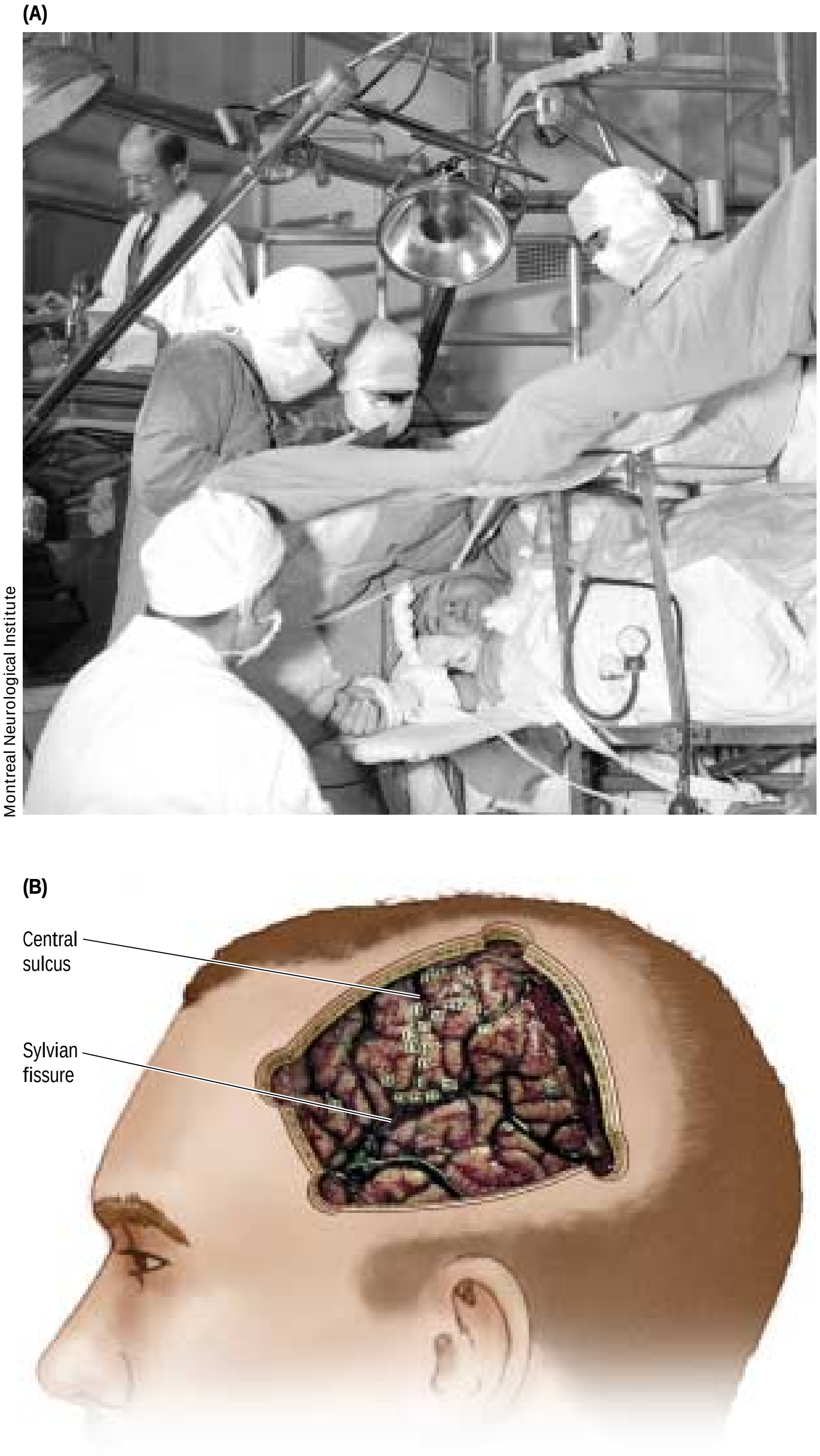
Figure 9-18 shows what happens in such surgery. Typically, two neurosurgeons perform the operation, and a neurologist analyzes the electroencephalogram in an adjacent room. Because patients are awake during the procedure, the effects of brain stimulation in specific regions can be determined in detail. Penfield placed little numbers on different parts of the brain's surface where the patient noted that stimulation had produced some noticeable effect.

When Penfield stimulated the auditory cortex, patients often reported hearing various sounds, such as a ringing sound like that of a doorbell, a buzzing noise, or a sound like that of birds chirping. This result is consistent with those of later studies of single-cell recordings from the auditory cortex in nonhuman primates. As mentioned earlier, findings in these later studies showed that the auditory cortex has a role in pattern recognition. Penfield also found that stimulation in A1 seemed to produce simple tones, ringing sounds, and so forth, whereas stimulation in the adjacent auditory cortex was more apt to cause some interpretation of a sound, such as ascribing it to a familiar source such as a cricket. There was no difference in the effects of stimulation of the left or right auditory cortex, and the patients heard no words when the brain was stimulated. Sometimes, however, stimulation of the auditory cortex produced effects other than the perception of sounds. Stimulation of one area, for example, might cause a patient to experience a sense of deafness, whereas stimulation of another area might produce a distortion of sounds actually being heard. As one patient exclaimed after a certain region had been stimulated, "All that you said was mixed up!"

Penfield was most interested in the effects of brain stimulation not on simple sound processing but on language. He mapped language areas in two ways. First, he stimulated different regions of the cortex while the patient was in the process of speaking. He expected that the electrical current might disrupt the ongoing speech by effectively "short-circuiting" the brain. In fact, this disruption did happen. The disruption of speech took several forms, such as slurred speech, confusion of words, or

Figure 9-18

(A) Neurosurgery performed on an awake epileptic patient. The patient is fully conscious, lying on his right side, with the left hemisphere of his brain exposed. He is kept comfortable with local anesthesia. In the background, the neurologist is studying the electroencephalogram being recorded from the patient's cortex, which will help in identifying the eleptogenic focus. The anesthetist is observing the effects of electrical stimulation of the cortex. **(B)** A drawing of the entire skull overlies a photograph of the patient's exposed brain at surgery. The numbered tickets identify the points that the surgeon (Wilder Penfield) stimulated. The application of a stimulating electrode at points 26, 27, and 28 produced interference with speech. Point 26 is presumably in Broca's area, 27 is in the face area, and 28 is in Wernicke's area in this patient.



Montreal Neurological Institute

difficulty in finding the right word. Such disruptions of speech are referred to as **aphasia**, which is a general term for any inability to comprehend or produce language. Electrical stimulation could also completely stop ongoing speech, a reaction that Penfield called *speech arrest*. Stimulation of regions well removed from the speech areas had no effect on ongoing speech, with the exception of regions of the motor cortex that control movements of the face. This exception makes sense because facial movement is required for talking.

The second way that Penfield mapped language areas in the brain was to stimulate the cortex when a patient was *not* speaking to see if he could cause the person to

Click on the Web site at **www.worthpublishers.com/kolb/chapter9** for current research on aphasia.

Aphasia. The inability to speak despite the presence of normal comprehension and intact vocal mechanisms.

Supplementary speech area. A region on the dorsal surface of the left frontal lobe that takes part in the production of speech.

utter some kind of speech sound. Penfield did not expect to trigger coherent speech, because the stimulation was not physiologically normal and so probably would not produce actual words or word combinations. This expectation was borne out. Stimulation of a region now known as the **supplementary speech area** produced a sustained vowel cry, such as “Oh” or “Eee.” Stimulation of the facial areas in the motor cortex and the somatosensory cortex also produced some vocalization because of their relation to movements of the mouth and tongue. Stimulation outside these speech-related zones produced no such effects. Figure 9-19 shows the areas of the left hemisphere that Penfield found were in some way engaged in processing language. Clearly, much of the left hemisphere takes part in this function.

In summary, Penfield and later researchers used electrical stimulation to identify four important cortical regions that control language. Two classic regions—Broca’s area and Wernicke’s area—are in the left hemisphere. Interestingly, although Broca’s area has traditionally been thought of as the site of speech production and Wernicke’s area as the site of language comprehension, electrical stimulation of either region disrupts both processes. Located on both sides of the brain are the other two major regions of language use: the supplementary speech area and the facial regions of the motor and somatosensory cortex. Although the effects on speech vary, depending on which of these four regions is stimulated, stimulation of any of them disrupts speech in some way. Not surprisingly, damage to each of these areas produces some form of aphasia, as described below in “Left-Hemisphere Dysfunction: The Story of Susan S.”

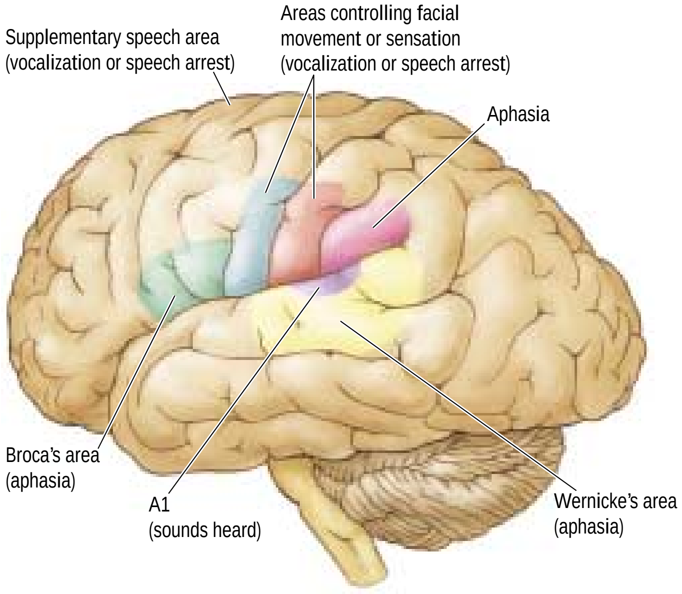


Figure 9-19

This map, based on Penfield and Rasmussen’s extensive study of patients who had surgery for the relief of intractable epilepsy, summarizes areas in which stimulation may interfere with speech or produce vocalization in the left hemisphere.

Adapted from *Speech and Brain Mechanisms* (p. 201), by W. Penfield and L. Roberts, 1956, London: Oxford University Press.

Focus on Disorders

Left-Hemisphere Dysfunction: The Story of Susan S.

Susan S. was a 25-year-old college graduate and mother of two who suffered from epilepsy. When she had a seizure, which was almost every day, she would lose consciousness for a short period, during which she would often engage in repetitive behaviors, such as rocking back and forth. Such psychomotor seizures can usually be controlled by medication, but the drugs were ineffective for Susan. The attacks were very disruptive to her life because they prevented her from driving a car and restricted the types of jobs that she could hold. So Susan decided to undergo neurosurgery to remove the region of abnormal brain tissue that was causing the seizures. This kind of surgery has a very high success rate. In her case, it entailed the removal of a part of the left temporal lobe, including most of the cortex in front of the auditory areas. Although this may seem to be a substantial amount of the brain to cut away, the excised tissue is usually abnormal, so any negative consequences are typically minor.

After her surgery, Susan did well for a few days, but then she suffered unexpected and unusual complications, which led to the death of the remainder of her left temporal lobe, including the auditory cortex and Wernicke’s area. As a result, she was no longer able to understand language, except for responding to the sound of her name, and she could say only one phrase: “I love you.” She was also unable to read, showing no sign that she could even recognize her own name in writing.

As we, the authors of this book, attempted to find ways to communicate with Susan, we tried humming nursery rhymes to her. She immediately recognized them and could say the words. We also discovered that she could sing. Although she was not ready for a concert performance, her singing skill was well within the normal range and she had a considerable repertoire of songs. She did not seem able to learn new songs, however, and she did not understand us if we “sang messages” to her. Apparently, Susan’s repertoire of songs and her singing ability were stored and controlled independently of her language system.

Positron emission tomography (PET). A technique whereby changes in blood flow can be detected by measuring changes in the uptake of compounds such as oxygen or glucose.

Go to the CD area on PET in the Research Methods module for a three-dimensional model of a PET camera and samples of PET scans.

THE AUDITORY CORTEX MAPPED BY POSITRON EMISSION TOMOGRAPHY

More recently, to analyze how the brain processes language, researchers have used a procedure known as **positron emission tomography (PET)** to study the metabolic activity of brain cells. Positron emission tomography is based on an idea that is surprisingly old. In the late 1800s, Angelo Mosso noticed pulsations in the living brain that kept pace with the heartbeat. Mosso was fascinated by this observation and believed that the pulsations were related to changes in blood flow in the brain. He later noticed that the pulsations appeared to be linked to mental activity. For example, when a subject was asked to perform a simple calculation, there was an immediate increase in brain pulsations and, presumably, in blood flow. But to demonstrate a relation between mental activity and blood flow within the brain, a more quantifiable measure than just visual observation would have to be used.

Various procedures for measuring blood flow in the brain were devised in the first 75 years of the twentieth century, one of which is described in "Arteriovenous Malformations." But not until the development of PET in the 1970s could blood flow in the brain of a human subject be measured safely and precisely (Posner and Raichle, 1997). This technique confirmed Mosso's observations.

A PET camera, like the one shown in Figure 9-20, is a doughnut-shaped array of radiation detectors that encircles a subject's head. A small amount of water, containing radioactive molecules to label it, is injected into the bloodstream. The person injected with these molecules is in no danger, because the molecules are very unstable

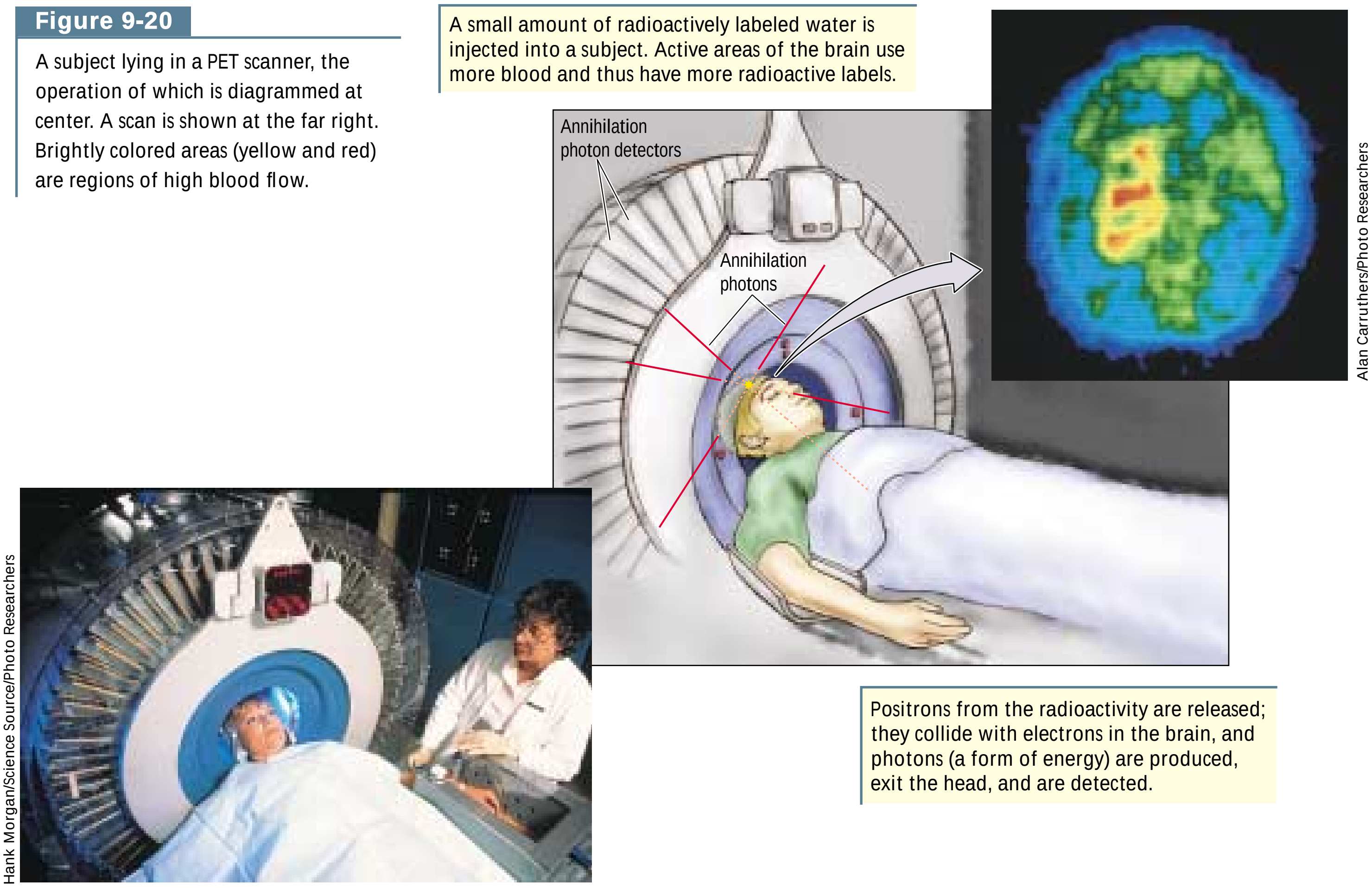


Figure 9-20

A subject lying in a PET scanner, the operation of which is diagrammed at center. A scan is shown at the far right. Brightly colored areas (yellow and red) are regions of high blood flow.

A small amount of radioactively labeled water is injected into a subject. Active areas of the brain use more blood and thus have more radioactive labels.

Positrons from the radioactivity are released; they collide with electrons in the brain, and photons (a form of energy) are produced, exit the head, and are detected.

Alan Carruthers/Photo Researchers

Hank Morgan/Science Source/Photo Researchers

Focus on Disorders

Arteriovenous Malformations

An arteriovenous malformation (also called an AV malformation or an angioma) consists of a mass of enlarged and tortuous cortical blood vessels that form congenitally. AV malformations are quite common, accounting for as many as 5 percent of all cases of cerebrovascular disease. Although these malformations may be benign, they often interfere with the functioning of the underlying brain and can produce epileptic seizures. The only treatment is to remove the malformation. This procedure carries significant risk, however, because the brain may be injured in the process.

Walter K. was diagnosed with an AV malformation when he was 26 years old. He had consulted a physician because of increasingly severe headaches, and a neurological examination had revealed an AV malformation over his occipital lobe. A surgeon attempted to remove the malformation, but the surgery did not go well; Walter was left with a defect in the bone overlying his visual cortex. This bone defect made it possible to listen to the blood flow through the malformation.

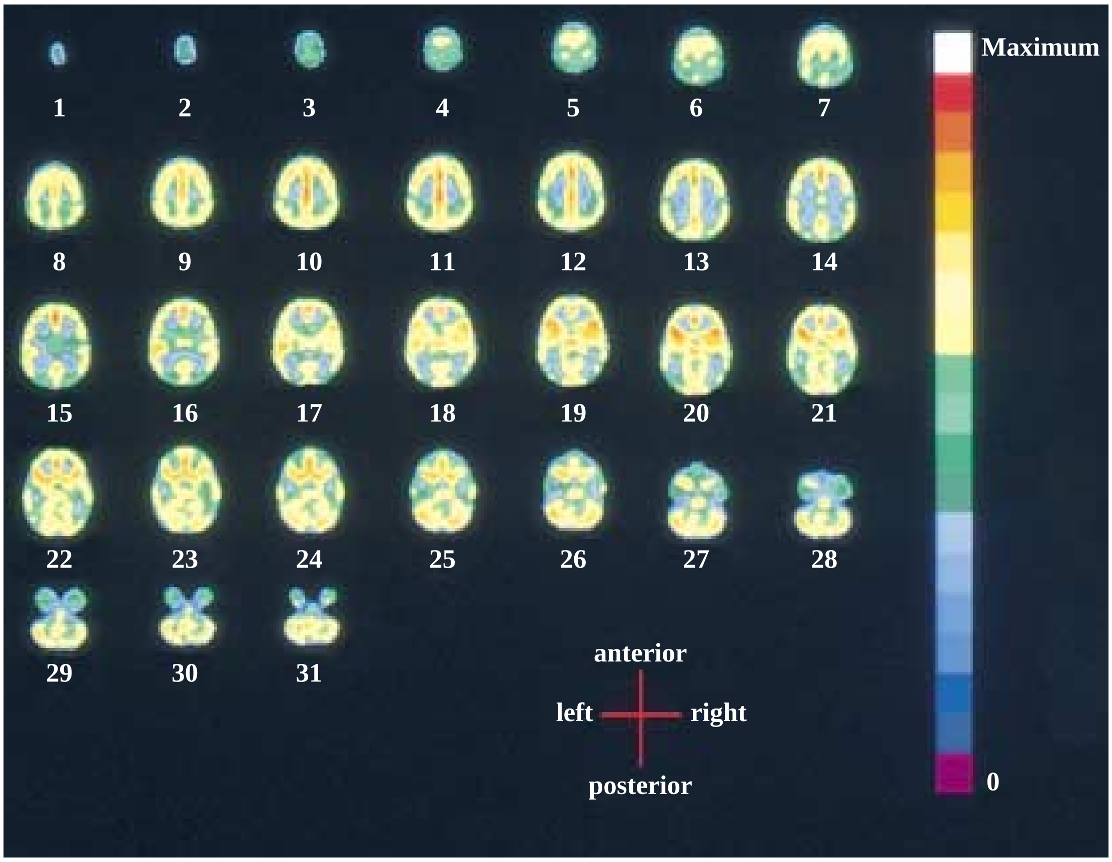
Dr. John Fulton noticed that when Walter suddenly began to use his eyes after being in the dark, there was a prompt increase in the noise (known as a bruit) associated with blood flow. Fulton documented his observations by recording the sounds of the bruit while Walter performed visual experiments. For example, if Walter had his eyes closed and then opened them to read a newspaper, there was a noticeable increase in blood flow through the occipital lobe. If the lights went out, the noise of the blood flow subsided. Merely shining light into Walter's eyes had no effect; nor was there an effect when he smelled vanilla or strained to listen to faint sounds. Apparently, the bruit and its associated blood flow were triggered by mental effort related to vision. To be able to reach this conclusion was remarkable, given that Fulton used only a stethoscope and a simple recording device for his study. Modern instrumentation, such as that of positron emission tomography, has shown that Fulton's conclusion was correct.

Simon Fraser/Royal Victoria Infirmary, Newcastle Upon Tyne/Science Photo Library/Photo Researchers

An MRI angiogram of an 18-year-old female with an AV malformation. The abnormal cerebral blood vessels (in white) formed a balloon-like structure (the blue area at lower right) that caused the death of the brain tissue around it.

and break down in just a few minutes. The radioactive molecules, such as the radioactive isotope oxygen-15 (^{15}O), release tiny positively charged particles known as positrons. These particles are emitted from an atom that is unstable because it is deficient in neutrons. The positrons are attracted to the negative charge of electrons in the brain, and the subsequent collision of these two particles leads to both of them being annihilated, thus creating energy. This energy, which is in the form of photons, leaves the head at the speed of light and is detected by the PET camera. The photons exit the head in exactly opposite directions at the same speed, so it is possible to identify where the source of the photons is.

How does this system enable the measurement of blood flow in the brain? The answer is that the unstable radioactive molecules accumulate in the brain in direct proportion to the rate of local blood flow. Local blood flow, in turn, reflects neural activity because potassium ions released from stimulated neurons dilate adjacent blood



M. E. Raichle, Mallinckrodt Institute of Radiology, Washington University School of Medicine

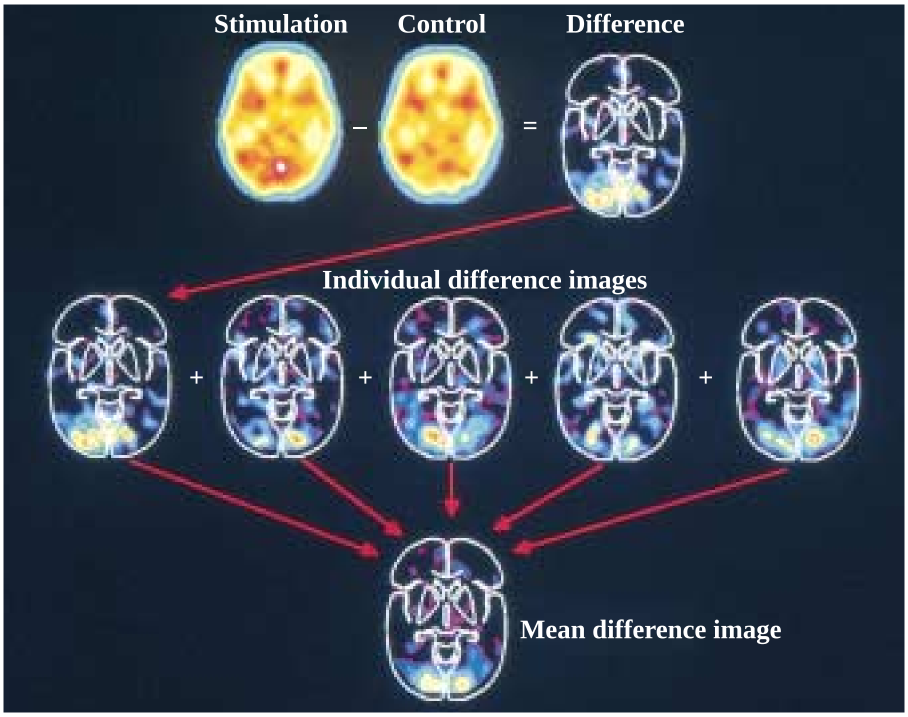
Figure 9-21

The PET images of blood flow obtained from a single subject while that subject was resting quietly with eyes closed. Each scan represents a horizontal plane, or section, from the top (1) to the bottom (31) of the brain.

vessels. The greater the blood flow, the higher the radiation counts recorded by the PET camera. With the use of sophisticated computer imaging, it is possible to create a map of the brain, like the one in Figure 9-21, which shows where the blood flow is highest.

Even though the distribution of blood is not uniform in Figure 9-21, it is still difficult to conclude very much from such a map. So PET researchers who are studying the link between blood flow and mental activity resort to a statistical trick. They take the pattern of blood flow when the subject is engaged in the experimental task and subtract from it the blood-flow pattern when the brain is in a carefully selected control state (such as when the person is lying quietly with eyes closed). As illustrated in Figure 9-22, this subtraction provides an image of the change in blood flow in the two states. The change can be averaged across subjects to yield an average image difference, revealing which areas of the brain are selectively active.

What happens when PET is used while subjects listen to sounds? Although there are many PET studies of auditory stimulation, a series conducted by Robert Zatorre and his colleagues (1992, 1995) serves as a good example. These researchers hypothesized that simple auditory stimulation, such as bursts of noise, would be analyzed by area A1, whereas more complex auditory stimulation, such as speech syllables, would be analyzed in adjacent secondary auditory areas. The researchers also hypothesized that performance of a speech-sound-discrimination task would selectively activate left-hemisphere regions. This is exactly what they found. Figure 9-23 shows that the primary auditory cortex experienced increased activity in response to noise bursts, whereas secondary auditory areas were activated by speech syllables. Both types of stimuli produced responses in both hemispheres, but there was greater activation in



M. E. Raichle, Mallinckrodt Institute of Radiology, Washington University School of Medicine

Figure 9-22

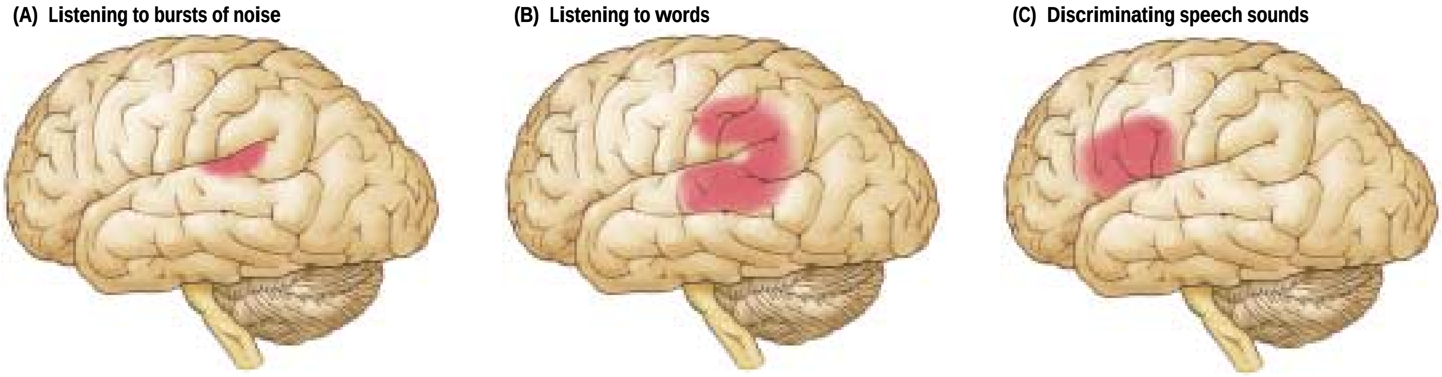
The procedure of subtraction. In the upper row of scans, the control condition (in this case, resting while looking at a static fixation point) is subtracted from the experimental condition of looking at a flickering checkerboard. The subtraction produces a somewhat different image for each of the five subjects shown in the middle row, but all show increased blood flow in the occipital region. The images are averaged to produce the image at the bottom.

the left hemisphere for the speech syllables. These results imply that auditory area A1 analyzes all incoming auditory signals, speech and nonspeech, whereas the secondary auditory areas are responsible for some higher-order signal processing required for the analysis of language sounds.

As Figure 9-23C shows, the speech-sound-discrimination task yielded an intriguing additional result: during this task, there was also activation of Broca's area in the left hemisphere. The involvement of this region in the frontal lobe during auditory analysis may seem surprising. In Wernicke's model, Broca's area is considered the place where the motor programs needed to *produce* words are stored. It is not normally a region thought of as the site of speech-sound discrimination. A possible explanation is that, to determine that the "g" in "bag" and "pig" is the same speech sound, the auditory stimulus must be related to how that sound is actually articulated. That is, the speech-sound perception requires a match with the motor behaviors associated with making that sound. This role for Broca's area in speech analysis is seen in further studies in which people are asked to determine if a stimulus is a word

Figure 9-23

Selective cortical areas are activated in different language-related tasks. **(A)** Passively listening to noise bursts activates the primary auditory cortex. **(B)** Listening to words activates the posterior speech area. **(C)** Making a phonetic discrimination activates the frontal region, including Broca's area.


(A) Listening to bursts of noise **(B) Listening to words** **(C) Discriminating speech sounds**

or a nonword (for example, "tid" versus "tin" or "gan" versus "tan"). In this case, information about how the words are articulated is irrelevant, and Broca's area would not need to be recruited. It is not.

Processing Music

Although Penfield did not study the effect of brain stimulation on musical analysis, there are many studies of musical processing in brain-damaged patients. "Cerebral Aneurysms: The Story of C. N." describes one of them. Collectively, the results of these studies confirm that musical processing is in fact largely a right-hemisphere specialization, just as language processing is largely a left-hemisphere one.

Maurice Ravel (1875–1937)

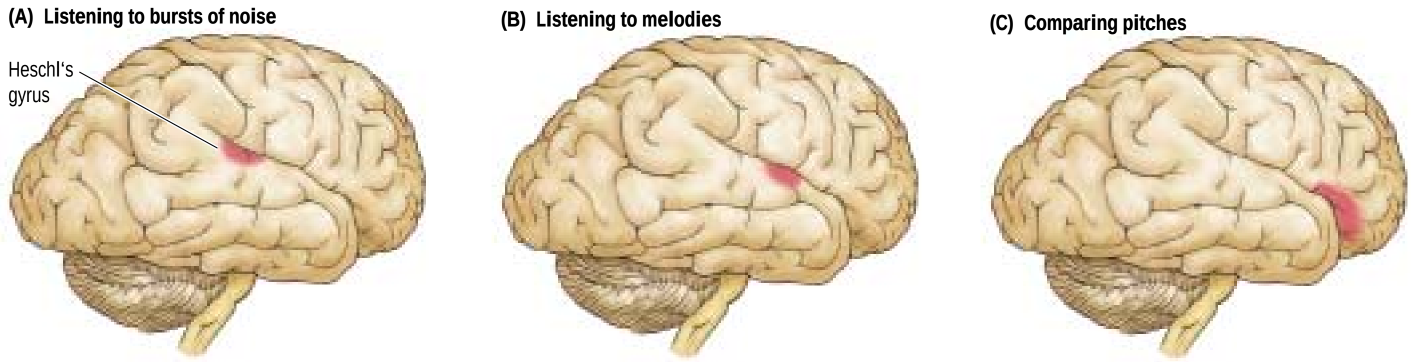
An excellent example of right-hemisphere predominance for the processing of music is seen in a famous patient—composer Maurice Ravel (1875–1937). Ravel suffered a left-hemisphere stroke and developed aphasia while at the peak of his career. Yet many of Ravel's musical skills remained intact after the stroke because they were localized to the right hemisphere. For instance, he could still recognize melodies, pick up tiny mistakes in music that he heard being played, and even judge the tuning of pianos. Interestingly, however, not all his musical skills were preserved. Skills that had to do with music production were among those that were destroyed. For instance, Ravel could no longer recognize written music, play the piano, or compose. This dissociation of music perception and music production is curious. Apparently, the left hemisphere plays at least some role in certain aspects of music processing, especially those that have to do with the "making" of music.

To find out more about how the brain carries out the perceptual side of music processing, Zatorre and his colleagues (1994) conducted PET studies. When subjects listened simply to bursts of noise, Heschl's gyrus became activated, but this was not the case when the subjects listened to melodies. As shown in Figure 9-24, the perception of melodies triggered major activation in the right-hemisphere auditory cortex lying in front of Heschl's gyrus, as well as minor activation in the same region of the left hemisphere. In another test, subjects listened to the same melodies but this time were asked to indicate whether the pitch of the second note was higher or lower than that of the first note. During this task, which requires short-term memory of what has just been heard, blood flow in the right frontal lobe increased. As with language, then, the frontal lobe plays a role in auditory analysis when short-term memory is used.

As noted earlier, the capacity for language appears to be innate. Sandra Trehab and her colleagues (1999) showed that this may be true for music as well. For example, infants show learning preferences for musical scales versus random notes. Furthermore, like adults, children are very sensitive to musical errors, presumably because they are biased for perceiving regularity in rhythms. Thus, it appears that, at birth, the brain is prepared for both music and language and, presumably, selectively attends to these types of auditory signals.

Figure 9-24

Selective cortical areas activated in different music-related tasks. **(A)** Passively listening to noise bursts activates Heschl's gyrus. **(B)** Listening to melodies activates the primary auditory cortex. **(C)** Making relative pitch judgments about two notes of each melody activates a right frontal lobe area.



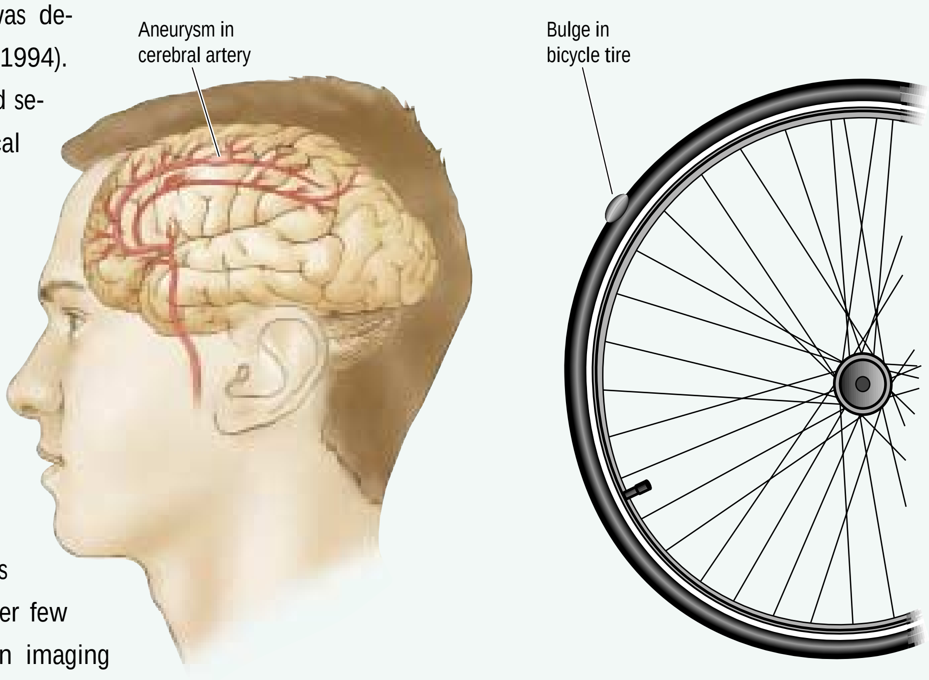
(A) Listening to bursts of noise

(B) Listening to melodies

(C) Comparing pitches

Focus on Disorders

Cerebral Aneurysms: The Story of C. N.



C. N. is a 35-year-old nurse whose case was described by Isabelle Peretz and her colleagues (1994). In December 1986, C. N. suddenly developed severe neck pain and headache. A neurological examination revealed an aneurysm in the middle cerebral artery on the right side of her brain. An *aneurysm* is a bulge in a blood vessel wall caused by a weakening of the tissue, much like the bulge that appears in a bicycle tube at a spot that is no longer strong. Aneurysms in a cerebral artery are dangerous because, if they burst, severe bleeding and subsequent brain damage result.

In February 1987, C. N.'s aneurysm was surgically repaired, and she appeared to suffer few adverse effects. However, postoperative brain imaging revealed that a new aneurysm had formed in the same location but on the opposite side of the brain. This second aneurysm was repaired 2 weeks later. After her surgery, C. N. had temporary difficulty finding the right word when she spoke, but, more important, her perception of music was deranged. She could no longer sing, nor could she recognize familiar tunes. In fact, singers sounded to her as if they were talking instead of singing. But C. N. could still dance to music. Because her music-related symptoms did not go away, she was given a brain scan. It revealed damage along the Sylvian fissure in both temporal lobes. The damage did not include the primary auditory cortex, nor did it include any part of the posterior speech zone. For these reasons, C. N. could still recognize nonmusical sounds and showed no evidence of language disturbance. Apparently, nonmusical sounds and speech sounds are analyzed in separate parts of the brain from music.

In Review

The auditory system has a complementary specialization in the cortex: left for language-related analyses and right for music-related ones. This asymmetry, however, appears to be relative, because there is good evidence that the left hemisphere plays a role in some aspects of music-related behaviors and that the right hemisphere has some language capabilities. The results of both electrical-stimulation and PET studies have shown that the left hemisphere contains several language-related areas. For instance, Wernicke's area identifies speech syllables and words, representations of which are stored in that location. Broca's area matches speech sounds to the motor programs necessary to articulate them, and, in this way, it plays a role in discriminating closely related speech sounds. The auditory cortex of the right hemisphere plays a major part in the comprehension of music.

AUDITORY COMMUNICATION IN NONHUMAN SPECIES

Many animals use sounds to aid their survival. Some, as we do, use sounds to communicate with other members of their species. Here we consider just two types of auditory communication in nonhumans: birdsong and bat echolocation. Each of these types provides a model for understanding different aspects of brain–behavior relations involving the auditory system.

Birdsong

There are about 8500 living species of birds, of which about half are considered songbirds. Birdsong has many functions, including attracting mates (usually by males), demarcating territories, and announcing location or even mere presence. Although birds of the same species all have a similar song, the details of that song vary markedly from region to region, much as dialects of the same human language vary. Figure 9-25 includes sound spectrograms for the songs of male white-crowned sparrows that live in three different localities near San Francisco. Notice how the songs of birds are nearly identical in a single region but quite different from region to region. These regional differences are due to the fact that song development in young birds is influenced not just by genes but also by early experience and learning. In fact, young birds can acquire more elaborate songs than can other members of their species if the young birds have a good tutor (Marler, 1991).

There are some broad similarities between birdsong and human language. For instance, if a young bird is not exposed to any songs until it is a juvenile and then listens to recordings of the songs of different species, there is a general preference for the song of the bird's own species. This preference must mean that there is some kind of species-specific song template in the brain of each species and that the details of this template are modified by experience. In a similar way, humans seem to have a basic structural template for language that is programmed into the brain, and a variety of specific grammatical forms are added to this template.

Another broad similarity between birdsong and human language is the great diversity of each. Among birds, this diversity can be seen in the sheer number of songs that a species possesses. Species such as the white-crowned sparrow have but a single song, whereas others such as the marsh wren have as many as 150 songs. The number of syllables in birdsong also varies greatly, ranging from 30 for the canary to about 2000 for the brown thrasher. In a similar way, even though all modern human languages are equally complex grammatically, they vary significantly in the type and number of elements that they employ. For instance, the number of meaningful speech sounds in human languages ranges from about 15 (for some Polynesian languages) to about 100 (for some dialects spoken in the Caucasus Mountains).

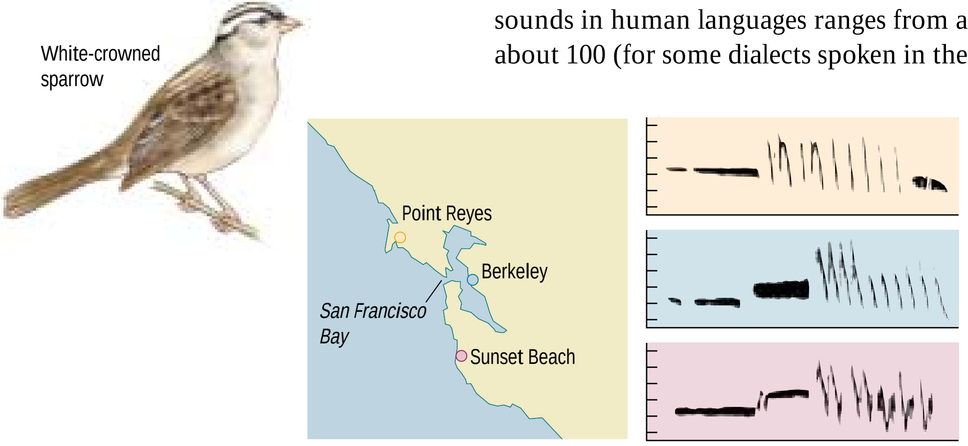


Figure 9-25

Sound spectrograms of three male white-crowned sparrows whose songs were recorded in three separate localities (Point Reyes in Marin County, Berkeley, and Sunset Beach) around San Francisco Bay. The sound spectrograms of males found in any one place are very similar but differ from those of males in other regions. Thus, birds raised in different regions have different dialects.

Adapted from P. Marler, The Instinct to learn. In S. Carey & R. German (Eds.). *The Epigenesis of mind: Essays on biology and cognition.* Hillsdale, NJ: Lawrence Erlbaum, 1991, p. 39.

A final broad similarity between birdsong and human language lies in how they develop. In many bird species, song development is heavily influenced by experience during a so-called sensitive period, just as it is in humans, as you learned in Chapter 7. Birds also go through stages in song development, just as humans go through stages in language development. Early in life, birds make noises that attract the attention of their parents, usually for feeding, and human babies, too, emit cries, often to signal hunger. When a bird is older and often ready to leave the nest, it begins to make noises that Charles Darwin compared to the prespeech babbling of human infants. These noises, called *subsong,* are variable in structure and low in volume, and they are often produced as the bird appears to doze. Presumably, subsong, like human babbling, is a type of practice for the later development of adult communication. As a young bird matures, it starts to produce sound patterns that contain recognizable bits of the adult song. Finally, the adult song emerges. In most species, the adult song remains remarkably stable, although a few species, such as canaries, can develop a new song every year.

The neural control of birdsong has been a topic of intense research, partly because it provides an excellent model of changes in the brain that accompany learning and partly because it can be a source of insight into how sex hormones influence behavior. Fernando Nottebohm and his colleagues first identified the major structures controlling birdsong in the late 1970s. These structures are illustrated in Figure 9-26. The largest are the *higher vocal control center* (HVC) and the *nucleus robustus archistriatalis* (RA). The axons of the HVC connect to the RA, which in turn sends axons to the 12th cranial nerve. This nerve controls the muscles of the syrinx, the structure that actually produces the song.

The HVC and RA have several important characteristics. First, they are asymmetrical in some species, with the structures in the left hemisphere being larger than those in the right hemisphere. In many cases, this asymmetry is similar to the lateralized control of language in humans: if the left-hemisphere pathways are damaged, the birds stop singing, but similar injury in the right hemisphere has no effect on song. Second, birdsong structures are sexually dimorphic. That is, they are much larger in males than in females. In canaries, they are five times as large in the male bird. This sexual difference is due to the hormone testosterone in males. Injection of testosterone into female birds causes the song-controlling nuclei to increase in size. Third, the size of the birdsong nuclei is related to singing skill. For instance, unusually talented singers among male canaries tend to have larger HVCs and RAs than do less-

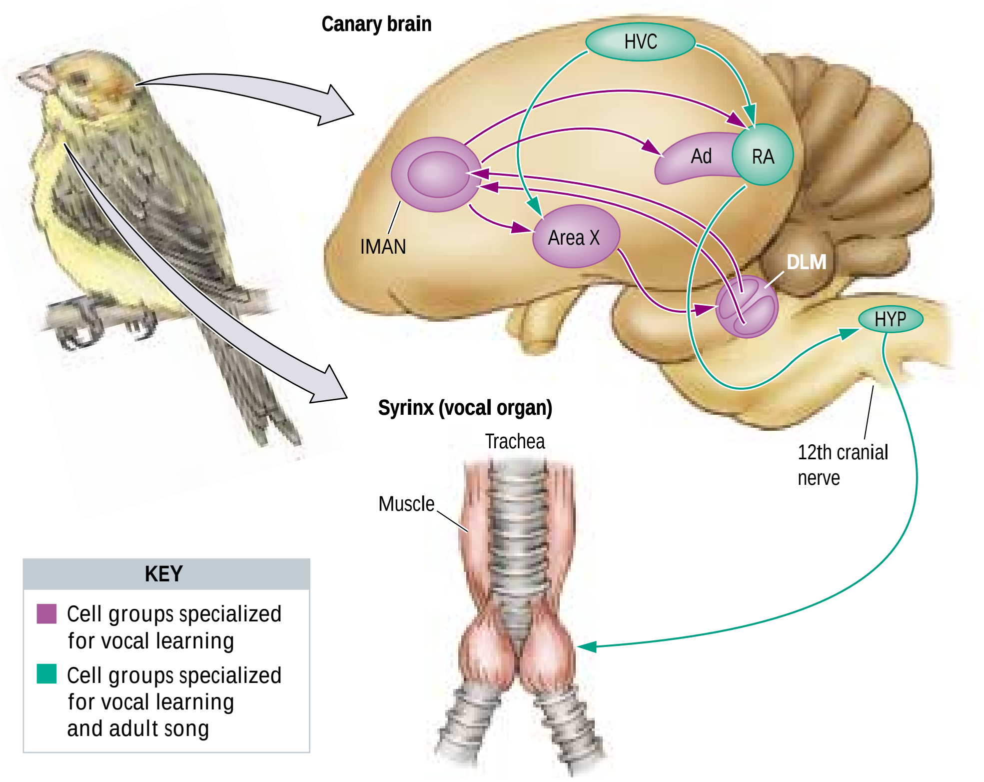


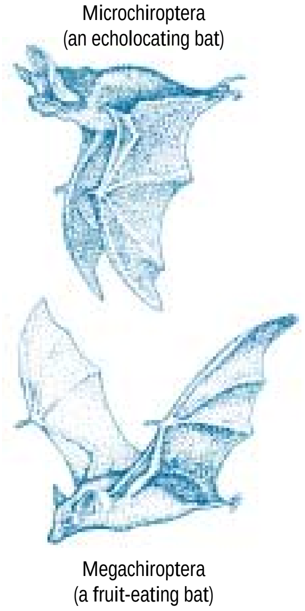
Figure 9-26

This side view of the canary brain shows several nuclei that control song learning, the two critical ones being the HVC (higher vocal control center) and the nucleus robustus archistriatalis (RA). These areas are necessary both for adult singing and for learning the song. Other regions necessary for learning the song during development but not required for the adult song include the dorsal archistriatum (Ad), the lateral magnocellular nucleus of the anterior neostriatum (IMAN), area X of the avian striatum, and the medial dorsolateral nucleus of the thalamus (DLM).

Echolocation. The ability to identify and locate an object by sound that bounces off the object.

gifted singers. Finally, the HVC and RA contain not only cells that produce birdsong but also cells responsive to hearing song, especially the song of a bird's own species. The same structures therefore play a role in both song production and song perception—comparable to the overlapping roles of Broca's and Wernicke's areas in language perception and production.

Echolocation in Bats



Next to rodents, bats are the most numerous order of mammals. There are two general groups of bats, known as suborders. One suborder consists of the smaller echolocating bats (Microchiroptera); the other comprises the larger fruit-eating and flower-visiting bats (Megachiroptera), sometimes called flying foxes. The echolocating bats interest us here because they use sound as a means of navigation.

There are 680 species of echolocating bats, most of which feed on insects. Some of the others live on blood (vampire bats), and some catch frogs, lizards, fishes, birds, and small mammals. These bats use echolocation for two purposes: for flying in the dark and for detecting and pursuing prey. The auditory system of bats is specialized for both tasks.

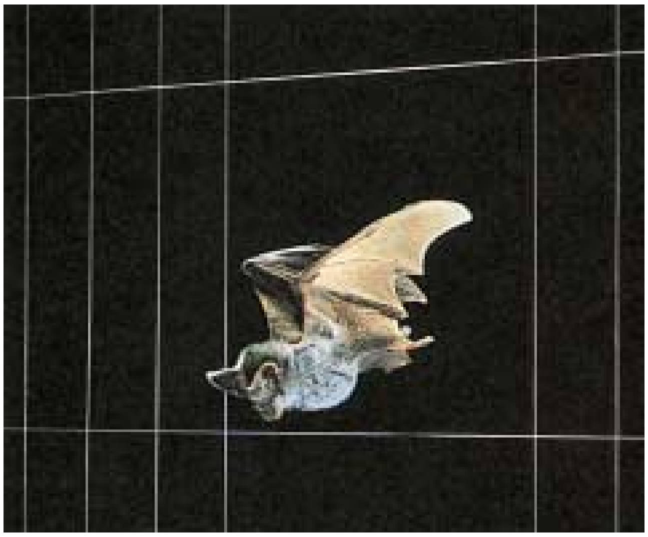
Echolocation allows bats not only to locate targets in the dark but also to analyze the features of targets, as well as features of the environment in general. Through echolocation, a bat identifies prey, navigates through the leaves of trees, and locates surfaces suitable to land on. Perhaps a term analogous to *visualization*, such as "audification," would be more appropriate.

How, exactly, does echolocation work? The answer is, Rather like radar. The larynx of a bat emits bursts of ultrasonic sound that bounce off objects and return to the bat's ears, allowing the animal to tell what is in the surrounding environment. The bat, in other words, navigates by the echoes that it hears, differentiating among the different characteristics of those echoes. Objects that are moving, such as insects, have a moving echo. Smooth objects give a different echo from that of rough objects, and so on. A key component of this echolocation system is the analysis of differences in the return times of echoes. Close objects return echoes sooner than more distant objects do, and there are minute differences in return times, depending on the texture of an object's surface.

A bat's cries are of short duration (from 0.3 to 200 milliseconds) and high frequency (from 12,000 to 200,000 hertz), mostly at too high a frequency for the human ear to detect. Different bat species produce sounds of different frequency, with the particular frequency used depending on the animal's ecology. For instance, bats that catch prey in the open use different frequencies from those used by bats that catch insects in foliage and from those used by bats that hunt for prey on the ground.

Figure 9-27

A bat with a 40-centimeter wingspan can navigate through a 14-centimeter mesh made of 80-micrometer nylon thread while flying in total darkness. The sonar of the bat is capable of creating an accurate map of the world that is based entirely on auditory information.


Steven Dalton/NHPA

The echolocation abilities of bats are impressive. The results of laboratory studies have shown that a bat with a 40-centimeter wing span can fly in total darkness through a 14-by-14-centimeter grid of nylon threads only 80 micrometers thick, as shown in Figure 9-27. Bats in the wild can be trained to catch small food particles thrown up into the air in the dark. These echolocating skills make the bat a very efficient hunter. For instance, the little brown bat can capture very small flying insects, such as mosquitoes, at the remarkable rate of two per second.

There has been considerable interest in the neural mechanisms of the bat echolocation system. Each bat species emits sounds in a relatively narrow range of frequencies, and a bat's auditory pathway has cells specifically tuned to echoes in the frequency range of its species. For example, the mustached bat sends out sounds ranging from 60,000 to 62,000 hertz, and its auditory system has a *cochlear fovea* (a maximally sensitive area in the organ of Corti) that corresponds to that frequency range. In this way, more neurons are dedicated to the frequency range used for echolocation than to any other range of frequencies. Analogously, our visual system dedicates more neurons to the retina's fovea, the area responsible for our most detailed vision. In the cortex of the bat's brain, there are several distinct areas for processing complex echo-related inputs. For instance, one area computes the distance of given targets from the animal, whereas another area computes the velocity of a moving target. This neural system makes the bat exquisitely adapted for nighttime navigation.

In Review

The analysis of birdsong has identified several important principles of auditory functioning. One principle is that specialized structures in the brain produce and perceive vocal stimuli. Another is that these structures are influenced by early experience. Third, an innate template imposes an important constraint on the nature of the songs that a bird produces and perceives. These principles underlying birdsong are similar to those observed in human language and reinforce the idea that many characteristics of human language may be innate. Insect-eating bats have evolved a different use of sound from that of birds. Such bats employ high-frequency sounds as a type of sonar that allows them to fly in the dark and to catch insects. An echolocating bat's auditory world is much richer than ours because it contains information about the shape and velocity of objects—information that our visual system provides.

SUMMARY

1. *What is the nature of the stimulus that the brain perceives as sound?* The stimulus for the auditory system is change in air pressure. The ear transduces changes in air pressure into what we perceive as sound. Sound has three fundamental physical qualities: frequency, amplitude, and complexity. Perceptually, these qualities translate into pitch, loudness, and timbre.
2. *How does the nervous system transform changes in air pressure into our impression of sounds?* Sound waves are transformed into perceptions of sound through a combination of mechanical and neural activities. The auditory receptor apparatus is the cochlea, located in the inner ear. The auditory receptors themselves are hair cells that are found on the basilar membrane. Changes in air pressure are conveyed in a chain reaction from the eardrum to the bones of the middle ear to the oval window of the cochlea and the cochlear fluid that lies behind it. Movements of the cochlear fluid produce movements in specific regions of the basilar membrane, leading to changes in the activity of the hair cells. The basilar membrane has a tonotopic organization. High-frequency sounds maximally stimulate hair cells at one of its ends, whereas low-frequency sounds maximally stimulate hair cells at the other end. In this way, cochlear neurons code for the various frequencies of sounds. The tonotopic organization of sound analysis is found at all levels of the auditory system. The auditory system also detects both sound amplitude and sound location. Sound amplitude is coded by the firing rate of cochlear neurons, with loud sounds producing higher firing rates than soft sounds do.

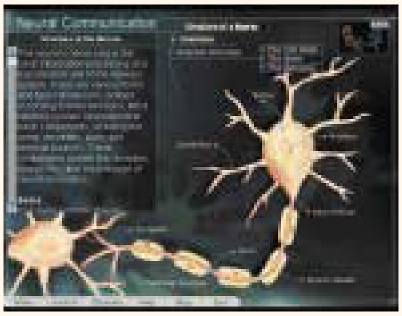
Sound location is detected by structures in the brainstem that compute differences in the arrival times and the loudness of a sound in the two ears.

3. *How does auditory information get from the receptors to the brain?* The hair cells of the cochlea synapse with bipolar neurons that form the cochlear nerve, which in turn forms part of the eighth cranial nerve. The cochlear nerve takes auditory information to three structures in the hindbrain: the cochlear nucleus, the olivary complex, and trapezoid body. Cells in these areas are sensitive to differences in sound intensity and sound arrival time in the two ears. In this way, they enable the location of a sound to be determined. The auditory pathway continues from the hindbrain areas to the inferior colliculus of the midbrain, then to the medial geniculate nucleus in the thalamus, and finally to the auditory cortex. Cells in the cortex are responsive to specific categories of sound, such as sounds used in communication by a given species.
4. *How does the brain understand language and music?* Despite their differences in speech sounds and grammar, all human languages have the same basic structure. This fundamental similarity implies that the brain possesses a basic template for creating language. The auditory areas of the cortex in the left hemisphere play a special role in analyzing language-related information, whereas those in the right hemisphere play a special role in analyzing music-related information. Studies have revealed several language-processing areas in the left hemisphere. For instance, Wernicke's area identifies speech syllables and words and so is critically engaged in speech comprehension. Broca's area matches speech sounds to the motor programs necessary to make them and so plays a major role in speech production. Broca's area also discriminates between closely related speech sounds. The primary auditory cortex of the right hemisphere plays a critical role in comprehending music. The right temporal lobe analyzes the musical qualities of speech, known as prosody.
5. *How does brain organization relate to the unique auditory worlds of other species?* Nonhuman species often use auditory analysis in specialized behaviors. One example is birdsong. Songbirds have regions of the brain that are specialized for producing and comprehending song. In many species, these regions are lateralized to the left hemisphere, just as the language areas are lateralized to the left hemisphere in humans. There are also striking similarities between the development of song in birds and the development of language in humans, as well as similarities in the neural mechanisms underlying both the production and the perception of song and language. Another example of the use of sound in nonhuman species is the auditory control of movements. Both owls and bats use auditory information to guide movement in the dark. These animals can fly and catch prey at night by using only auditory information. Bats have the added ability to produce a type of sonar that allows them to create an auditory map of the objects in their world. This type of sensory reality we can only try to imagine.

KEY TERMS

aphasia, p. 340
basilar membrane, p. 327
Broca's aphasia, p. 338
Broca's area, p. 338
cochlear implant, p. 333
echolocation, p. 350
Heschl's gyrus, p. 330
hertz, p. 321
positron emission tomography (PET), p. 342
prosody, p. 325
supplementary speech area, p. 341
tonotopic representation, p. 332
tuning curve, p. 332
Wernicke's aphasia, p. 338
Wernicke's area, p. 330

neuroscience Interactive



There are many resources available for expanding your learning on line:

- **www.worthpublishers.com/kolb/chapter9**

Try some self-tests to reinforce your mastery of the material. Look at some of the updates reflecting current research on the brain. You'll also be able to link to other sites which will reinforce what you've learned.

- **www.nad.org**

Link to the National Association of the Deaf and learn about living with deafness.

- **www.aphasia.org**

Do more research about aphasia at the site for the National Aphasia Association.

On your CD-ROM you'll be able to quiz yourself on your comprehension of the chapter. The module on the Central Nervous System also provides important review of the auditory pathways and cortical anatomy important for understanding this chapter.

REVIEW QUESTIONS

1. What are the three principal qualities of sound, and how does the auditory system code them?
2. How does the auditory system code the location of a sound?
3. Why do all human languages have the same basic structure?
4. How is language perception organized in the brain?
5. How is blood flow measured in the brain, and what does it tell us about brain function?
6. Give a simple neurobiological explanation of how we understand and produce language.
7. What can we learn from birdsong that is relevant to human auditory function?

FOR FURTHER THOUGHT

1. Different species have different ranges of hearing. Why would this be adaptive?
2. What is special about language and music?

RECOMMENDED READING

Drake-Lee, A. B. (1992). Beyond music: Auditory temporary threshold shift in rock musicians after a heavy metal concert. *Journal of the Royal Society of Medicine, 85,* 617–619. Have you ever wondered what listening to loud music might be doing to your hearing? This paper looks at the effects of hearing loud music at a rock concert on hearing thresholds in the musicians. What is important to remember is that the musicians are standing beside the speakers, so those in the front rows are likely to hear even louder music.

Gazzaniga, M. S. (1992). *Nature's mind*. New York: Basic Books. Michael Gazzaniga is an eminent cognitive neuroscientist who has an easy writing style. He has written several popular books, such as *Nature's Mind*, each of which is chock full of interesting ideas about how the brain works. This book is a pleasure to read and introduces the reader to Gazzaniga's ideas about why the brain is asymmetrically organized and what the fundamental differences between the hemispheres might be.

Luria, A. R. (1972). *The man with a shattered world.* Chicago: Regnery. Alexander Luria wrote many neuropsychology books in his long career, but this one is perhaps the most interesting and accessible to the nonspecialist. This book describes the effect of a bullet wound to the head of a university student who was recruited to defend Leningrad in World War II. The book has many anecdotes, often humorous, that show how the young man's mental world was severely altered by this traumatic experience. Reading this book can be a source of insight into what it is like to cope with brain damage.

Pinker, S. (1997). *How the mind works.* New York: Norton. Stephen Pinker gives us a provocative look at theories of how brain activity produces mental events. For those interested in cognitive neuroscience, this book is a good introduction to questions we might ask in everyday life. For example, why does a face look more attractive with makeup? Or, why is the thought of eating worms disgusting?