No matter how a listener analyzes a message, the data on which he or she operates are the acoustic patterns of speech. The essential step, then, is that the listener hears the speech. Because the hearing mechanism per se is somewhat removed from the main concerns of this book, we say only a few words about the peripheral reception of speech, as the auditory system itself imposes certain changes on speech sounds.

HEARING

The human auditory mechanism analyzes sound according to changes in frequency and intensity as a function of time. As a receptor, the ear falls short of the eye in sensitivity, but it seems to be remarkably responsive to the sounds that humans produce, the sounds of speech. These sounds change not only in amplitude but also in their mode of transmission as they travel through the outer ear, middle ear, cochlea, and auditory nerve to the brain. Figure 9.1 differentiates these parts of the mechanism. As we know from Chapter 3, the pressure waves of speech are usually disturbances in air and thus they continue in the outer ear. In the middle ear, however, they are converted from pressure waves to mechanical vibrations by a series of small bones leading to the cochlea of the inner ear. In the cochlea, a snail-shaped cavity within the temporal bone of the skull, the vibrations are again transformed. This time the transformation is from mechanical vibrations to vibrations in the fluid with which the cochlea is filled. Finally, the nerve endings in the cochlea act to transform the hydraulic vibrations into electrochemical changes that are sent to the brain in the form of nerve impulses.

The Outer Ear

The outer ear is composed of two parts: the external part you can readily see, called the

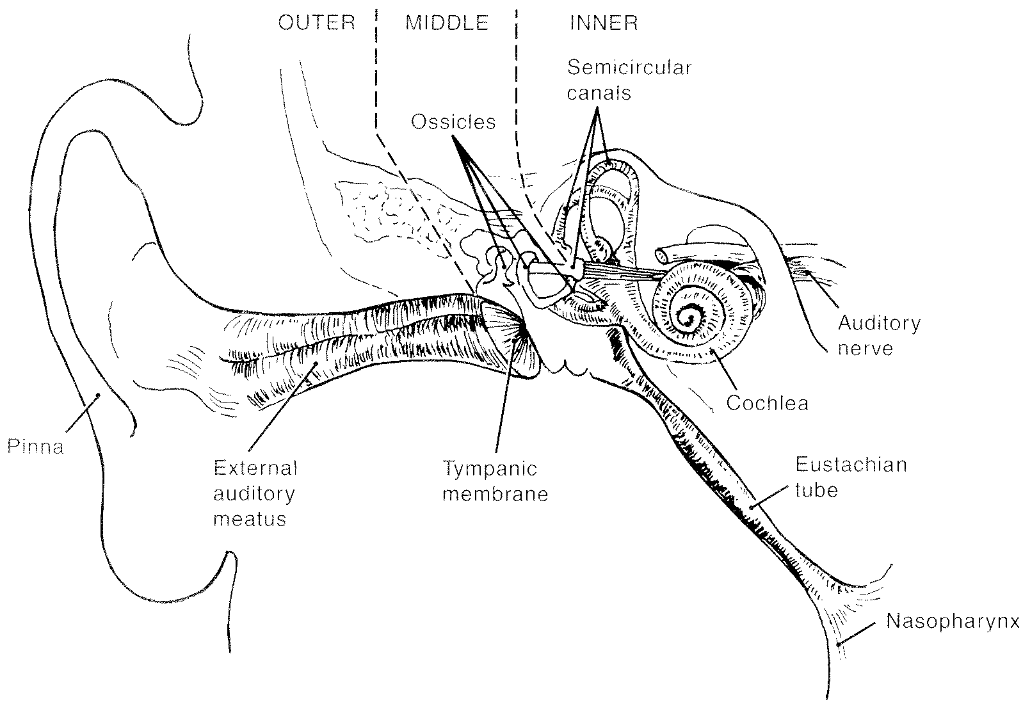

FIGURE 9.1 Frontal section of the outer, middle, and inner ear.

auricle or pinna, and the ear canal, named the external auditory meatus, leading from the pinna to the eardrum. Meatus means "channel," and the channel of the outer ear is specified by the term "external," distinguishing it from the internal auditory meatus that runs from the inner ear out of the temporal bone to the brain. The pinna funnels the sound somewhat, being a little more receptive to sounds in front of the head than behind it. The pinna also serves to protect the entrance to the canal, especially the small projection of the pinna, situated over the opening to the canal, called the tragus. One way to reduce the intensity of a loud sound is to press the tragus into the entrance to the auditory meatus with your finger.

The external auditory meatus protects the more delicate parts of the ear from trauma and from the intrusion of foreign objects. A waxy substance, cerumen, is secreted into the canal, and aided by the hairs (cilia) lining the canal, it filters out dust and any flying insects that intrude into the canal. Some people constantly clean out the cerumen, but they are depriving themselves of their natural protection.

In addition to offering protection to the more critical parts of the ear, the external auditory meatus functions to boost the high frequencies of the sounds it receives. The meatus is an air-filled cavity about 2.5 cm long and open at one end. It therefore acts as a quarter-wave resonator, with a lowest resonant frequency of about 3,440 Hz.

$$f = \text{velocity of sound}/4(\text{length}) = 34{,}400/10 = 3{,}440 \text{ Hz}$$

A woman's or child's ear canal would probably be shorter than 2.5 cm and would resonate at higher frequencies. The high-frequency emphasis provided by the outer ear is useful for the perception of fricatives because much of the sound energy that distinguishes fricatives from each other is above 2,000 Hz.

To conclude our discussion of the outer ear, let us consider the advantage of having binaural hearing and ears on both sides of our heads. That advantage lies in our ability to localize a source of sound. A person who becomes deaf in one ear can hear perfectly well in a quiet environment but has difficulty monitoring a large group conversation because localization of sound is impaired. In a meeting room with voices coming from all directions, the unilaterally deaf person seeking to locate the speaker may look in the wrong direction.

The Middle Ear

The outer ear is separated from the air-filled middle ear cavity by the eardrum, more properly called the tympanic membrane (Figs. 9.1 and 9.2). The tympanic membrane is slightly concave as seen from the outer ear and is responsive to small pressure variations across a wide range of frequencies. The tension of the eardrum can be altered by a muscle, the tensor tympani, which pulls on the manubrium, or handle, of a small bone, the malleus, attached to the inside of the drum. At low frequencies, the

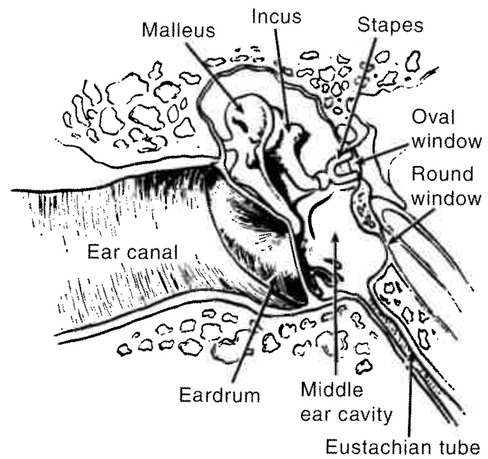

FIGURE 9.2 Cross section of the middle ear and ossicles. (Modified with permission from Denes, P. B., and Pinson, E. N., *The Speech Chain*. New York: Doubleday, 1963.)

tympanic membrane vibrates as a whole, but at high frequencies, different areas of the membrane are responsive to different frequency ranges. On the internal side of the tympanic membrane is the ossicular chain, three tiny connected bones called the ossicles. The aforementioned malleus (hammer) is attached to the tympanic membrane; the incus (anvil) acts as a fulcrum between the other two bones; and the stapes (stirrup) inserts into the membranous oval window leading to the inner ear. Thus, the ossicular chain bridges the space between the tympanic membrane and the cochlea. The chain is suspended in the air-filled cavity of the middle ear by ligaments and is held in such a delicate balance that no matter what position the body takes, the tiny bones are held in suspension, free to vibrate in response to sound. The vibrations in the outer ear take the form of disturbances of air molecules, but in the middle ear, they take the form of mechanical vibrations of the bony ossicles. The tympanic membrane and the ossicular chain taken together are especially responsive to the frequencies of acoustic signals that are important for speech.

Why have a middle ear at all? Why not have the fluid-filled cochlea on the other side of the tympanic membrane? The problem is a mismatch in impedance. Impedance is a force determined by the characteristics of a medium itself (gas, liquid, or solid) and is a measure of the resistance to transmission of signals. Liquid offers a higher impedance, or resistance, to sound pressure than does gas. When sound pressure waves traveling through air (a gas) suddenly come to a fluid, most of the sound energy is reflected back, with very little admitted into the liquid. The cochlea is filled with fluid, and to overcome the difference in impedance between air and fluid, a transformer is needed to increase the sound pressure so that more of it will be admitted into the liquid. The middle ear is that transformer.

The middle ear increases sound pressure by approximately 30 dB. The ossicles by themselves cannot effect such a large amplification of the signal, although they do act as a lever to increase the sound pressure by about 5 dB (Fig. 9.3). Leverage is the force long used by farmers to remove a heavy rock from a field. If the rock is too heavy for the farmer to lift, leverage can be used by placing a pole over a fulcrum, with the shorter part of the pole under the heavy object. A given pressure applied by the farmer results in a much larger pressure under the rock. In somewhat the same way, the pressures applied to the relatively long malleus are transmitted by the incus, which acts something like a fulcrum to the much smaller stapes. The result is an increase of a few decibels in transmission because of the increase in the pressure exerted on the oval window.

The leverage applied along the ossicles helps to overcome some of the impedance

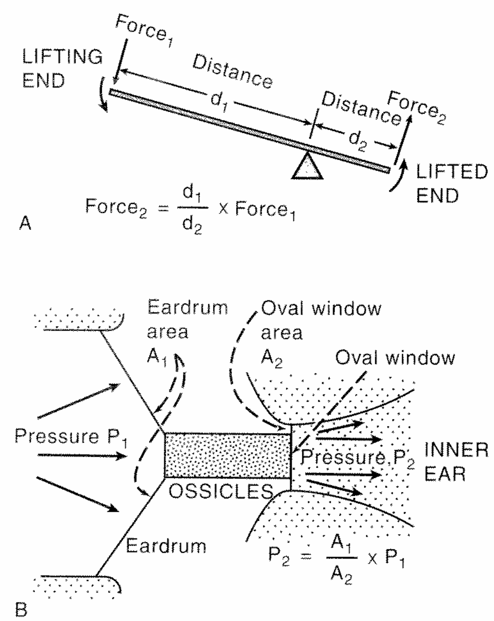

FIGURE 9.3 *A.* The lever principle of the ossicles. *B.* The effect of the area difference between the tympanic membrane and the oval window. (Modified with permission from Denes, P. B., and Pinson, E. N., *The Speech Chain*. New York: Doubleday, 1963.)

mismatch, but the larger part of the increase in pressure comes from the design of the tympanic membrane relative to the oval window. The area of the tympanic membrane is about 0.85 cm^2 (although only about 0.55 cm^2 of that area is active in vibration), whereas the area of the oval window is 0.03 cm^2. When a given force (F) is applied to a small area (A), the pressure (p) is greater than if it is applied to a larger area. This is expressed by the formula $p = F/A$. Thus, as area (A) increases, the absolute value of the fraction F/A, which is the pressure, decreases.

Consider the following example: If your friend were to fall through the ice, you would be well advised to spread your weight over a large area in attempting to reach the victim. By lying flat or, better, by distributing your weight over an even larger area, by crawling along on a ladder, you are in much less danger of falling through the ice yourself. The pressure on any point is much less than if you were to attempt to walk to your friend on the ice, focusing all the pressure at the points on the ice beneath your feet. In an analogous manner, the sound vibrations occurring over the larger vibrating area of the tympanic membrane are focused by the stapes to the smaller area of the oval window, resulting in an increase in pressure of approximately 25 dB. Thus, the impedance-matching function of the middle ear is accomplished by the area difference between the tympanic membrane and the oval window and by the leverage afforded by the ossicular design, which adds a few more decibels.

Besides the important function of impedance matching between the air and the cochlear fluid, the middle ear serves two other functions. First, it attenuates loud sounds by action of the acoustic reflex. Second, by the action of the eustachian tube, it works to maintain relatively equal air pressure on either side of the eardrum despite any changes in atmospheric pressure.

The acoustic reflex is elicited when a sound having a pressure level of 85 or 90 dB reaches the middle ear. This causes a contraction of the smallest muscle in the body, the stapedius muscle, which is attached to the neck of the smallest bone in the body, the stapes. There are two theories to account for the function of this acoustic reflex. The first, that it protects the inner ear from loud sounds, posits that the contraction of the stapedius muscle pulls the stapes to one side, changing the angle of its coupling to the oval window and reducing the pressure it applies. The second theory is that the stapedius muscle, along with the tensor tympani muscle, stiffens the ossicular chain, thereby regulating intensity changes much as the eye adjusts to changes in light. In either case, the stapedius muscle takes a few milliseconds to act, allowing sounds with sudden onset to penetrate the inner ear before the reflex occurs. Also, like any muscle, it eventually fatigues, so that in a noisy environment, the reflexive attenuation of the sound will gradually lessen, allowing the full impact of the sound pressure to impinge again on the inner ear. The stapedius muscle is innervated by the facial (seventh cranial) nerve but is somehow associated with the innervation of the larynx (vagus; tenth cranial nerve), because phonation activates the acoustic reflex. The acoustic reflex attenuates frequencies below 1 kHz by about 10 dB, and the spectral energy of the human voice is also largely below 1 kHz. The acoustic reflex may keep us from hearing ourselves too loudly, for we hear our own voice not only by the air-conducted sound coming through our outer ears but also by bone-conducted sound, as our facial and skull bones vibrate in response to our own voices.

The middle ear also equalizes differences between internal and external air pressures. This is accomplished by the eustachian tube, which leads from the middle ear to the nasopharynx. The eardrum does not vibrate properly if the air pressure in the middle ear is different from the air pressure in the external auditory meatus. Relatively high pressure within the middle ear pushes out on the tympanic membrane,

causes discomfort, and attenuates outside sounds. A sudden change in pressure, as when one drives up into the mountains or descends in an airplane, can create this pressure difference if the eustachian tube, normally closed, fails to open. The outside air pressure is suddenly lower whereas the air pressure in the middle ear cavity (containing the same air as when one was at sea level) is relatively higher. Swallowing, yawning, and chewing facilitate the opening of the tube, which is why airline attendants sometimes offer chewing gum to passengers on landing.

The Inner Ear

Within the temporal bone of the skull are several coil-shaped tunnels filled with fluid called perilymph. The fluid is like seawater in many of its properties. Floating in the fluid are coiled tubes made of membrane and filled with a more viscous fluid called endolymph. Figure 9.4 depicts the membranous labyrinth. The snail-shaped coil is the cochlear duct, containing the sensory receptor for hearing, and the system of three coils is the vestibular system, consisting of the semicircular canals, which, along with the vestibule (utricle and saccule) connecting them, contain organs that sense changes in body position and movement.

We limit our description to the cochlea, for audition is the first step in speech perception. As the footplate of the stapes vibrates in the oval window, the vibrations set up disturbances in the perilymph of the cochlea. These pressure waves in the perilymph surrounding the cochlear duct set up vibrations in the duct itself. Especially important are the resulting vibrations of the "floor" of the duct, which is called the basilar membrane.

The cochlea in humans is a cavity within bone that coils around a bony core almost three times. The membranous duct (or cochlear duct) within is attached to the bony core on the inside and by a ligament to the bony wall on the outside. It is perhaps easier to visualize if we imagine the cochlear chambers uncoiled as in Figure 9.5.

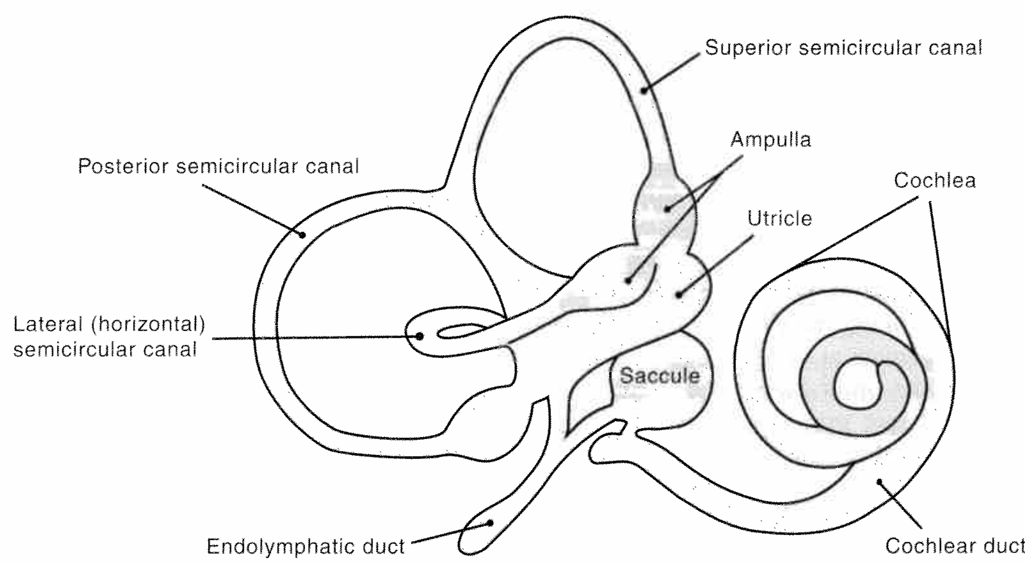

FIGURE 9.4 The parts of the membranous labyrinth. The three semicircular canals, the ampulla, utricle, and saccule make up the vestibular organs, which sense body position and movement. The cochlea contains the organ of hearing. (Modified with permission from Durrant, J. D., and Lovrinic, J. H., *Bases of Hearing Science*. Baltimore: Williams & Wilkins, Baltimore, 1977.)

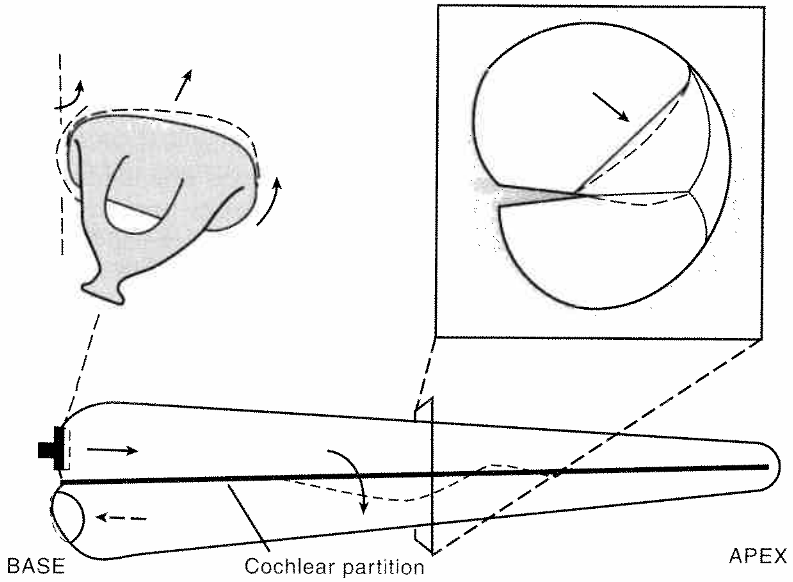

FIGURE 9.5 Uncoiled cochlea (*bottom*); a cross section is shown in the upper right. The stapes, upper left, rocks in the oval window, leading to displacement of the cochlear partition, the basilar membrane in particular. (Modified with permission from Durrant, J. D., and Lovrinic, J. H., *Bases of Hearing Science.* Baltimore: Williams & Wilkins, 1977.)

Pressure variations applied by the stapes rocking in the oval window are translated into pressure variations within the fluids of the cochlea, which in turn lead to displacements of the basilar membrane. The beauty of the system is that different parts of the basilar membrane respond to different frequencies. The membrane is narrow and stiff at the base, gradually getting wider and less stiff at the apex (the opposite of what one might expect). As a result, low-frequency sounds produce traveling waves in the fluid that stimulate the basilar membrane to vibrate with the largest amplitude of displacement at the wider, more flaccid tip. On the other hand, high-frequency sounds create pressure waves with the largest displacement of the basilar membrane at the thinner, stiffer base (Fig. 9.6).

The basilar membrane is not the sense organ of hearing, however. The organ of Corti, lying on the basilar membrane for the length of the cochlear duct, is the auditory receptor. It consists of rows of hair cells, along with other cells for support. Above the rows of thousands of hair cells is a gelatinous mass called the tectorial membrane. The basilar membrane and the tectorial membrane are attached to the cochlear duct at different points and therefore move somewhat independently. Figure 9.7 shows a cross section of the cochlea. The scala vestibuli and scala tympani containing perilymph lie on either side of the cochlear duct. Pressure waves in the perilymph set up traveling waves within the cochlear duct. In some imperfectly understood way, the undulating motions of the basilar membrane cause the hair cells to be stimulated. The tectorial membrane above the hairs shears across the hairy endings of the cells, and the result is electrochemical excitation of the nerve fibers serving the critical hair cells.

The cochlea performs a Fourier analysis of complex sounds into their component frequencies. The sound of [i] as in "see" results in many traveling waves

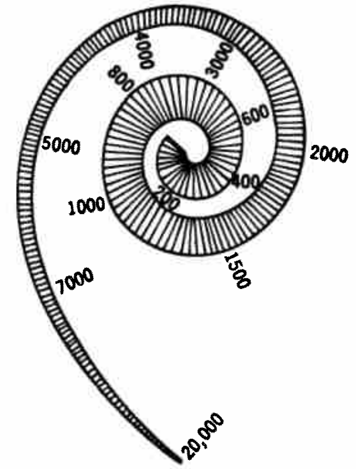

FIGURE 9.6 The width of the basilar membrane (somewhat exaggerated) as it approaches the apex. The approximate positions of maximum amplitude of vibration in response to tones of different frequency are also indicated.

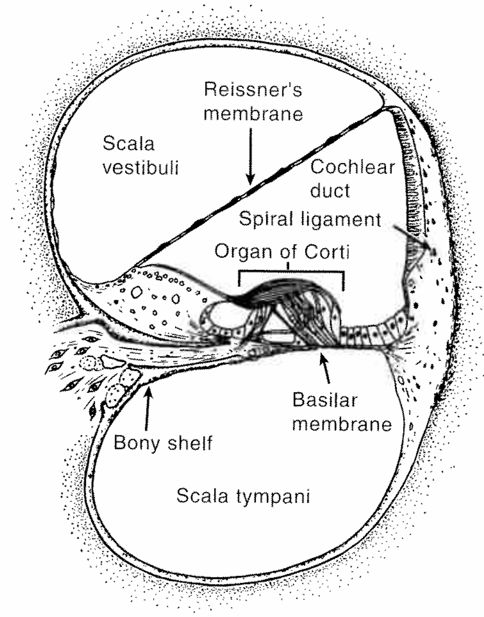

FIGURE 9.7 Cross section through the cochlea showing the scala vestibuli, the scala tympani, and the cochlear duct. The organ of Corti lies within the cochlear duct. (Modified with permission from Denes, P. B., and Pinson, E. N., *The Speech Chain.* New York: Doubleday, 1963.)

moving along the basilar membrane with at least two maxima of displacement: one near the apex for the lower resonance and one near the base of the cochlea for the higher resonance. If the speaker were to say [si] "see," the membrane displacement would initially be maximum even closer to the base of the cochlea for the high frequency [s]. Also, the traveling waves would be aperiodic during [s] and become periodic during the phonated part of the word. Both the traveling wave theory and the description of the stiffness gradient of the basilar membrane are the result of the work of the late Georg von Békésy.

Frequency information is extracted from the signal by the combined factors of the place of stimulation (which activates the sensory nerve fibers at a particular location along the basilar membrane–the place theory) and also by timing of impulses along the nerve fibers. Ernest Glen Wever theorized that at low frequencies the displacement is not sharp enough to distinguish the frequencies by place; rather, they may be signaled by the number of cycles per second translated into a corresponding number of clusters of nerve impulses per second (Fig. 9.8). At high frequencies, place is probably important for indicating frequency, because neurons cannot fire at high frequencies. Another possibility is Wever's volley theory, by which several neurons would cooperate in the neural transmission of high frequencies (Fig. 9.9). The coding of intensity may well be as complicated as frequency coding. It is thought, though, that it is primarily transmitted by relative rate of nerve impulse spikes, as it is throughout the body.

The Auditory Nerve

There are 30,000 nerve fibers serving the cochlea, each fiber coming from a few hair cells and each hair cell exciting several

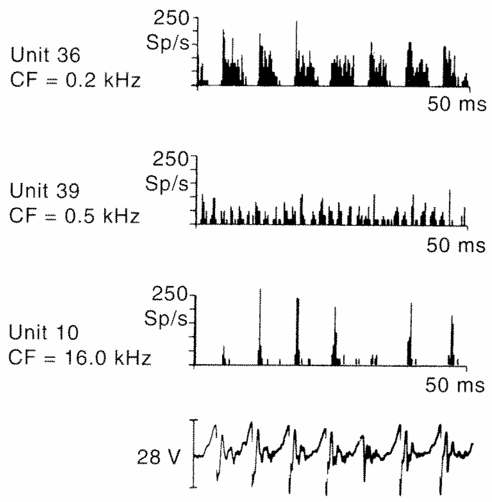

FIGURE 9.8 Responses of single neurons of the auditory nerve of a cat to a presentation of a segment of the vowel [æ]. Bottom display shows the acoustic signal. The three upper displays show the number of spikes per second (Sp/s) for three neural units. Although different units have different firing frequencies, they maintain a fixed temporal relationship to the signal. (Modified with permission from Kiang, N. Y. S., and Moxon, E. C., *J. Acoust. Soc. Am.* *55,* 1974, 620–630.)

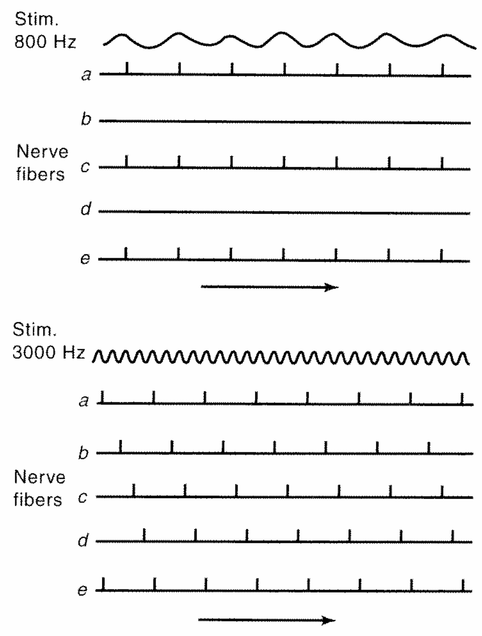

FIGURE 9.9 Wever's volley principle. At low frequencies (800 Hz), individual neurons can fire for each cycle of the stimulus, but at high frequencies (3,000 Hz), the organized firing of groups of neurons indicates frequency. (Modified with permission from Osgood, C. E., *Method and Theory in Experimental Psychology.* Oxford, England: Oxford University Press, 1953.)

nerve fibers. They form a bundle known as the auditory nerve, or eighth cranial nerve. Another branch of the eighth cranial nerve relays information from the semicircular canals. When the nerve fibers are excited by the stimulation of the hair cells, the frequency analysis performed by the organ of Corti is further refined because of lateral inhibition: when a certain place along the basilar membrane is maximally stimulated, to sharpen the effect, surrounding cells and nerve fibers are inhibited in their response.

The eighth cranial nerve does not have far to go between the cochlea and the temporal lobe of the brain. It exits the temporal bone by the internal auditory meatus and enters the brainstem where the medulla meets the pons. In the brainstem, most nerve fibers from each ear decussate (cross) to the contralateral pathway. At that point, comparisons can be made between signals from each ear to localize a sound source. It is thought that eighth cranial nerve fibers in the brainstem may be specialized to detect certain auditory features. Such a specialization would be useful in detecting distinctions important to speech processing. From the brainstem, the eighth cranial nerve courses to the midbrain and then to the temporal lobe. Along the way, fibers go off to the cerebellum and to a network of the brainstem that acts to focus attention. Motor fibers of the auditory nerve also descend to control the sensitivity of the cochlea.

When signals arrive at the auditory cortex of the temporal lobe, the impulses have

CLINICAL NOTE

It is important for the clinician to understand and maintain the distinction between hearing and perception, especially with regard to speech. We can appreciate the difference between speech audition and speech perception when we compare the effects of deafness with those of developmental aphasia. When a child is born deaf or hard of hearing, the difficulty in learning language is based on dysfunction of the peripheral hearing mechanism. Without the auditory input of a normal-hearing child, the hearing-impaired child lacks the normal basis of information on which to base his or her own vocalizations and to match them, eventually, to a reasonable facsimile of the speech of others.

In contrast, when a child is born with brain damage that specifically interferes with speech perception, the child has normal hearing but is unable to interpret the sounds in any linguistically useful way. Although several syndromes are called by such terms as developmental aphasia or auditory agnosia, a common difficulty seems to lie in the processes leading to the discrimination and identification of speech sounds rather than in the auditory processes themselves. We can all, of course, experience the difference between hearing speech and perceiving the linguistic message it carries if we find ourselves in a place where the language being spoken is one we do not know. We hear exactly the same acoustic signal as the native speaker of the unknown language, but we are not sure how the sounds differ from those of our language, nor can we tell where phrases and words begin or end or, of course, what those words mean.

Given the differences between hearing sounds and perceiving speech and language, it is obvious that the clinical strategies used to ameliorate dysfunctions of the peripheral hearing mechanism must be largely different from those used to treat dysfunctions in the central nervous system. The clinician must always be aware of which type of disorder is presented by a patient so that an appropriate clinical regimen can be planned and administered.

preserved the place–frequency arrangement of the basilar membrane. In a three-dimensional display along the superior part of the temporal lobe, low-frequency stimulation near the apex of the cochlea excites the layers of cortical cells along the lateral part of the primary auditory area, whereas high-frequency stimulation of the base of the cochlea is registered in columns of cells within the lateral fissure. This topographic representation is present in both temporal lobes. Most of the contribution to each lobe comes from the contralateral ear. Thus, hearing is accomplished, but the signals must be processed further if the hearer is to understand what is heard. Cortical processing of speech sounds will be discussed further in Chapter 11 when we consider the neurophysiology of speech perception.

REFERENCES

General Readings

Denes, P., and Pinson, E. N., *The Speech Chain,* 2nd ed. New York: W. H. Freeman & Co., 1963.

Fant, G., Descriptive Analysis of the Acoustic Aspects of Speech. *Logos 5,* 1962, 3–17.

Kent, R. D., Atal, B. S., and Miller, J. L. (Eds.), *Papers in Speech Communication: Speech Production.* Woodbury, NY: Acoustical Society of America, 1991.

Pickett, J. M., *The Acoustics of Speech Communication.* Boston: Allyn and Bacon, 1999.