

CHAPTER

3

Structure and Function of the Auditory System

CHAPTER OBJECTIVES

- To be able to identify basic anatomical landmarks of the outer ear, middle ear, inner ear, and auditory portions of the central nervous system;
- To understand the primary functions of the outer ear, middle ear, inner ear, and auditory portions of the central nervous system;
- To gain insight into how the primary functions of each portion of the auditory system are supported by underlying anatomy and physiology; and
- To understand some basic perceptual aspects of sound such as hearing threshold, loudness, pitch, and masking.

KEY TERMS AND DEFINITIONS

- **Auditory periphery:** The outer ear, middle ear, and inner ear, ending at the nerve fibers exiting the inner ear.
- **Auditory central nervous system:** The ascending and descending auditory pathways in the brainstem and cortex.
- **Tonotopic organization:** The systematic mapping of sound frequency to the place of maximum stimulation within the auditory system that begins in the cochlea and is preserved through the auditory cortex.
- **Transducer:** A device or system that converts one form of energy to another. The cochlea can be considered a mechano-electrical transducer because it converts mechanical vibrations to electrical energy to stimulate the afferent nerve fibers leading to the brainstem.

52 SECTION 1 ■ The Communication Chain

Recall that the primary purpose of the communication chain is to support the transfer of thoughts, ideas, or emotions between two people, the sender and the receiver. Thus far, we have seen that the sender can translate the thought to be communicated into a code of sound sequences representing meaningful words and sentences in the language shared by these two individuals. How is this acoustic code converted into information that can be understood by the mind of the receiver or listener? The mind of the receiver makes use of neural electrical signals generated by the central nervous system. Thus, one of the primary functions of the auditory system is to convert the acoustic energy produced by the talker—human speech sounds—into neural energy that can be deciphered by the brain of the listener. This process is the central topic of this chapter.

This chapter is divided into three main sections. The first two deal with the anatomy, physiology, and functional significance of the peripheral and central sections of the auditory system. The peripheral portion of the auditory system is defined here as the structures from the outer ear through the auditory nerve. The auditory portions of the central nervous system begin at the cochlear nucleus and end at the auditory centers of the cortex. The third section of this chapter details some fundamental aspects of the perception of sound.

PERIPHERAL AUDITORY SYSTEM

Figure 3.1 shows a cross-section of the **auditory periphery**. This portion of the auditory system is usually further subdivided into the outer ear, middle ear, and inner ear.

Outer Ear

The outer ear consists of two primary components: the pinna and the ear canal. The pinna is the most visible portion of the ear, which extends laterally from the side of the head. It is composed of cartilage and skin. The ear canal is the long, narrow canal leading to the eardrum (or tympanic membrane). The eardrum represents the boundary between the outer ear and the middle ear. The entrance to this canal is called the external auditory meatus. The deep bowl-like portion of the pinna adjacent to the external auditory meatus is known as the concha.

The outer ear serves a variety of functions. First, the long (2.5-cm), narrow (5- to 7-mm) canal makes the more delicate middle and inner ear less accessible to foreign objects. The outer third of the canal is composed of skin and

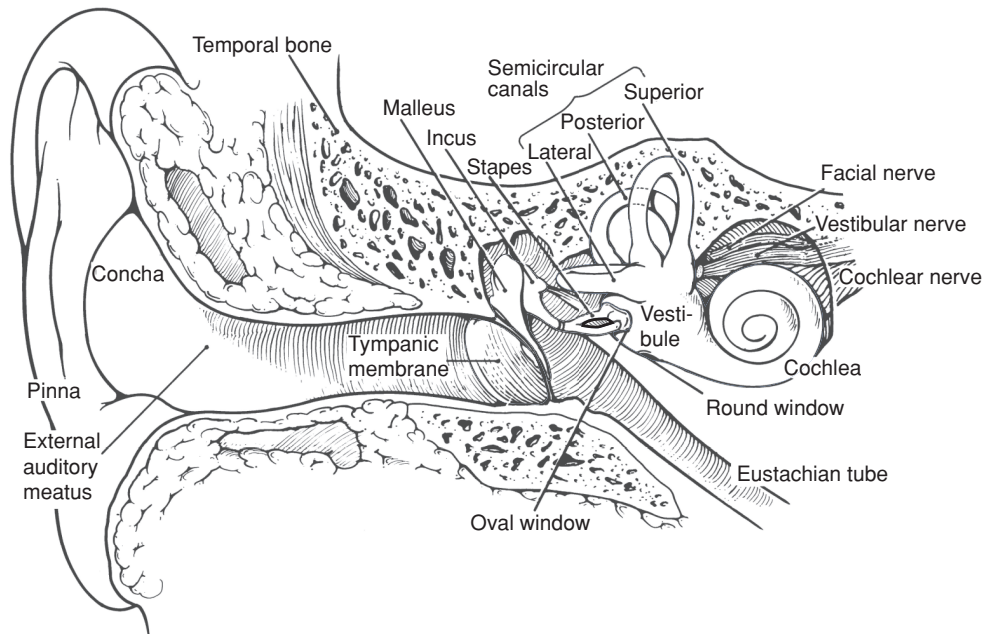


FIGURE 3.1 A cross-section of the peripheral portion of the auditory system revealing some of the anatomic details of the outer, middle, and inner ear. (Adapted from Kessel RG, Kardon RH. *Tissues and Organs: A Text-Atlas of Scanning Electron Microscopy*. San Francisco: WH Freeman; 1979, with permission.)

cartilage lined with glands and hairs. These glands, known as ceruminous glands, secrete a substance that potential intruders, such as insects, find terribly noxious. So both the long, narrow, tortuous path of the canal and the secretions of these glands serve to protect the remaining portions of the peripheral auditory system.

Second, the various air-filled cavities composing the outer ear, the two most prominent being the concha and the ear canal, have a natural or resonant frequency to which they respond best. This is true of all air-filled cavities and tubes. For an adult, the resonant frequency of the ear canal is approximately 2,500 Hz, whereas that of the concha is roughly 5,000 Hz. The resonance of each of these cavities is such that each structure increases the sound pressure at its resonant frequency by approximately 10 to 12 dB. This gain or increase in sound pressure provided by the outer ear can best be illustrated by considering the following hypothetical experiment.

Let us begin the experiment by using two tiny microphones. One microphone will be placed just outside (lateral to) the concha, and the other will be

54 SECTION 1 ■ The Communication Chain

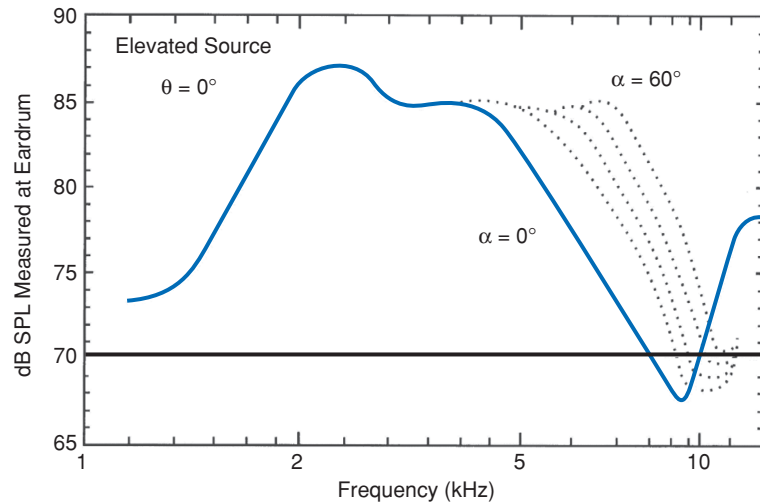


FIGURE 3.2 The response of the head, outer ear, and ear canal for various angles of elevation (a) of the sound source. Zero degrees corresponds to a sound source at eye level and straight ahead (0 degrees azimuth, 0 degrees elevation), whereas 60 degrees represents a source located straight ahead but at a higher elevation. If the listener's head and outer ear had no influence on the sound level measured at the eardrum, a flat line at 70 dB SPL would result. This figure illustrates the amplification of high-frequency sound by the outer ear and the way that this amplification pattern changes with elevation of the sound source. (Adapted from Shaw EAG. The external ear. In: Keidel WD, Neff WD, eds. *Handbook of Sensory Physiology*, vol 1. New York: Springer Verlag; 1974:463, with permission.)

positioned very carefully inside the ear canal to rest alongside the eardrum. Now if we present a series of sinusoidal sound waves, or pure tones, of different frequencies, which all measure 70 dB SPL at the microphone just outside the concha, and if we read the sound pressure levels measured with the other microphone near the eardrum, we will obtain results like those shown by the solid line in Figure 3.2. Notice that, at frequencies less than approximately 1,400 Hz, the microphone that is located near the eardrum measures sound levels of approximately 73 dB SPL. This is only 3 dB higher than the sound level just outside the outer ear. Consequently, the outer ear exerts little effect on the intensity of low-frequency sound. As the frequency of the sound is increased, however, the intensity of the sound measured at the eardrum increases to levels considerably above 70 dB SPL. The maximum sound level at the eardrum is reached at approximately 2,500 Hz and corresponds to a value of approximately 87 dB SPL. Thus, when sound waves having a frequency of

2,500 Hz enter the outer ear, their sound pressure levels are increased by 17 dB by the time they strike the eardrum. The function drawn with a solid line in Figure 3.2 illustrates the role that the outer ear serves as a resonator or amplifier of high-frequency sounds. The function shown is for the entire outer ear. Experiments similar to the one just described can be conducted to isolate the contribution of various cavities to the resonance of the total outer ear system. Again, the results of such experiments suggest that the two primary structures contributing to the resonance of the outer ear are the concha and the ear canal.

The resonance of the outer ear, which is represented by the solid line in Figure 3.2, was obtained with a sound source located directly in front of the subject at eye level. If the sound source is elevated by various amounts, a different resonance curve is obtained. Specifically, the notch or dip in the solid function that is located at 10 kHz in Figure 3.2 moves to a higher frequency and the peak of the resonant curve broadens to encompass a wider range of frequencies as the sound source is increased in elevation. This is illustrated in Figure 3.2 by the dotted lines. Each dotted line represents a different angle of elevation. An elevation of 0 degrees corresponds to eye level, whereas a 90-degree elevation would position the sound source directly overhead. The response of the outer ear changes as the elevation of the sound source changes. This results from the angle at which the incident sound wave strikes the various cavities of the outer ear. The result is that a code for sound elevation is provided by the outer ear. This code is the amplitude spectrum of the sound, especially above 3000 Hz, that strikes the eardrum. Thus, the outer ear plays an important role in the perception of the elevation of a sound source.

Finally, the outer ear also assists in another aspect of the localization of a sound source. The orientation of the pinnae is such that the pinnae collect sound more efficiently from sound sources located in front of the listener than from sources originating behind the listener. The attenuation of sound waves originating from behind the listener assists in the front/back localization of sound. This is especially true for high-frequency sounds (i.e., sounds with short wavelengths).

In summary, the outer ear serves four primary functions. First, it protects the more delicate middle and inner ears from foreign bodies. Second, it boosts or amplifies high-frequency sounds. Third, the outer ear provides the primary cue for the determination of the elevation of a sound's source. Fourth, the outer ear assists in distinguishing sounds that arise from in front of the listener from those that arise from behind the listener.

Middle Ear

The middle ear consists of a small (2-cm³) air-filled cavity lined with a mucous membrane. It forms the link between the air-filled outer ear and the fluid-filled

56 SECTION 1 ■ The Communication Chain

inner ear (Fig. 3.1). This link is accomplished mechanically via three tiny bones, the ossicles. The lateral-most ossicle is the malleus. The malleus is in contact with the eardrum or tympanic membrane. At the other end of the outer ear/inner ear link is the smallest, medial-most ossicle, the stapes. The broad base of the stapes, known as the footplate, rests in a membranous covering of the fluid-filled inner ear referred to as the oval window. The middle ossicle in the link, sandwiched between the malleus and stapes, is the incus. The ossicles are suspended loosely within the middle ear by ligaments extending from the anterior and posterior walls of the cavity.

We have mentioned that the middle ear cavity is air-filled. The air filling the cavity is supplied via a tube that connects the middle ear to the upper part of the throat, or the nasopharynx. This tube, known as the auditory tube or the eustachian tube, has one opening located along the bottom of the anterior wall of the middle ear cavity. The tube is normally closed but can be readily opened by yawning or swallowing. In adults, the eustachian tube assumes a slight downward orientation. This facilitates drainage of fluids from the middle ear cavity into the nasopharynx. Thus, the eustachian tube serves two primary purposes. First, it supplies air to the middle ear cavity and thereby enables an equalization of the air pressure on both sides of the eardrum. This is desirable for efficient vibration of the eardrum. Second, the eustachian tube permits the drainage of fluids from the middle ear into the nasopharynx.

What is the purpose of the elaborate link between the air-filled outer ear and the fluid-filled inner ear formed by the three ossicles? If the middle-ear cavity did not exist and the oval window on the cochlea was placed at the location of the eardrum, the acoustic energy would directly strike the fluid-filled inner ear and 99.9% of the energy in the impinging sound wave would be reflected away, with only 0.1% transmitted through to the inner ear. This loss amounts to a decrease in sound level of approximately 30 dB. Consequently, if such an arrangement existed, there would be a considerable loss of sound energy.

The middle ear compensates for this loss of sound energy when going from air to a fluid medium through two primary mechanisms. The first of these, the areal ratio (ratio of the areas) of the tympanic membrane to the footplate of the stapes, accounts for the largest portion of the compensation. The second mechanism, referred to as a lever action of the ossicle, makes a very small contribution to the recovery of sound energy and is not considered further in this text. Regarding the areal ratio, the effective area of the tympanic membrane (i.e., the area involved directly in the mechanical link between the outer ear and inner ear) is approximately 55 mm². The corresponding area of the stapes footplate is 3.2 mm². Pressure (p) may be defined in terms of force (F) per unit area (A): $p = F/A$. If the force applied at the eardrum is the same as that reaching

the stapes footplate, then the pressure at the smaller footplate must be greater than that at the larger eardrum. As an analogy, consider water being forced through a hose. If the area of the opening at the far end of the hose is the same as that of the faucet to which it is connected, the water will exit the hose under the same water pressure as it would at the faucet. If a nozzle is now attached to the far end of the hose and it is adjusted to decrease the size of the opening at that end of the hose, the water pressure at that end will be increased in proportion to the degree of constriction produced by the nozzle. The smaller the opening, the greater the water pressure at the nozzle (and the farther the water will be ejected from the nozzle). Applying the same force that exists at the faucet to push the water through a smaller opening created by the nozzle has increased the water pressure at the nozzle. Another analogy explaining the pressure gain associated with areal ratios is one that explains why carpentry nails have a broad head and sharp, narrow point (Vignette 3.1).

The middle ear system, primarily through the difference in area between the eardrum and the stapes footplate, compensates for much of the loss of sound energy that would result if the airborne sound waves impinged directly on the fluid-filled inner ear. Approximately 25 to 27 dB of the estimated 30-dB loss has been compensated for by the middle ear. The ability of the middle ear system to amplify or boost the sound pressure depends on signal frequency. Specifically, little pressure amplification occurs for frequencies below 100 Hz or above 2,000 to 2,500 Hz. Recall, however, that the outer ear amplified sound energy by 20 dB for frequencies from 2,000 to 5,000 Hz. Thus, taken together, the portion of the auditory system peripheral to the stapes footplate increases sound pressure by 20 to 25 dB in a range of approximately 100 to 5,000 Hz. This range of frequencies happens to correspond to the range of frequencies in human speech that are most important for communication. As is noted later in this chapter, a decrease of 10 dB in sound level for moderate-intensity sounds, such as conversational speech, corresponds to a halving of loudness. Without the amplification of sound waves provided by the outer and middle ears, conversational speech produced by the talker would sound like soft speech or whispered speech to the listener and effective communication would be jeopardized.

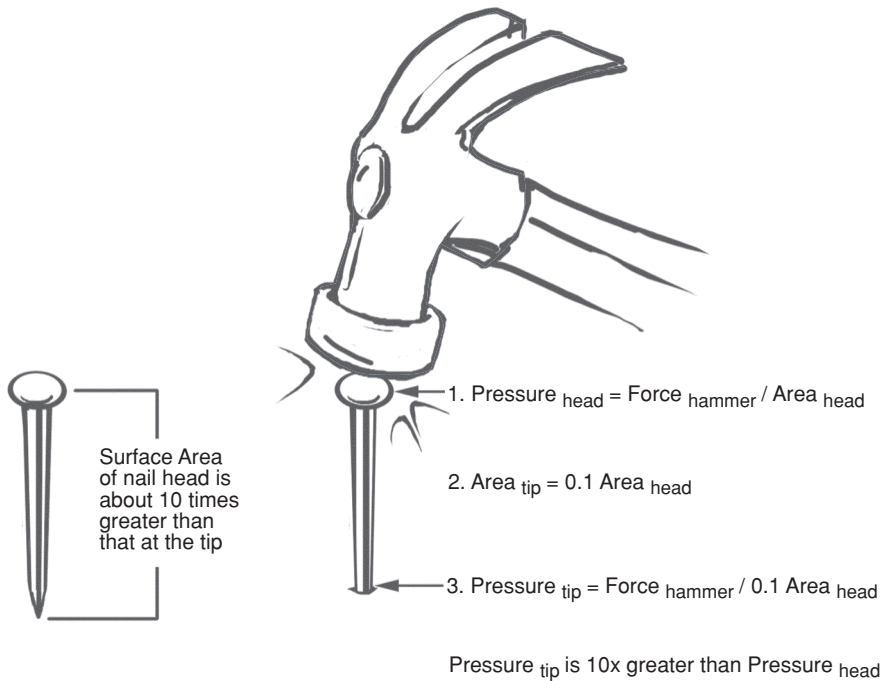
Another less obvious function of the middle ear also involves the outer ear/inner ear link formed by the ossicles. Because of the presence of this mechanical link, the preferred pathway for sound vibrations striking the eardrum will be along the chain formed by the three ossicles. Sound energy, therefore, will be routed directly to the oval window. There is another membranous window of the inner ear that also lies along the inner or medial wall of the middle ear cavity. This structure is known as the round window (Fig. 3.1). For the inner ear to be stimulated appropriately by the vibrations of the sound waves, it is important that the oval window and round window not be displaced



VIGNETTE 3.1 CONCEPTUAL DEMO

AN ANALOGY FOR PRESSURE

The common carpentry nail provides an illustration of the pressure amplification that occurs when the same force is applied over both a larger and smaller surface area. As shown in the figure, the head of the nail has a greater surface area than the narrow point of the nail. When force is applied to the broader head of the nail with a hammer, that same force is channeled through to the narrow point of the nail. However, because the area of the narrow point at the end of the nail is approximately 10 times smaller than the area at the head of the nail, the pressure applied at the narrow point is 10 times greater. (Recall that pressure is equal to force per unit area or $\text{Pressure} = \text{Force}/\text{Area}$.) The increased pressure at the point of the nail enables it to better penetrate the material (wood, for example) into which it is being driven. There are, of course, other practical factors affecting the design of the common carpentry nail (for example, the broader head also makes it easier to strike and to remove), but the resulting pressure amplification is one of the key elements in the design and represents a nice analogy to the pressure amplification achieved by the areal ratio of the tympanic membrane and stapes footplate in the middle ear.



in the same direction simultaneously. This situation would arise frequently, however, if the sound wave impinged directly on the medial wall of the middle ear cavity where both the oval window and the round window are located. Thus, routing the vibrations of the eardrum directly to the oval window via the ossicles assures appropriate stimulation of the inner ear.

Inner Ear

The inner ear is a complex structure that resides deep within a very dense portion of the skull known as the petrous portion of the temporal bone. Because of the complexity of this structure, it is often referred to as a labyrinth. The inner ear consists of a bony outer casing, the osseous labyrinth. Within this bony structure is the membranous labyrinth that contains the actual sensory structures. The osseous labyrinth, as shown in Figure 3.3, can be divided into three major sections: the *semicircular canals* (*superior*, *lateral*, and *posterior*), the *vestibule*, and the *cochlea*. The first two sections house the sensory organs for the vestibular system. The vestibular system assists in maintaining balance and posture. The focus here, however, will be placed on the remaining portion of the osseous inner ear, the cochlea. It is the cochlea that contains the sensory organ for hearing. The coiled, snail-shaped cochlea has approximately 2.75 turns in human beings. The largest turn is called the basal turn, and the smallest turn at the top of the cochlea is referred to as the apical turn. Two additional

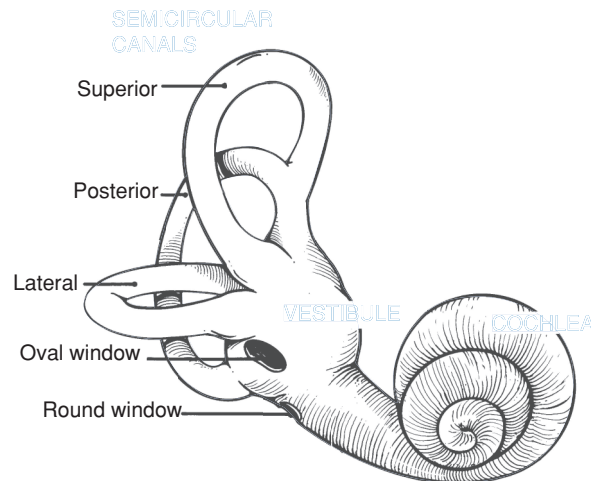


FIGURE 3.3 Illustration of the osseous labyrinth and its landmarks. (Adapted from Durrant JD, Lovrinic JH. *Bases of Hearing Science*. 2nd ed. Baltimore: Williams & Wilkins; 1984:98, with permission.)

60 SECTION 1 ■ The Communication Chain

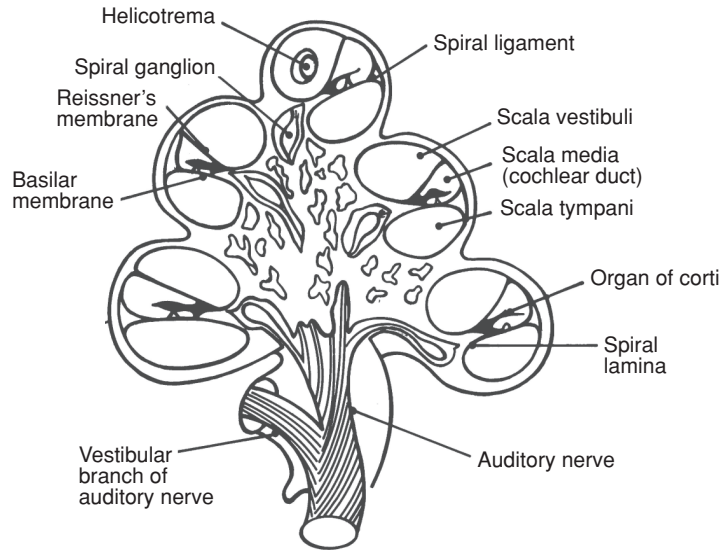


FIGURE 3.4 Modiolar cross-section of the cochlea illustrating the scalae through each of the turns. (Adapted from Zemlin WR. *Speech and Hearing Science: Anatomy and Physiology*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall; 1988:464, with permission.)

anatomic landmarks of the inner ear depicted in Figure 3.3 are the *oval window* and the *round window*. Recall that the footplate of the stapes, the medial-most bone of the three ossicles in the middle ear, is attached to the oval window.

The cochlea is cut in cross-section from top (apex) to bottom (base) in Figure 3.4. The winding channel running throughout the bony snail-shaped structure is further subdivided into three compartments. The middle compartment, the *scala media*, is sandwiched between the other two, the scala tympani and scala vestibuli. The scala media is a cross-section of the membranous labyrinth that runs throughout the osseous labyrinth. All three compartments are filled with fluid, although the fluids are not identical in each compartment. The scala media is filled with a fluid called endolymph, which differs considerably from the fluid in the other two compartments, called perilymph. The *scala media* houses the sensory organ for hearing, the organ of Corti. When the oval window vibrates as a consequence of vibration of the ossicles, a pressure wave is established within the fluid-filled inner ear that causes the scala media, and the structures within it, to move in response to the vibration. This displacement pattern for the scala media is usually simplified by considering the motion of just one of the partitions forming the scala media, the *basilar membrane* (Fig. 3.4). Although the displacement pattern of the basilar membrane is depicted, it should be noted that the entire fluid-filled scala media is undergoing similar displacement.

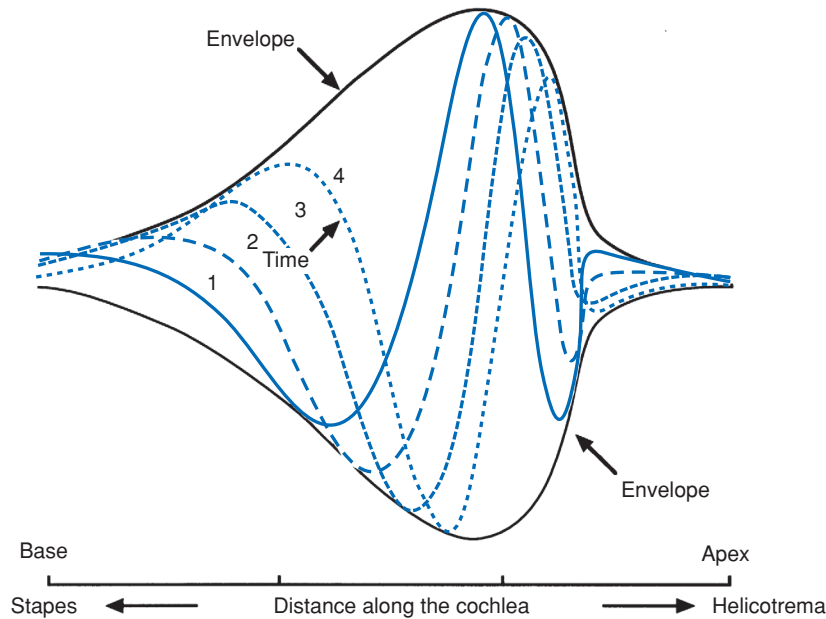


FIGURE 3.5 Illustration of the traveling wave pattern at four instants in time in response to a mid-frequency sound. The *thin solid lines* connect the maximum and minimum displacements at each location along the cochlea and for each instant in time. These lines represent the envelope of the displacement patterns. (Adapted from von Békésy G. *Experiments in Hearing*. New York: McGraw-Hill; 1960:462.)

Figure 3.5 illustrates the displacement pattern of the basilar membrane at four successive instants in time. When this displacement pattern is visualized directly, the wave established along the basilar membrane is seen moving or traveling from the base to the apex. The displacement pattern increases gradually in amplitude as it progresses from the base toward the apex until it reaches a point of maximum displacement. At that point, the amplitude of displacement decreases abruptly. The thin solid line connecting the amplitude peaks at various locations for these four instants in time describes the displacement envelope. The envelope pattern is symmetrical in that the same pattern superimposed on the positive peaks can be flipped upside down and used to describe the negative peaks. As a result, only the positive or upper half of the envelope describing maximum displacement is usually displayed.

Figure 3.6 displays the envelopes of the displacement pattern of the basilar membrane that have been observed for different stimulus frequencies. Note that as the frequency of the stimulus increases, the peak of the displacement pattern moves in a basal direction closer to the stapes footplate. At low frequencies,

62 SECTION 1 ■ The Communication Chain

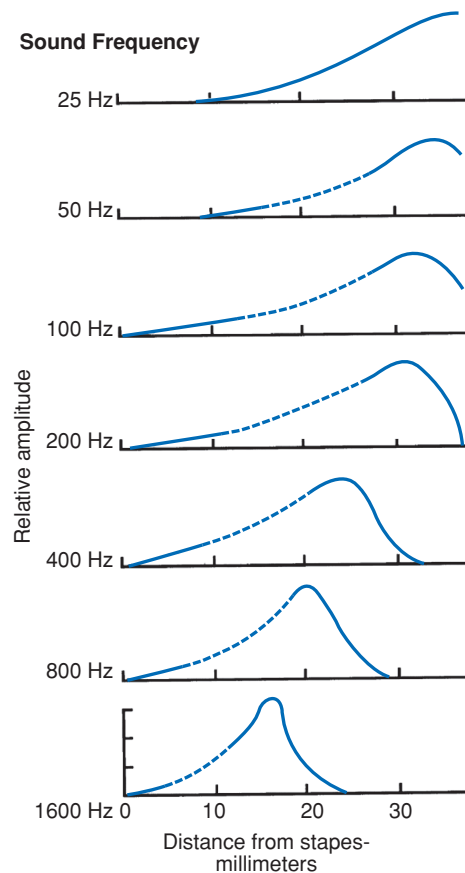


FIGURE 3.6 Envelopes of traveling wave displacement patterns illustrated for different stimulus frequencies. Notice that low frequencies (*top*) produce a maximum displacement of the basilar membrane in the apex (farthest distance from the stapes), whereas higher frequencies (*bottom*) produce maximum displacement in the basal portion of the cochlea (nearer to the stapes). (Adapted from von Békésy G. *Experiments in Hearing*. New York: McGraw-Hill; 1960:448, with permission.)

virtually the entire membrane undergoes some degree of displacement. As stimulus frequency increases, a more restricted region of the basilar membrane undergoes displacement. Thus, the cochlea is performing a crude frequency analysis of the incoming sound. In general, for all but very low frequencies (≤ 50 Hz), the place of maximum displacement within the cochlea is associated with the frequency of an acoustic stimulus. The frequency of an acoustic stimulus striking the eardrum and displacing the stapes footplate will be analyzed or distinguished from sounds of different frequency by the location of the displacement pattern along the basilar membrane. This precise mapping of sound frequency to place of

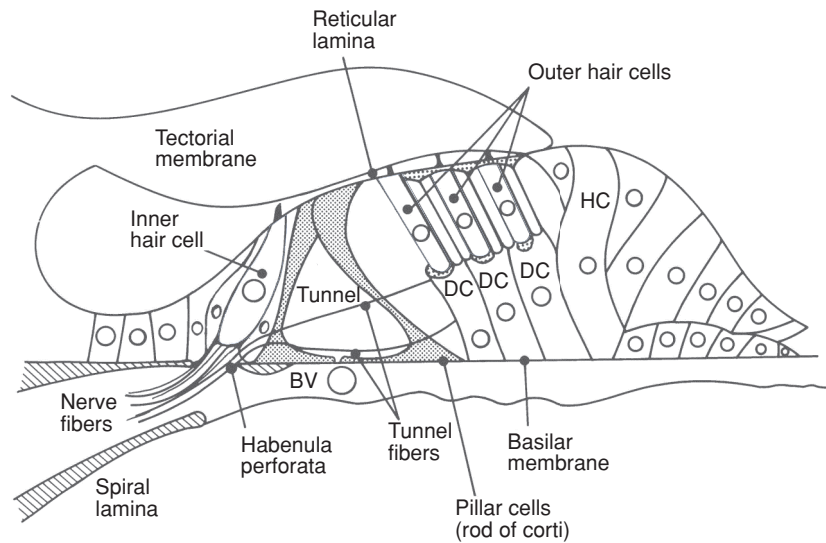


FIGURE 3.7 Detailed cross-section of the organ of Corti. *BV*, basilar vessel; *DC*, Deiter cell; *HC*, Hensen cell. (Adapted from Pickles JO. *An Introduction to the Physiology of Hearing*. 2nd ed. London: Academic Press; 1988:29.)

maximum mechanical activity in the cochlea is referred to as **tonotopic organization** (*tono* = frequency; *topic* = place; *tonotopic* = mapping of frequency to place of stimulation). Tonotopic organization is an important property of the auditory system, one that begins in the cochlea and is preserved to the auditory cortex, and forms the basis for viable code for sound frequency.

A more detailed picture of the structures within the scala media is provided in Figure 3.7. The sensory organ of hearing, the organ of Corti, is seen to rest on top of the basilar membrane. The organ of Corti contains several thousand sensory receptor cells called *hair cells*. Each hair cell has several tiny hairs or cilia protruding from the top of the cell. As shown in Figure 3.7, there are two types of hair cells within the organ of Corti: inner hair cells and outer hair cells. Approximately 90% to 95% of the auditory nerve fibers that carry information to the brain make contact with the inner hair cells. The outer hair cells are much greater in number, but do not have as many nerve-fiber connections to the upper portions of the auditory system (only 5% to 10% of nerve fibers connect to the outer hair cells). One of the critical functions of the hair cells in the organ of Corti is the conversion (**transducer** function) of mechanical energy from the middle-ear vibration to electrical energy that stimulates the connecting nerve fibers. As noted previously, this is a critical link in the communication chain in converting the acoustic signal from the sender to a neural electrical signal that can be deciphered by the brain of the receiver. The conversion takes

64 SECTION 1 ■ The Communication Chain

place within the inner ear in the organ of Corti. Intact inner and outer hair cells are needed for this conversion to take place. As we will see later, many agents—such as noise, genetic disorders, or aging—can harm or destroy these hair cells and cause a breakdown in the communication chain. Although exciting research is underway that may alter this statement in the near future, at the moment, such loss of hair cells in the organ of Corti is irreversible.

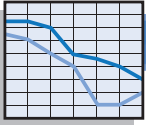
If only a small percentage of outer hair cells (5% to 10%) are actually connected to the brain, what purpose do they serve? A complete answer to this question is beyond the scope of this book. Suffice it to say that the outer hair cells serve to augment the mechanical response of the cochlea when stimulated by vibrations. They enable the inner hair cells to be stimulated by much lower sound levels than would be possible otherwise. Without them, the inner hair cells would not be able to respond to sound levels lower than about 50 to 60 dB SPL. In addition, in the course of altering the mechanical response of the cochlea for incoming sound or mechanical vibration, the outer hair cells generate their own vibrations that can be recorded as sounds in the ear canal using a tiny microphone. These sounds, known as otoacoustic emissions (OAEs) can be measured and their presence can be used to verify a normal functioning cochlea (Vignette 3.2).

The two primary functions of the auditory portion of the inner ear can be summarized as follows. First, the inner ear performs a frequency analysis on incoming sounds so that different frequencies stimulate different regions of the inner ear (i.e., tonotopic organization). Second, mechanical vibration is amplified and converted into electrical energy by the hair cells. The hair cells are frequently referred to as mechano-electrical transducers. That is, cells that convert mechanical energy (vibration) into electrical energy that leads to the generation of electrical activity in the auditory nerve.

Auditory Nerve

The electrical potentials generated by auditory nerve fibers are called all-or-none action potentials because they do not vary in amplitude when activated. If the nerve fibers fire, they always fire to the same degree, reaching 100% amplitude. The action potentials, moreover, are very short-lived events, typically requiring less than 1 to 2 ms to rise in amplitude to maximum potential and return to resting state. For this reason, they are frequently referred to as *spikes*. Information is coded in the auditory portions of the central nervous system via patterns of neural spikes. For example, as the intensity of sound is increased, the rate at which an auditory nerve fiber fires also increases, although only over a limited range of variation in sound intensity.

By varying the intensity and frequency of the sound stimulus, it is possible to determine the frequency to which a given nerve fiber responds best; that is,



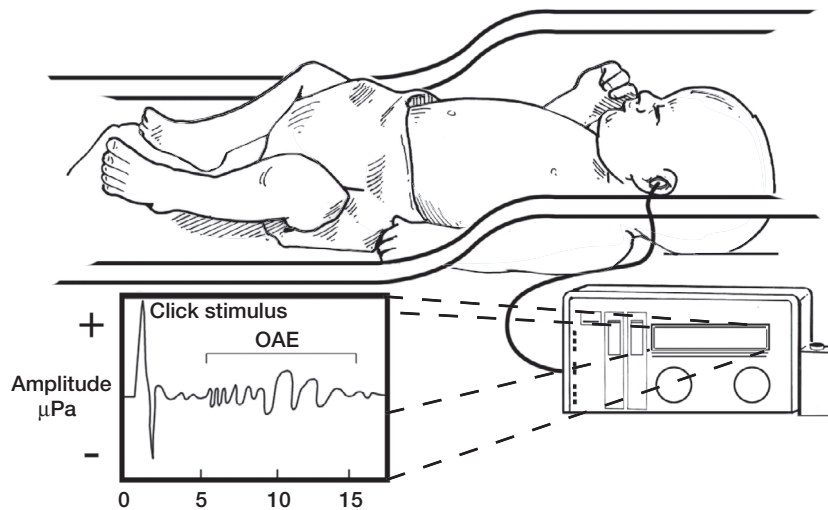
VIGNETTE 3.2 CLINICAL APPLICATIONS

USE OF OAEs AS A SCREENING TOOL FOR BABIES

The development of tools to measure OAEs clinically has led to many exciting applications. For the first time in the history of clinical audiology, a tool was made available that measured the integrity of the sensory receptors in the generally inaccessible inner ear.

One of the clinical applications explored and refined during the 1990s was the use of OAEs as a screening tool for hearing loss in infants. Some of the advantages associated with the use of OAEs for this application included the ability to record these responses without requiring the active participation of the infant, the ability to assess each ear separately, and the capability of evaluating a fairly broad range of frequencies. In addition, normal OAE responses would only be possible with normal outer, middle, and inner ears. The focus in OAE-based screening is clearly on the auditory periphery. Problems localized to any of these sections of the auditory periphery, moreover, are most amenable to intervention, either medical (for outer and middle ears) or audiologic (in the form of hearing aids or other prosthetic devices for those with inner ear problems).

Many states have adopted legislation in recent years that requires that all infants be screened for hearing loss at birth (referred to as "universal screening"). Screening tools based on OAEs represent one of the most commonly used devices in universal newborn screening programs.



66 SECTION 1 ■ The Communication Chain

the frequency that requires the least amount of intensity to produce an increase in firing rate. This frequency is often referred to as the *best frequency* or *characteristic frequency* of the nerve fiber. Some nerve fibers have a low characteristic frequency and some have a high characteristic frequency. Fibers with high characteristic frequencies come from (inner) hair cells in the base of the cochlea, whereas those with low characteristic frequencies supply the apex. As the nerve fibers exit through the bony core of the cochlea on their way to the brainstem, they maintain an orderly arrangement. The bundle of nerve fibers composing the cochlear branch of the auditory nerve is organized so that fibers with high characteristic frequencies are located around the perimeter, whereas fibers with low characteristic frequencies make up the core of the cochlear nerve. Thus, the auditory nerve is organized, as is the basilar membrane, so that each characteristic frequency corresponds to a specific anatomical location or place. This mapping of the frequency of the sound wave to place of maximum activity within an anatomic structure is referred to as *tonotopic organization*. The neural place that responds to a particular sound stimulus provides the brain with important information regarding the frequency content of that sound.

Temporal or time-domain information is also coded by fibers of the auditory nerve. Consider, for example, the discharge pattern that occurs within a nerve fiber when that stimulus is a sinusoid that lies within the response area of the nerve fiber. The pattern of spikes that occurs under such conditions is illustrated in Figure 3.8. Note that when the single nerve fiber discharges, it always does so at essentially the same location on the stimulus waveform. In Figure 3.8, this happens to be the positive peak of the waveform. Notice also that it may not fire during every cycle of the stimulus waveform. Nonetheless, if one were to record the interval between successive spikes and examine the number of times each interval occurred, a histogram of the results would look like that shown in Figure 3.8. This histogram, known simply as an interval histogram, indicates that the most frequent interspike interval corresponds to the period of the waveform. All other peaks in the histogram occur at integer multiples of the period. Thus, the nerve fiber is able to encode the period of the waveform. This holds true for nerve fibers with characteristic frequencies less than approximately 5,000 Hz. As discussed in Chapter 2, if we know the period of a sinusoidal waveform, we know its frequency ($f = 1/T$). Hence, the nerve fibers responding in the manner depicted in Figure 3.8 could code the frequency of the acoustic stimulus according to the timing of discharges. For frequencies up to 5,000 Hz, the neural firing is synchronized to the sound stimulus, as in Figure 3.8A. By combining synchronized firings for several nerve fibers, it is possible to encode the period of sounds up to 5,000 Hz in frequency. This type of frequency coding may be useful for a wide variety of sounds, including the coding of the fundamental frequency for complex sounds like vowels or musical notes.

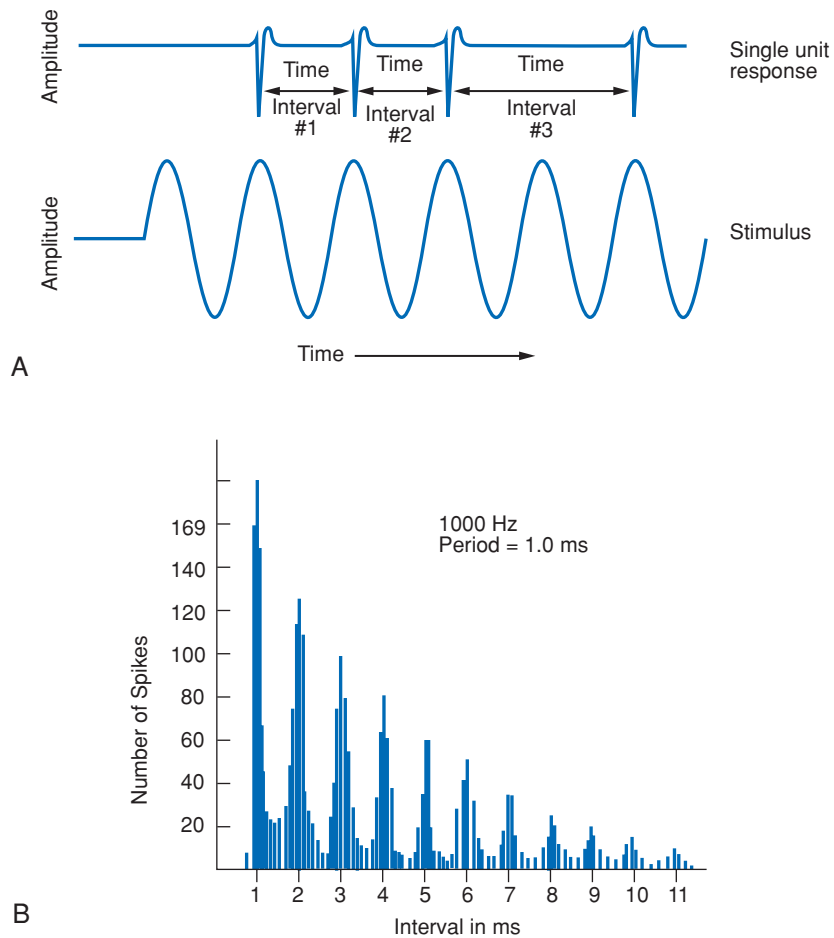


FIGURE 3.8 A, Illustration of the synchronization of nerve fiber firings to the stimulus waveform. Note that the nerve fiber always fires at the same point of the waveform, although it may not fire every cycle. The intervals between successive firings of the nerve fiber can be measured and stored for later analysis. B, A histogram of the intervals measured in (A), referred to as interspike intervals. An interval of 1 ms was the most frequently occurring interval, having occurred approximately 180 times. This corresponds to the period of the waveform in (A).

The electrical activity of the auditory nerve can also be recorded from more remote locations. In this case, however, the action potentials are not being measured from single nerve fibers. Recorded electrical activity under these circumstances represents the composite response from a large number of nerve fibers. For this reason, this composite electrical response is referred to frequently as the *whole-nerve action potential*. Because the whole-nerve action potential represents

68 SECTION 1 ■ The Communication Chain

the summed activity of several nerve fibers, the greater the number of fibers that can be made to fire simultaneously, the greater the summed amplitude of the response. For this reason, brief abrupt acoustic signals, such as clicks, are used as stimuli. A click stimulus rises to maximum amplitude in a very short period of time and lasts for a short period of time (e.g., less than 1 ms). A click is also comprised of many frequencies from 0 to 10,000 Hz. As such, a click stimulus stimulates nerve fibers from the high-frequency basal region to the low-frequency apical region of the cochlea and does so almost instantaneously. With such a large number of nerve fibers firing nearly simultaneously, the electrical activity generated by the cochlea's nerve fibers can be measured with recording electrodes in the ear canal or on the scalp.

AUDITORY CENTRAL NERVOUS SYSTEM

Once the action potentials have been generated in the cochlear branch of the auditory nerve, the electrical activity progresses up toward the cortex. This network of nerve fibers is frequently referred to as the **auditory central nervous system** (auditory CNS). The nerve fibers that carry information in the form of action potentials up the auditory CNS toward the cortex form part of the ascending or afferent pathways. Nerve impulses can also be sent toward the periphery from the cortex or brainstem centers. The fibers carrying such information compose the descending or efferent pathways.

Figure 3.9 is a simplified schematic diagram of the ascending auditory CNS. All nerve fibers from the cochlea terminate at the cochlear nucleus on the same side. From here, however, several possible paths are available. The majority of nerve fibers cross over or decussate at some point along the auditory CNS, so that the activity of the right ear is represented most strongly on the left side of the cortex and vice versa. The crossover, however, is not complete. From the superior olives through the cortex, activity from both ears is represented on each side. In fact, this represents one of the key functions of the auditory CNS: combining or integrating the neural information from both ears. This is referred to as binaural (or “two-ear”) processing and is critical for some aspects of sound perception, including our ability to locate sounds in space. All ascending fibers terminate in the medial geniculate body before ascending to the cortex. Thus, all ascending fibers within the brainstem portion of the auditory CNS synapse at the cochlear nucleus and at the medial geniculate body, taking one of several paths between these two points, with many paths having additional intervening nerve fibers. Vignette 3.3 describes measurement of the auditory brainstem response and its correlation with anatomic structure.

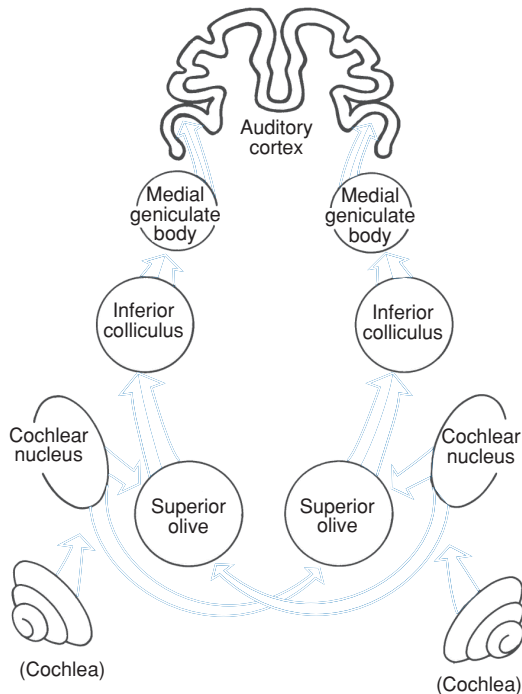
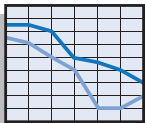


FIGURE 3.9 The ascending pathways of the auditory central nervous system. The *blue arrows* represent input from the right ear; the *white arrows* represent input from the left ear. (Adapted from Yost WA, Nielsen DW. *Fundamentals of Hearing: An Introduction*. 2nd ed. New York: Holt, Rinehart, & Winston; 1985:98, with permission.)

We have already reviewed the simple coding of information available in the responses of the auditory nerve fibers. The mapping of frequency to place within the cochlea, for example, was preserved in the responses of the nerve fibers such that fibers having high characteristic frequencies originated from the high-frequency base of the cochlea. The period of the waveform could also be coded for stimulus frequencies less than 5,000 Hz. In addition, the intensity of the stimulus is coded over a limited range (30 to 40 dB) by the discharge rate of the fiber. At the level of the cochlear nucleus, it is already apparent that the ascending auditory pathway begins processing information by converting this fairly simple code into more complex codes. The coding of timing information, for example, is much more complex in the cochlear nucleus. In addition, some nerve fibers within the cochlear nucleus have a much broader range of intensities (up to 100 dB) over which the discharge rate increases steadily with sound intensity. As one probes nerve fibers at various centers within the auditory CNS, a tremendous diversity of responses is evident.

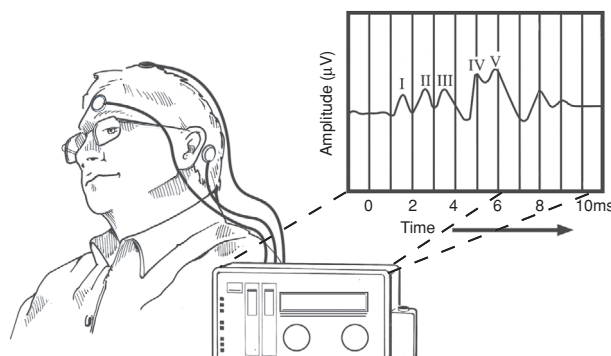


VIGNETTE 3.3 CLINICAL APPLICATIONS

AUDITORY BRAINSTEM RESPONSE MEASUREMENTS

It was mentioned earlier that the whole-nerve action potential represented the summed response of many single nerve fibers firing synchronously in response to an abrupt acoustic signal. It was also mentioned that this potential could be recorded remotely from the ear canal. The left-hand portion of the drawing that accompanies this vignette shows a patient with electrodes pasted to the skin of the forehead, the top of the head (vertex), and the area behind the pinna (the mastoid prominence). The tracing in the right-hand portion of this illustration shows the electrical activity recorded from the patient. The acoustic stimulus is a brief click that produces synchronized responses from nerve fibers in the cochlea and the brainstem portion of the auditory CNS. The tracing represents the average of 2,000 stimulus presentations presented at a moderate intensity at a rate of 11 clicks per second. Approximately 3 minutes is required to present all 2,000 stimuli and to obtain the average response shown below. Note that the time scale for the x axis of the tracing spans from 0 to 10 ms. This represents a 10-ms interval beginning with the onset of the click stimulus. The tracing shows several distinct bumps or waves, with the first appearing at approximately 1.5 ms after stimulus onset. This first wave, wave I, is believed to be a remote recording of the whole-nerve action potential from the closest portion of the auditory nerve. Approximately 1 ms later, 2.5 ms after stimulus onset, wave II is observed. This wave is believed to be the response of the more distant portion of the auditory nerve. One millisecond later, the electrical activity has traveled to the next center in the brainstem, the cochlear nucleus, and produces the response recorded as wave III. Wave IV represents the activity of the superior olivary complex. Wave V represents the response of the lateral lemniscus, a structure lying between the superior olives and the inferior colliculus. Waves VI and VII (the two unlabeled bumps after wave V) represent the response of the latter brainstem structure.

The response shown in the right-hand tracing is known as an auditory brainstem response (ABR). It has proven very useful in a wide variety of clinical applications, from assessment of the functional integrity of the peripheral and brainstem portions of the ascending auditory CNS to assessment of hearing in infants or difficult-to-test patients.



Despite this increasing anatomic and physiologic complexity, one thing that appears to be preserved throughout the auditory CNS is tonotopic organization. At each brainstem center and within the auditory portions of the cortex, there is an orderly mapping of frequency to place. This can be demonstrated by measuring the characteristic frequency of nerve fibers encountered at various locations within a given brainstem center or within the cortex.

Another principle underlying the auditory CNS is that of redundancy. That is, information represented in the neural code from one ear has multiple representations at various locations within the auditory system. Every auditory nerve fiber, for example, splits into two fibers before entering the cochlear nucleus, with each branch supplying a different region of the cochlear nucleus. In addition, from the superior olives through the auditory areas of the cortex, information from both ears is represented at each location in the auditory CNS. This redundancy in the auditory CNS helps to protect the communication chain against breakdowns. A lesion, such as a tumor, on one side of the brainstem, for example, will not necessarily prevent information from both ears reaching the auditory portions of the cortex.

Perception of Sound

The basic structure and some key aspects of the physiologic function of the auditory system have been reviewed; the remainder of this chapter reviews some fundamental aspects of the perception of sound by humans. We have examined the acoustics involved in the generation and propagation of a sound wave and have reviewed its conversion into a complex neural code. This code of incoming sensory information can influence the behavior of a human subject, whether the sound is the loud whistle of an oncoming train or the soft voice of a precocious toddler.

The lowest sound pressure that can be reliably heard by a human listener is referred to as the listener's hearing threshold. The results obtained from the measurement of hearing thresholds at various frequencies are depicted in Figure 3.10. Note that the sound pressure level that is just detectable varies with frequency, especially below 500 Hz and above 8,000 Hz. The range of audibility of the normal-hearing human ear is described frequently as 20 to 20,000 Hz. Acoustic signals at frequencies above or below this range typically cannot be heard by the normal human ear.

The phenomenon of masking has also been studied in detail. Masking refers to the ability of one acoustic signal to obscure the presence of another acoustic signal so that it cannot be detected. A whisper might be audible, for example, in a quiet environment. In a noisy industrial environment, however, such a weak acoustic signal would be masked by the more intense factory noise.

72 SECTION 1 ■ The Communication Chain

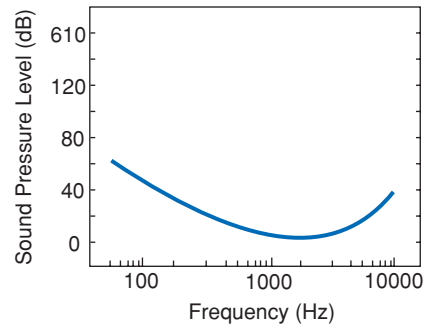


FIGURE 3.10 Average normal threshold sound pressure level plotted as a function of frequency for binaural (two-ear) listening in a free field. (Adapted from Sivian LJ, White SD. Minimum audible pressure and minimum audible field. *J Acoust Soc Amer.* 1933;4:288–321.)

Masking is used clinically to aid the audiologist in the diagnosis of hearing difficulties, as will be discussed in more detail in a subsequent chapter.

Loudness is clearly one of the more salient perceptual features of a sound and it also has considerable clinical importance. As noted previously, there are many agents that can destroy the outer hair cells of the cochlea, including high-intensity noise and aging. When this occurs, the listener loses the ability to hear soft sounds in the frequency regions associated with the loss of hair cells. For example, if the hair cell loss was in the basal portion of the cochlea, then the listener would have difficulty hearing low-intensity, high-frequency sounds. The loudness of high-intensity sounds, however, is not impacted much by the loss of outer hair cells and this same individual would hear loud sounds normally. As will be noted later, this represents a potential “break” in the communication chain with input to the higher centers distorted by a loss of hair cells in the cochlea. In this particular case of loss of outer hair cells only, the conventional hearing aid represents a reasonable approach to the repair of the communication chain.

Pitch is another very salient perceptual feature of sound. In general, pitch is the perceptual correlate of sound frequency. We have already seen that the auditory system is capable of coding sound frequency via place of stimulation (tonotopic organization) and timing information, such as the period of the waveform. This perceptual pitch information can aid in the identification of specific talker voices or fundamental frequencies, as well as elements of the speech sounds produced by a talker, such as the formant frequencies of vowels. It is generally believed that the pitch associated with sound frequencies below 50 Hz are mediated exclusively by timing information because, as noted previously, tonotopic organization breaks down at these very low frequencies. On the other hand, because timing information breaks down above 5,000 Hz due to the loss of syn-

chronous neural firing of the auditory nerve, place coding predominates at these higher frequencies. For frequencies between 50 and 5,000 Hz, both place-based and timing-based codes for sound frequency may underlie the perceived pitch. An overview of some interesting early research on timing-based and place-based codes for pitch perception is provided in Vignette 3.4.

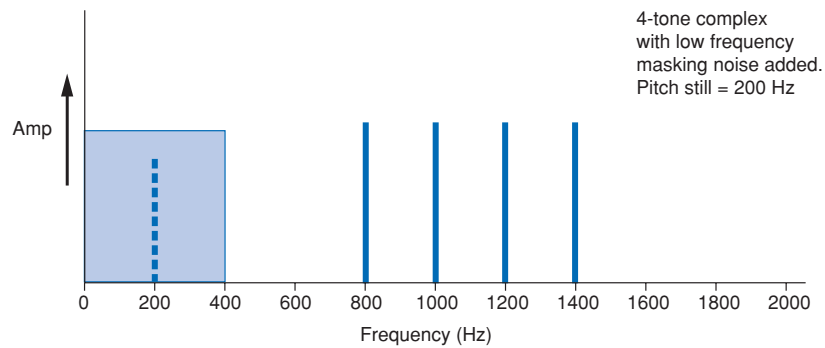
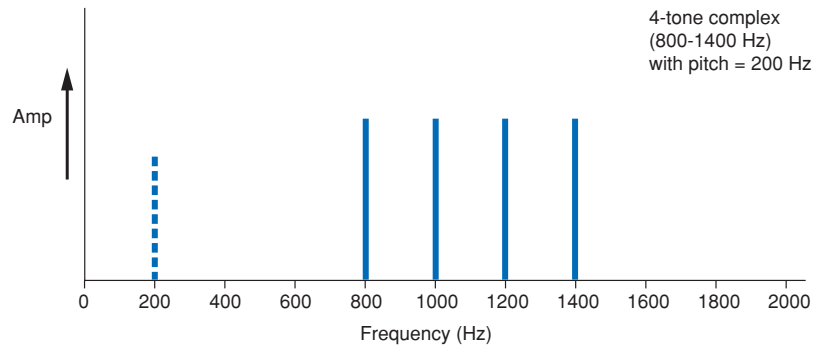


VIGNETTE 3.4 EXPERIMENT

PITCH PERCEPTION AND “THE CASE OF THE MISSING FUNDAMENTAL”

As noted, the pitch of a sound is often one of its most salient perceptual characteristics. Musicians and hearing scientists have been interested in how humans perceive pitch for centuries, but probably the greatest progress in our understanding of pitch perception occurred during the 20th century.

In the late 19th century, probably the most accepted theory of pitch perception was based on a “place principle” akin to tonotopic organization described previously in the discussion of basic auditory structure and function. Although many details of early



74 Audiology, Hearing Loss, and Communication Disorders

place-principle theories of pitch perception were incorrect, the basic premise that certain portions of the auditory system, especially the inner ear, were “tuned” to specific frequencies was correct. Basically, low-frequency pure tones produced a low pitch sensation and high-frequency pure tones produced a high pitch sensation because each frequency stimulated a different region or place within the inner ear.

In the late 19th and early 20th centuries, however, several researchers in the Netherlands produced a low-frequency pitch using sounds that were comprised of only higher frequencies. An example of the amplitude spectrum for one such sound is shown in the figure. Notice that this sound is comprised of pure tones at frequencies of 800, 1,000, 1,200, and 1,400 Hz. Yet, the pitch of this sound was judged to be 200 Hz by listeners. A frequency of 200 Hz would correspond to the fundamental frequency of these sounds with 800 Hz being the fourth harmonic of 200 Hz (i.e., $800 \text{ Hz} = 4 \times 200 \text{ Hz}$), 1,000 Hz being the fifth harmonic, 1,200 Hz being the sixth harmonic, and 1,400 Hz being the seventh harmonic of 200 Hz. For this reason, the pitch of such a series of pure tones is often referred to as the “missing fundamental.”

Now, the place theorists simply argued that the equipment or the ear (the middle ear was believed to be the culprit at the time) generated a distortion product that corresponded to the 200-Hz frequency, and the listener’s pitch perception was dominated by this low-frequency distortion tone. Therefore, the fundamental frequency (200 Hz) was not really missing after all. In this way, the results were entirely consistent with the place theory. The complex of four pure tones from 800 to 1,400 Hz generated a distortion product at 200 Hz; this 200-Hz distortion product stimulated the place associated with a 200-Hz pure tone yielding a corresponding pitch.

To counter this argument, the Dutch researchers conducted the following experiment. A low-frequency noise was introduced with enough acoustical energy below 400 Hz to mask low-frequency tones of moderate intensity (as demonstrated with a pure tone at 200 Hz, for example). The four pure tones from 800 to 1,400 Hz were played in this background of low-frequency noise, and a pitch of 200 Hz was again perceived. Basically, the low-frequency masking noise rendered the low-frequency region of the inner ear unusable, and yet, a low-frequency pitch was still perceived when the four tones from 800 to 1,400 Hz were presented. Clearly, the place theory could not account for these findings.

Recognizing the limitations of the place theory as a single explanation for the perception of pitch, hearing scientists began to focus on possible timing or “periodicity” cues present in the four-tone complex. Today, although several details of pitch perception remain unexplained, the majority of findings can be described by a “duplex” theory that relies on exclusive use of place cues above approximately 5,000 Hz (recall that synchronization of nerve fiber firing breaks down at frequencies above approximately 5,000 Hz, making timing or “periodicity” cues unavailable), exclusive use of timing cues

for very low frequencies below approximately 50 Hz, and a combination of both periodicity and place cues for frequencies from 50 to 5,000 Hz.

Incidentally, there is a standardized scale of pitch sensation referred to as the Mel scale (no, not named after a famous scientist named “Mel” but derived from “melody”). A pure tone of 1,000 Hz at 40 dB SPL is said to have a pitch of 1,000 mels (as well as a loudness of 1 sone and a loudness level of 40 phons). A sound with a pitch judged to be twice as high as this standard would have a pitch of 2,000 mels, whereas one with half the pitch of the standard sound would be assigned a pitch value of 500 mels. Often, however, pitch is measured by matching the frequency of a pure tone to the perceived pitch of the sound under evaluation.

SUMMARY

In this chapter, we have seen how the acoustic information, serving as the link between the sender and the receiver in the communication chain, is encoded by the auditory system. Initially, the pressure fluctuations corresponding to the acoustic speech sounds are converted into mechanical vibrations of the middle ear and organ of Corti. These mechanical vibrations are converted to electrical signals by the organ of Corti and ultimately lead to the generation of electrical signals in the auditory nerve. The sensory input representing the acoustic speech stimulus is now in the “electrical language” of the brain and is subject to further processing by the auditory centers of the brainstem and cortex. This neural auditory information in the auditory cortex is then processed, as needed, by other centers of the brain, such as various language-processing centers, to make sense of the incoming sensory information. Impairments in the auditory periphery or auditory portions of the central nervous system can result in “breaks” in the communication chain, with degraded neural information reaching the higher centers of the auditory system. The identification of such impairments is the focus of the Chapter 4.

CHAPTER REVIEW QUESTIONS

1. What are the primary functions of the outer ear?
2. What is the primary function of the middle ear and what is the main anatomical feature that supports this function?
3. The Eustachian tube serves two main functions. What are they?
4. Define tonotopic organization. Where does it originate? Based on this anatomical organization, if an individual lost all inner and outer hair cells in the base of the cochlea, what sound frequencies would this person have difficulty hearing?
5. What is meant by the “transducer function” of the cochlea?
6. Frequency can be encoded by the auditory periphery using a place-based code or a timing-based code. Explain what this means.
7. What structures represent the beginning and ending of the auditory central nervous system?
8. Aside from the initial letter in their spelling, what is the difference between “afferent” and “efferent” nerve fibers?
9. In very general terms, loudness and pitch can be considered to be the perceptual correlates of what two physical aspects of sound waves?

REFERENCES AND SUGGESTED READINGS

- Durrant JD, Lovrinic JH. *Bases of Hearing Science*. 2nd ed. Baltimore: Williams & Wilkins; 1995.
- Geisler CD. *From Sound to Synapse*. New York: Oxford University Press; 1998.
- Gelfand SA. *Hearing: An Introduction to Psychological and Physiological Acoustics*. 3rd ed. New York: Marcel Dekker; 1997.
- Hall JW. *Handbook of Auditory Evoked Responses*. Needham, MA: Allyn & Bacon; 1992.
- Kidd J. Psychoacoustics. In Katz J, ed. *Handbook of Clinical Audiology*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2002.
- Möller AR. *Auditory Physiology*. New York: Academic Press; 1983.
- Moore BCJ. *An Introduction to the Psychology of Hearing*. 4th ed. London: Academic Press; 1997.
- Pickles JO. *An Introduction to the Physiology of Hearing*. 2nd ed. London: Academic Press; 1988.
- Probst R, Lonsbury-Martin BL, Martin G. A review of otoacoustic emissions. *J Acoust Soc Amer*. 1991;89:2027–2067.

CHAPTER 3 ■ Structure and Function of the Auditory System 77

- Shaw EAG. The external ear. In Keidel WD, Neff WD, eds. *Handbook of Sensory Physiology*. Vol. 1. New York: Springer Verlag; 1974: 455–490.
- von Békésy G. *Experiments in Hearing*. New York: McGraw-Hill; 1960.
- Wever EG, Lawrence M. *Physiological Acoustics*. Princeton: Princeton University Press; 1954.
- Yost WA. *Fundamentals of Hearing: An Introduction*. 4th ed. New York: Academic Press; 2000.
- Zemlin WR. *Speech and Hearing Science: Anatomy and Physiology*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall; 1988.
- Zwicker E, Fastl H. *Psychoacoustics*. 2nd ed. Berlin: Springer-Verlag; 1999.

