INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O. B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460392 - I

SMR/302-53

COLLEGE ON NEUROPHYSICS:
"DEVELOPMENT AND ORGANIZATION OF THE BRAIN"
7 November - 2 December 1988

"The Auditory System: Acoustics, Psychoacoustics and the Periphery"

Edwin W. RUBEL
University of Washington
Department of Otolaryngology
Seattle, WA
USA

CHAPTER 17

THE AUDITORY SYSTEM: ACOUSTICS, PSYCHOACOUSTICS AND THE PERIPHERY

Robert A. Dobie and Edwin W Rubel

INTRODUCTION

The sense of hearing, along with vision and olfaction, provides vertebrates with information about events occurring at a distance. In humans and in other animals with a poorly developed sense of smell, hearing is the only sense which constantly monitors the entire environment, receiving information from all directions. Even when an animal is asleep or directing its attention elsewhere, it can still respond to sounds (e.g., the sleeping mother is aware of her baby's slightest movement). The sense of hearing, thus, serves as a primitive alarm or warning system.

The vertebrate ear probably evolved from the lateral line organ (still seen in fishes and larval amphibians), which is specialized for detecting mechanical disturbances in a fluid environment. In the lateral line organ, highly specialized mechanoreceptors called hair cells transduce displacement of the fluid surrounding them to graded changes in membrane potential, and ultimately to excitation of afferent neurons. The actual subcellular elements responsible for the transduction are a group of tiny actin-filled cilia called stereocilia, which protrude as a tuft from the top of the cell. Similar hair cells and subcellular elements are also responsible for transduction in the auditory end organs of the vertebrate inner ear and the vestibular end organs (specialized for sensing head position and movement; see Chapter 27). In the auditory and vestibular organs, however, the hair cells detect the motion of fluids within the temporal bone; these fluids are in turn driven by sound

waves and head movements, respectively. Together, these systems are often referred to as the acoustico-lateralis system (Hudspeth, 1983).

In non-mammalian vertebrates, the auditory end organ has undergone major evolutionary changes related to the transition from an aquatic to a terrestial environment (Popper and Fay, 1980). The mammalian ear, by contrast, is surprisingly constant across species, displaying primarily quantitative rather than qualitative variations related to the specific behavioral and ecological needs of different species. For example, the mechanical characteristics of mammalian middle ear structures are specialized to enhance either high-frequency or low-frequency hearing (Webster and Webster, 1984), while both the middle and inner ears of certain bats are constructed to provide exceptional sensitivity and resolution at the ultra-high frequency of the bats' sonar vocalizations (Bruns, 1980).

Although many mammals and birds use the auditory sense to receive communication signals from other members of their species, the ears of such animals do not appear to differ in structure or function from those of other, less vocal, species. However, there is usually a good match between the sound frequencies to which an animal is most sensitive and those frequencies which carry communication information from the animal's vocalizations.

The role of hearing in intraspecies communication reaches its zenith in humans. Loss of the ability to understand speech is by far the most serious consequence of acquired deafness. In congenital deafness the ability to learn spoken language is severely diminished, so that the

vast majority of congenitally deaf children never acquire intelligible speech. Despite the crucial role of the hearing sense in the acquisition and use of speech and language, the human ear, up to and including the auditory nerve, does not appear to differ in any important way from the general plan of the mammalian ear.

The structure and function of the adult mammalian ear is discussed in this chapter. The next chapter considers how the brain processes the afferent signals provided by the auditory nerve.

ANATOMICAL OVERVIEW

Figure 1 shows the location and structure of the mammalian ear. It is divided for convenience of description into outer, middle, and inner parts. The outer ear includes the pinna (or auricle) and ear canal. The pinna is a skin-covered cartilaginous flap, often intricately shaped, and possesses muscles that permit it to be moved to a greater or lesser extent in different species. The ear canal, which is also lined by skin, is a roughly tubular structure with a cartilaginous outer portion and bony inner portion. The ear canal terminates at the tympanic membrane (or eardrum), the boundary between the outer ear and middle ear. The tympanic membrane is a trilaminar structure. Its outer lining is skin, its fibrous middle lamina lends rigidity to the drum membrane, and its inner layer is respiratory mucosa continuous with the lining of the middle ear.

The air-containing middle ear is connected to the nasopharynx by the eustachian tube, which provides aeration and is responsible for equilibration of pressure on the two sides of the eardrum. The middle ear is also continuous with additional air-filled, mucous membrane-lined spaces in the temporal bone (bullae or mastoid air cells, depending on the species). Although these spaces act as resonant cavities to very slightly affect the transmission of certain frequencies through the middle ear, their functional significance is otherwise obscure. A chain of three delicately suspended bones (ossicles) connects the tympanic membrane to the oval window. In order, they are named the malleus (hammer), incus (anvil), and stapes (stirrup); the stapes through its footplate is in direct contact with the fluid of the inner ear.

Sound entering the ear canal sets the tympanic membrane into vibration. This vibration is in turn mechanically coupled by the ossicular chain to the cochlear fluids. Two middle ear muscles, the tensor tympani and the stapedius, attach to the malleus and stapes respectively; the stapedius in particular seems to play an active role in the regulation of sound transmission through the ossicular chain.

The inner ear contains the cochlea (also called the auditory labyrinth), the spiral end-organ of hearing. The inner ear also houses the vestibular labyrinth, a group of end organs specialized for sensation of head position and movement (see Chapter 27). The cochlea can be thought of as a tunnel (about 1 inch long in humans) spiraling through the dense temporal bone. Figure 2 is a section along the axis of the cochlear spiral; the apex of the cochlea is shown at the top and the base

at the bottom. Although the axis of the spiral actually points anterolaterally rather than upward, the nomenclature for the cochlea has been based on diagrams like Fig. 2; therefore, the terms "up" and "down" describe the dimension parallel to the cochlear axis. The terms "basal" and "apical" describe position along the cochlear spiral.

At every point along its 2-1/2 turns, the cochlea has a roughly circular cross-section; it is divided by two membranes into three fluid-filled compartments, each of which runs the full length of the cochlea. The upper-most compartment, called the scala vestibuli, contains perilymph, a fluid resembling cerebrospinal fluid. At the basal end of the cochlear spiral, the scala vestibuli opens into a space called the vestibule, whose fluid is set into motion by the movements of the stapes footplate. The lower-most of the three cochlear compartments, the scala tympani, also contains perilymph and is in fact continuous with the scala vestibuli through an opening at the apex of the cochlea called the helicotrema. At the basal end of the cochlea, the scala tympani terminates in a membranous partition (the round window), which is in contact with the air space of the middle ear. The middle comparment, the scala media (or cochlear duct), is separated from the other two by the basilar membrane below and by Reissner's membrane above. The scala media contains endolymph, a unique extracellular fluid with a high potassium concentration and a high positive electrical potential relative to perilymph.

Resting on the basilar membrane along the entire length of the cochlear duct is the organ of Corti, which contains the auditory neuroepithelium. It is this organ which contains the receptor cells and supporting structures that transduce mechanical disturbances of the inner ear fluids into electrical signals in the auditory nerve. Viewed from above (i.e., from the scala vestibuli), the organ of Corti has three rows of outer hair cells and a single row of inner hair cells (about 15,000 total hair cells in the human cochlea). The afferent auditory nerve fibers run in the axis (modiolus) of the cochlea. Their cell bodies are located in the spiral ganglion within a bony ledge called the osseous spiral lamina. In the cat, but probably other mammals as well, ninety to 95% of the 30,000 afferent neurons innervate inner hair cells. The auditory nerve also contains about 1000 efferent fibers, most of which innervate outer hair cells.

ACOUSTICS

Sound can be defined in a physical sense as a propagated change of density and pressure in an elastic medium. From some vibrating source (e.g., a tuning fork, a loudspeaker, or vocal cords) alternating waves of condensation and rarefaction spread out like ripples in a body of water. A sound source which oscillates back and forth like a tuning fork creates a periodic sound. If a tuning fork is struck, it oscillates at a single repetition rate or frequency. The ensuing variations in sound pressure can be described by a sine (or cosine) function:

$$p = A \sin(2 f t)$$

where p = instantaneous sound pressure, A = maximum sound pressure, f = the frequency of oscillation, and t = time. The frequency of such a sine wave or pure tone is the number of cycles/second or Hertz (Hz). The period of a sine wave (the time between successive condensations or rarefactions) is simply the inverse of its frequency (i.e., seconds/cycle) and is measured in seconds, or more often milliseconds (msec). For example, a 1000 cycle/second tone is said to have a frequency of 1000 Hz (1 kHz) and a period of 1 msec; the musical note middle C has a frequency of 256 Hz and a period of about 3.9 msec. The phase of a sine wave specifies its relationship in time to some temporal reference point or in some cases to other periodic sounds. If the frequency, phase, and intensity (see below) of a sine wave are given, one has completely and adequately described that sound; this is called a frequency domain description. Alternatively, one can completely describe a sound by specifying the instantaneous values of sound pressure as a function of time for as many cycles of the sound as desired; the latter is called a time domain representation.

Real sounds of biological interest are never pure tones. Some complex sounds (like sustained notes from a musical instrument, or speech sounds such as vowels) are periodic and can be described as the sums of many pure tones whose frequencies are integer multiples (called harmonics) of a single fundamental frequency. For example, an organ note with a fundamental frequency of 100 Hz will have harmonics (also called overtones) of 200, 300, 400 Hz, and so on. The relative strength of

these overtones determines the tonal quality. Most real sounds, however, are not periodic. Although it is easiest to think of aperiodic sounds such as handclaps or consonants in the time domain (pressure as a function of time), these sounds can also be specified as sums of sine waves (i.e., in the frequency domain).

Intensity

As described earlier, the "amplitude" of a sound is usually specified by its pressure. It is important to realize, however, that sound waves consist of propagated disturbances not only of sound pressure but also of the molecules of the conducting medium (The importance of this fact will be apparent in our subsequent discussion of middle ear function). The intensity of sound is defined as the average rate of flow of sound energy across a given area perpendicular to the direction of propagation of the sound. It is therefore equal to the product of sound pressure and a quantity called volume velocity (the product of the average particle velocity and the cross-sectional area). By analogy to the flow of electricity, intensity is similar to electrical power and indeed is expressed in units of acoustic watts. Sound pressure is analogous to electrical potential or voltage, while volume velocity is analogous to electrical current.

Impedance

Acoustical impedance is defined as the ratio of sound pressure to volume velocity and represents the opposition to movement offered by an acoustical system (acoustical impedance is similar to electrical impedance, which is equal to the ratio of voltage to current). Sound is

propagated in media which, like electrical conducting materials, have characteristic impedances. In an ideal medium with no boundaries, acoustical impedance is defined by density and elasticity alone. In air, for example, the molecules are far apart and compressible; relatively little sound pressure is required to cause high particle velocities. Because the ratio of pressure to volume velocity is low, air can thus be characterized acoustically as a low-impedance medium. Water, on the other hand, is a dense, high-impedance medium in which the ratio of pressure to volume velocity is high. The impedance of a conducting medium is independent of the frequency of sound being propagated; it is thus analogous to the resistive component of electrical impedance. "Bounded" media, like the cochlear fluids, have impedances which are higher than would be predicted by these parameters, because they are less compressible at their boundaries.

The above discussion refers to sound propagated in fluid (including gaseous) media; the situation is different for solid materials. Although sound waves of pressure and density can be propagated in solids, sound also can cause a solid object to vibrate or undergo deformation as a whole. For suitably suspended solid objects like the tympanic membrane and ossicles, these mechanical responses predominate. The opposition offered by such systems to applied sound is mechanical impedance, the ratio of sound pressure to the velocity of the system. Unlike acoustical impedance, the mass and stiffness of a system cause its mechanical impedance to vary with the frequency of sound.

If a system is massive, it has an increased opposition to movement (impedance), especially for high frequencies; low-frequency sounds are less impeded because their slower pressure changes allow more time to overcome inertia. Conversely, if a system is stiff, it preferentially impedes low-frequency sounds, because they require greater displacement for the same amount of sound energy, and the opposition offered by a stiff system (like a spring) is proportional to its displacement from a rest position. The relation of mechanical impedance to these factors is summarized in the following equation:

$$Z_f = R^2 + (2\ fM - \frac{S}{2\ f})^2$$

where Z_f = impedance for a given frequency (f), R = the resistive or frictional (non-frequency-dependent) component of impedance, M = mass, and S = stiffness.

Decibel

The loudest sound to which one can listen without discomfort is more than 1 trillion times (10^{12}) as intense as the softest sound which is audible. Because of this extremely wide operating (or dynamic) range, it is necessary to use a logarithmic measure for sound intensity. This logarithmic measure is called the decibel (dB), after Alexander Graham Bell, and is defined as:

$$dB = 10 \log_{10} (I/I_0),$$

Where I = the intensity of the sound of interest, and I_0 = the intensity of some reference sound. Since the ratio between sound intensities of painful and barely-audible sounds is about 10^{12}, the human dynamic

range can be restated as 120 dB [i.e., $10(\log_{10} 10^{12})$]. Since acoustic intensity is proportional to the square of sound pressure and since sound pressure is more conveniently measured than sound intensity or power, the more common formulation for the decibel is:

$$dB = 20 \log_{10} (P/P_0).$$

When using a logarithmic relationship like this it is essential to specify the value of P_0 and to understand that 0 dB does not mean the absence of sound. When dB = 0, $P = P_0$, and when $P < P_0$, dB will have a negative value. Unless otherwise stated, dB values usually refer to the sound pressure level (SPL) standard for which P_0 = 20 micropascals, or 0.0002 dyne/cm^2. Other standards are also used. For example, dB HL (hearing level) measures sound pressure or intensity relative to the normal human threshold of hearing.

Psychoacoustics

To appreciate the contributions and limitations that the peripheral components of the auditory system impose upon an animal's hearing, it is essential to first understand the acoustic processing capabilities of the entire organism. The study of the relationship of acoustic stimuli to behavior, specifically the abilities to detect, discriminate, and identify sounds, is called psychoacoustics (a branch of psychophysics). A brief review of human psychoacoustics is presented so that the reader can better understand the relevance of the subsequent discussions of auditory physiology.

The levels of sound pressure required by a human to hear different frequencies (the minimum audibility curve) are shown in Figure 3. For

any frequency, a range of intensities can be discriminated up to a level that is painful (threshold of feeling). As previously noted, the dynamic range for the best frequencies (2-4 kHz) is extremely large, about 120 dB. Within much of this range, intensity differences of one dB or less can be reliably detected, both for tones and for complex sounds. Frequency differences as small as 2-3 Hz can be discriminated for frequencies up to about 3 kHz. Not only can the auditory system discriminate very small changes in these parameters, it can do so across a whole range of stimulus frequencies and intensities. For example, small frequency and intensity changes can be detected for both very soft and very loud sounds. These capabilities should be kept in mind during our subsequent discussion of coding in the auditory nerve.

Identification of speech sounds is one of our most important abilities. Figure 4 shows the power contained at different frequencies for three vowels. The differences among the vowels, like the differences among the tones of different musical instruments playing the same note, is in the distribution and relative intensity of the different overtones, which are shaped by the positions of the tongue, palate, cheeks, and lips. The maxima in these curves of intensity as a function of frequency are called formants. Vowels are periodic sounds which can be indefinitely prolonged and have formants that do not change over the duration of the vowel. Consonants, on the other hand, are dynamic sounds which exhibit rapid changes in their formant structure over time. One can easily discriminate one vowel from another across a wide dynamic range from very soft to very loud sounds, and in the presence of white noise

(random sound pressure fluctuations containing energy across a wide band of frequencies) nearly as intense as the vowels themselves. Vowels cannot be recognized simply by frequency or intensity but rather by patterns of intensity as a function of frequency. Although vowels are the simplest speech sounds to discriminate, we will see that it is not simple to understand how they can be recognized on the basis of cochlear physiology alone.

SOUND TRANSMISSION TO THE INNER EAR

The outer and middle ear are essentially passive, linear mechanical systems and the way in which they transmit sounds of different frequencies are predicted reasonably well by their physical properties (mass, stiffness, etc.). The combined properties of the outer and middle ear predict quite well the range of frequencies to which a given animal will be most sensitive (Rosowski et al., 1986).

Outer Ear

The pinnas of some animals have considerable sound-collecting capabilities which are facilitated by strong muscular control. They can be moved to "focus" hearing in a particular direction, achieving much the same effect as a cupped hand behind the ear does for humans. The pinna is also a complex sound baffle which accentuates or attenuates certain frequencies depending on the angle at which the sound waves are approaching the head. In man, the pinna is necessary for sound localiza-

tion in the vertical plane, and for some sound localization in the horizontal plane when only one ear is functional.

The external ear canal acts essentially like a rigid tube closed at one end. It therefore resonates at a frequency whose wave length is four times the length of the ear canal (the dependence of the resonant frequency of a closed tube on its length is easily demonstrated by blowing over the lip of a bottle filled with different amounts of water). The resonant frequency of the adult human ear canal is about 3 kHz.

The combined effect of the pinna, the external ear canal and the head on sound reaching the eardrum is shown in Figure 5. The graph shows the ratio of the sound pressure measured by a small microphone placed near the tympanic membrane to the sound pressure measured by a microphone placed in a free field in front of the speaker. A 15 dB improvement in pressure occurs between 2.5 and 4.0 kHz, largely due to the characteristics of the ear canal.

Middle Ear

As described earlier, air and water have very different specific acoustic impedances. Since sounds of interest to terrestial vertebrates travel in air while the hair cells are bathed in an aqueous medium, sound must traverse an air-water interface at the boundary of the middle ear and inner ear. Sound traveling in air has insufficient pressure to displace the densely packed water molecules. On the other hand, sound traveling from water to air has insufficient volume velocity to adequately displace the air molecules. The ratio of specific impedances of water and air is approximately 10,000. Because the transmission ratio across

any acoustic interface is approximately 4 divided by the impedance ratio, only 0.04% of the sound energy will be transmitted across an air- water interface, in either direction. This impedance mismatch, which causes more than 99.9% of the acoustic energy to be reflected rather than transmitted, produces a transmission loss of approximately 34 decibels.

Fortunately, nature solved this problem by creating a middle ear which acts as an impedance-matching transformer, converting the low-pressure/high-volume velocity excursions of sound in air to high-pressure/low-volume velocity excursions in the perilymphatic fluid of the inner ear. Impedance matching is accomplished in two ways (Fig. 6). First, the handle of the malleus is slightly longer than the long process of the incus, resulting in a lever ratio of approximately 1.3. Far more importantly, the effective area of the tympanic membrane is about 17 times larger (in humans) than the area of the stapes footplate. The equivalent increase in sound pressure from the eardrum to the footplate is the product of these two ratios (1.3 x 17), a 22-fold (or 28 dB) increase. The pressure increase is accompanied by an equivalent decrease in volume velocity, so that the total energy or power remains constant (this must, of course, be true for any passive system according to the principle of conservation of energy). Consequently, most of the energy which would have been lost or reflected back at the eardrum in fact crosses into the cochlear fluids.

This description is over-simplified but captures the function of what the middle ear for those frequencies at which the mass and stiffness

of the eardrum and ossicular chain are negligible. However, the picture is complicated by several factors. Above 2000 Hz, the tympanic membrane does not move as a unit and thus transmits energy less efficiently. In addition, the ossicular mass begins to impair transmission, and also a small amount of energy is dissipated by loose coupling between the individual ossicles. At low frequencies, the stiffness of the eardrum and ossicular chain can impair transmission. For example, unequal air pressure across the tympanic membrane, due to eustachian tube blockage and air absorption in the middle ear, can stiffen the tympanic membrane. Resonances of the middle ear cavity and of the mastoid and bulla cavities can also affect middle ear sound transmission.

Defects occurring anywhere from the outer ear to the stapes cause what are called conductive hearing losses since they impair sound transmission into the inner ear. Examples of such problems would be blockage of the ear canal with cerumen (ear wax), perforations in the tympanic membrane, ear infections that fill the middle ear with fluid, disruption of the joints between ossicles, or impairment of ossicular motion. Otosclerosis, for example, is a common hereditary hearing disorder in which the stapes footplate becomes immobilized by bone growth bridging the gap from the footplate to the surrounding bone of the otic capsule. All of these forms of conductive hearing loss are potentially correctable by medical or surgical means. In contrast, hearing loss arising from disorders of the cochlea or auditory nerve (sensorineural hearing loss) is usually irreversible.

INNER EAR

Cochlear Duct

Figure 7 shows a schematic cross-section of the cochlear duct. As described previously, the scala vestibuli and scala tympani contain perilymph, while the scala media contains endolymph. The high potassium concentration of the scala media is maintained by the stria vascularis which actively pumps K^+ ions into and Na^+ out of the scala media. Reissner's membrane, which separates the scala vestibuli and scala media, is impermeable to charged ions (but not to water) and protects the unique ionic and electrical composition of the endolymph. In contrast, the semi-permeable basilar membrane permits perilymph from the scala tympani to bathe the bodies of the hair cells and supporting cells in the organ of Corti. The tectorial membrane is also permeable, so the stereocilia at the tops of the hair cells are bathed in endolymph. The impermeable barrier between endolymph and the scala tympani perilymph is at the cuticular plate, at the level of the tops of the hair cells (Hunter-Duvar et al., 1981). Since the hair cells have a resting potential of approximately -70 mV and the endolymph is maintained at +80 to +90 mV by the high K^+ concentration, a very large net electrochemical potential exists across the apical surface of the hair cell.

The stereocilia of the three rows of outer hair cells are attached to the tectorial membrane. Although gel-like, the tectorial membrane apparently contains abundant collagen as its main protein component, and thus probably possesses some rigidity (Thalmann, 1986). The single row

of inner hair cells apparently is not directly attached to the tectorial membrane (Lim, 1980).

The Traveling Wave

An impulsive sound (like a click) displaces the stapes, creating a pressure gradient between the scala vestibuli and the scala tympani (Fig. 8). The result is a wave of displacement which begins at the basal end of the basilar membrane and travels towards its apex. This "traveling wave" displaces not only the basilar membrane but the entire organ of Corti. These structures move together and the mechanical characteristics of each is important in determining the response of the whole; together, they are referred to as the cochlear partition. The traveling wave begins at a high velocity which decreases exponentially as it travels toward the apex; in the human ear, it traverses the cochlea in about 4-5 msec. It is not a sound wave (the travel time for sound in the cochlea is on the order of 20 sec), but is somewhat analogous to the propagated disturbance that can be created by whipping a rope attached at one end to a wall. Since perilymph and endolymph, like all liquids, are incompressible and the cochlear scalae are housed in rigid bone, any net displacement of the cochlear partition toward the scala tympani (as occurs during condensation) must be accompanied by a compensatory outward movement of the round window. Similarly, when the stapes moves outward (in response to acoustic rarefaction), there is a net displacement of the cochlear partition toward the scala vestibuli, and the round window membrane moves inward.

Frequency Tuning of the Basilar Membrane

For a continuous tone, the picture is more complex (see Fig. 9). Two important points should be noted from Figure 9. First, each moment-to-moment change in sound pressure can be thought of as a separate impulsive sound that sets up its own traveling wave. Thus, a point very near the base of the cochlea may be moving upward (toward the scala vestibuli) in response to the rarefaction phase of the sound, while at the same instant a more apical point is moving downward in response to a condensation phase that occurred earlier. The portions of the cochlea responding to the tone will all be vibrating at the frequency of the tone, but not in phase (synchronized in time), since the traveling wave is delayed by different amounts at different points along the cochlear partition. Figure 9 shows an example of the displacement pattern of the basilar membrane in response to a pure tone at several successive moments in time (solid lines). Second, the amplitude of the vibrations is not the same at each point. The maximum excursions of each point along the basilar membrane can be described by an envelope (dotted line in Fig. 9) whose amplitude increases gradually to a maximum and then drops sharply. For different tones, the maxima occur at different positions. For low frequencies the envelope maxima are near the apex and for progressively higher frequencies the peak amplitude of the traveling wave shifts nearer the base of the cochlea (von Bekesy, 1960).

Complex sounds containing several different frequencies are represented in a similar manner along the cochlea. The high frequency components produce peaks of vibration near the base, lower frequencies excite

more apical regions, and so on. Thus, the cochlea acts like a spectrum analyzer, "dissecting out" the component frequencies of a complex sound to be separately displayed at different places along the basilar membrane.

Each successive point along the cochlear partition is most sensitive to a slightly different frequency (Fig. 10). This is explained primarily by changes in the structure (and impedance) of the basilar membrane along its length. The basilar membrane is narrow near the base and wide near the apex. More significantly, its stiffness decreases 100-fold from base to apex. These mechanical characteristics impart a degree of tuning to the cochlear partition, so that the basal portion of the cochlea (with its hair cells and neural elements) responds best to high frequencies, and the apical part responds best to low frequencies. The principle of an orderly spatial array of responsive elements according to an increasing or decreasing order of frequencies is called the place code or place principle; it is similar in all species of birds and mammals. Moreover, as will be discussed in the next chapter, this orderly representation of frequency is preserved in the projection from the cochlea to the central nervous system and then at each successive level of the auditory pathways where it is referred to as tonotopic organization. Along the cochlear partition, the distance from the apex to the position that responds maximally to a particular frequency is proportional to the logarithm of that frequency.

The place principle and the properties of the traveling wave were first established by the elegant experiments of Georg von Bekesy (1960),

for which he received a Nobel prize. His experiments were carried out on cadaver ears with relatively primitive instrumentation consisting of a microscope and stroboscopic illumination. In order to see the movements of the cochlear partition, he used sounds as loud as 130-140 dB SPL. Only recently have two new techniques emerged to measure the tiny excursions of the basilar membrane in response to near-threshold sounds. The Mossbauer technique measures Doppler shifts in radiation emitted by a gamma source placed on the basilar membrane, while laser interferometry uses a mirror (10^{-8} gm) placed on the membrane; the latter technique can measure movements as small as 10^{-9} cm. These new techniques have shown much sharper tuning than initially described by von Bekesy.

A convenient way to describe the sharpness of tuning shown at any locus within the auditory system is to plot the minimum sound intensity required to obtain some criterion response at each frequency. Figure 11 shows the sound intensity required for a criterion displacement (3.5 A) and velocity (0.04 mm/sec) for a point on the basilar membrane near the cochlear base. These tuning curves from the live guinea pig (and similar curves from cats) show that the basilar membrane is exquisitely sharply tuned with high-frequency slopes exceeding 100 dB/octave (Khanna and Leonard, 1982; Sellick et al., 1982). Active processes seem to be involved in the maintenance of this tuning, as its sharpness is considerably degraded by anoxia.

Two properties of the inner ear limit the range of audible frequencies. Very low frequencies (below about 20 Hz in man) are inaudible because the slow pressure changes of these sounds are transmitted from

the scala vestibuli to the scala tympani through the helicotrema. This small opening between the two perilymphatic scalae serves primarily to protect the cochlear duct from excessive displacement during very slow or even static middle ear pressure changes, such as occur during changes in atmospheric pressure, nose-blowing, etc. Very high frequencies (above about 20 kHz in man) are inaudible because even the most basal portion of the basilar membrane is not stiff enough to respond to them. It is important to remember that like the middle ear, mechanical properties of the inner ear are differentially specialized in different species. Some birds and mammals hear ultra-low frequencies (down to .1 Hz) while others, such as some bats, hear sounds in the ultrasonic range (e.g. up to 120 kHz).

It has generally been assumed that the place code is fixed throughout an organism's lifetime. However, recent experiments suggest that during development there is a systematic shift in the place code (Rubel et al., 1984). Early in development, the basal part of the cochlea is the first to mature, but it responds optimally to mid-range rather than high frequencies. As the inner ear matures the base gradually encodes higher frequencies, and the best locations for middle or low frequencies shift toward the apex. These developmental findings suggest that the place code is not immutable even in the life of a single individual and suggests the possibility that other conditions such as aging or pathology may alter its organization.

Frequency Tuning by Hair Cells

Although the mechanical tuning properties of the cochlear partition in mammals appear sharp enough to account for the tuning described later for auditory nerve fibers and for behavioral responses, there is emerging evidence that the hair cells may also contribute to tuning. The stiffness gradient observed in cadaver ears explains only rather coarse tuning, while living hair cells are needed for the sharp tuning curves seen in Fig. 11. In some reptiles, amphibians, and fish, the hair cells themselves may be the only tuned elements. In alligator lizards, for example, the basilar membrane moves as a unit; there is no traveling wave and different points on the membrane do not respond maximally to different frequencies (Peake and Ling, 1980). Instead, the height of the stereocilia bundles varies systematically with the position of individual hair cells along the basilar membrane (Mulroy, 1974), and the preferred frequency of individual cells depends on the heights of the ciliary bundles (cells responding to low frequencies have long stereocilia while cells tuned to higher frequencies have shorter stereocilia) (Frischkopf and DeRosier, 1983). In addition, both bullfrogs and turtles have hair cells whose membranes exhibit a frequency-dependent electrical resonance, which appears to contribute to the acoustic tuning of the hair cells (Crawford and Fettiplace, 1980; Lewis and Hudspeth, 1983).

The hair cells of mammals and birds also demonstrate morphological features that are correlated with tuning. In chicks, the length, diameter, and number of stereocilia per cell all vary systematically from basal to apical locations along the basilar membrane (Tilney and

Saunders, 1982). Stereociliary stiffness (and presumably responsiveness to high tones) decreases from base to apex in guinea pigs (Strelioff and Flock, 1982) The relationship between height of the stereociliary bundles and best frequency for a given cochlear region has also been confirmed in humans (Lim, 1980). Thus, it seems likely that, in all vertebrate species, some tuning is contributed by the mechanical and/or electrical characteristics of individual hair cells.

Transduction and Synaptic Transmission

The traveling waves along the cochlear partition must somehow activate the hair cells. In the current theory, the basilar membrane and tectorial membrane can be considered as being hinged at different points on the medial wall of the cochlea; consequently, either upward or downward displacements of the cochler partition create a shearing force that tends to bend the hair cell stereocilia bundle at the apical hair cell surface (Fig. 12). (Since the inner hair cells are not attached to the tectorial membrane, their stereocilia must be deflected by subtectorial fluid currents rather than directly by the tectorial membrane.) The stereocilia bundles appear to pivot in a relatively rigid fashion at their points of attachment to the apical surface of the hair cell. This motion opens and closes ionic channels which appear to be located near the tips of the stereocilia (Hudspeth, 1983; see Fig. 13). Recall that the stereocilia are exposed to the large positive endolymphatic potential, while the intracellular potential is of course negative. This large potential difference constitutes a "battery" (first described by Davis, 1958) which drives potassium ions into the cell (unopposed by any

concentration gradient; see Chapter 5) and depolarizes it. Depolarization of the hair cell membrane causes voltage-dependent calcium channels at the base of the hair cell to open. Calcium entry, in turn, presumably initiates fusion of synaptic vesicles with the synaptic specialization at the base of the hair cell. Neurotransmitter release will then effect spike initiation in the afferent neuron. The nature of the afferent transmitter that is released from the base of the hair cell is as yet undetermined. Glutamate and aspartate appear to be the best candidates (Bobbin, 1979), but many doubts remain (Drescher et al., 1983).

Sound-evoked intracellular potentials from hair cells are as highly tuned as cochlear partition displacement tuning curves (Russell and Sellick, 1978; Fettiplace and Crawford, 1978, Goodman et al., 1982). In response to continuous tones, intracellular responses have two components. An AC (alternating current) component faithfully follows the oscillations of the stimulating tone while a DC component contributes a net depolarization of the hair cell membrane. Presumably the DC component occurs because the stereocilia bundles display a non-linear behavior that produces more potential change when the stereocilia are deflected toward the tallest row than when they are deflected toward the shortest row (Hudspeth and Corey, 1977). A net depolarization, of course, enhances the probability that neurotransmitter will be released at the hair cell synapse.

The Role of Outer Hair Cells

Because they are present in all vertebrates and receive almost all of the afferent innervation of the cochlea, inner hair cells are

considered to be the "true" (or at least the predominant) sensory receptors in the cochlea. Outer hair cells, on the other hand, receive scant afferent innervation. Although most terrestial vertebrates have two types of hair cell, the difference in their afferent innervation is clearest in mammals.

It has been known for many years that exposures to excessively loud sounds or ototoxic drugs can produce widespread loss of outer hair cells with no loss of inner hair cells and afferent neurons. Animals so treated exhibit 40-50 dB hearing losses (elevations of response thresholds) for certain frequencies (Ryan and Dallos, 1975). Initially, this was thought to mean that very soft sounds were detected via stimulation of the outer hair cells. It now seems more likely that outer hair cells are needed to tune the cochlear partition, which in turn permits stimulation of inner hair cells by very soft sounds. In support of this idea, the sound intensity required to produce a criterion amount of basilar membrane movement (10^{-8} cm) increases with increasing damage to outer hair cells (Leonard and Khanna, 1984).

Recent evidence suggests that the outer hair cells may be viewed as primarily effector or motor structures. Isolated guinea pig outer hair cells contain contractile proteins (actin, myosin) and shorten in response to electrical depolarization, potassium, calcium and ATP, and acetylcholine (Brownell, 1983; Gitter et al., 1986; Flock et al., 1986). Since the outer hair cells receive most of the efferent (cholinergic) innervation of the cochlea it is possible that efferent activity regulates outer hair cell length, tension or the mechanical properties of

their stereocilia so as to alter the tuning properties of the cochlear partition. Indeed, electrical stimulation of the cochlear efferents appears to alter the micromechanical properties of the cochlea (Mountain, 1980).

Kim (1986) has postulated an "active bidirectional transduction mechanism" linking the outer hair cells with the remainder of the cochlear partition. Since bending of stereociliary bundles induces graded changes in transmembrane potential (mechanical to electrical), and since electrical stimuli applied to outer hair cells induce motile responses (electrical to mechanical), it is possible that, especially for very soft tones near a cell's best frequency, an active resonance occurs. If true, sound-induced stereociliary deflection would open ionic channels, causing a relatively large depolarization (because of the large potential difference between endolymph and intracellular fluid). This depolarization would in turn cause the hair cell (or its hair bundle) to "push back" by shortening the contractile proteins, thus reinforcing the mechanical response of the cochlear partition.

Near-field (Gross) Cochlear Potentials

In some clinical and experimental situations, it is desirable to monitor the stability and health of the cochlea without entering it. Electrodes placed on the round window or in other locations relatively near the cochlea detect a variety of sound-induced cochlear potentials (Fig. 14). The first of these, the cochlear microphonic (CM, Fig. 14), rather faithfully follows the wave form of the stimulating sound. Since it nearly disappears after lesions which destroy the outer hair cells

while leaving the inner hair cells intact, it is believed to come primarily from the outer hair cells, and to represent a summation of the AC intracellular potentials discussed above ("Transduction"). A second stimulus-related potential is the summating potential (SP, Fig. 14), a DC potential which may be either positive or negative depending on the properties of the stimulus and the recording electrode. It is probably due to summation of the DC or "net" depolarizing hair cell potentials. Finally, compound action potentials (N_1, N_2, Fig. 14) can be recorded which reflect synchronous activation of afferent neurons at the onset of a tone burst or other impulsive sound. Because the traveling wave moves much more quickly near the base of the cochlea, individual hair cells and their afferent neurons respond much more synchronously there than toward the cochlear apex. For this reason, gross potentials recorded outside the cochlea preferentially reflect the activity of synchronously responding hair cells and neurons in the base of the cochlea.

AUDITORY NERVE

The auditory nerve in man contains about 30,000 neurons (Rasmussen, 1940); all afferent fibers have their cell bodies in the spiral ganglion of the cochlea. About 95% have large cell bodies (type I neurons) whose dendrites pass radially (perpendicular to the cochlear axis) to form afferent synapses at the bases of inner hair cells. The remaining 5% (type II neurons) have small cell bodies and thin fibers

which pass radially, then turn to form the outer spiral bundle, and eventually synapse on outer hair cells (Spoendlin, 1979). The great disparity in afferent innervation of the inner and outer hair cells has already been noted. Approximately 20 radial fibers form afferent synapses on each inner hair cell, while the sparse outer spiral fibers each branch to supply 10-60 outer hair cells (Fig. 15). The role of the few afferent fibers supplying outer hair cells is unknown because to date, all auditory nerve fibers from which it has been possible to record sound-evoked spike activity and which were subsequently morphologically traced to their peripheral origins innervated inner hair cells (Robertson, 1984).

Frequency Threshold Curves

Each afferent neuron can code information either by changes in the rate of spike discharge or by changes in the timing of spike discharges, or both. Of course, information can also be coded by the spatial-temporal pattern of excitation of an array of neurons responding to a stimulus. We will first consider the discharge patterns of a single afferent neuron to sinusoidal tones of differing frequency and intensity.

A frequency-threshold curve plots the intensity required to produce a small increase in spike rate (over resting rate) at different frequencies. Typical frequency-threshold curves for auditory neurons are shown in Figure 16 (one is also shown in Fig. 11, along with basilar membrane tuning curves). As might be expected from the discussion of the traveling wave, auditory afferents have frequency-threshold curves whose

best frequencies correspond to the loci that they innervate along the basilar membrane. Like hair cell and basilar membrane tuning curves, they are sharply tuned to exclude frequencies above the fiber's best, or characteristic frequency (CF). For low frequencies, the tuning is less sharp, and auditory nerve fibers respond to sounds well below the CF once stimulus intensity is raised 40-50 dB above CF threshold. Apparently, frequency-threshold curves of auditory nerves directly reflect the tuning present in the inner hair cells to which these afferent neurons connect (Fettiplace and Crawford, 1978; Russell and Sellick, 1978).

Auditory nerve fibers have relatively similar response characteristics. Other than differences in CF, their frequency-threshold curves are more or less alike (although, as seen in Fig. 16, fibers with low CF tend to have somewhat broadly-tuned curves. One important quantitative difference is in their unstimulated firing rate and response thresholds: most auditory nerve fibers have high unstimulated firing rates and respond to very soft sounds, while a few fibers have relatively low spontaneous rates and higher thresholds.

The frequency-threshold curves of auditory nerves are clearly sharp enough to permit the discrimination of very soft (near threshold) tones of only slightly different frequencies. Even for moderate to high intensity tones, which would activate a large fraction of auditory nerve fibers, the stimulating frequency could be deduced from the CF of the most apical fibers responding, because the frequency-threshold curves have such sharp high-frequency slopes. The general notion that the auditory system encodes stimulus frequency according to the spatial position

of responding auditory neurons along the basilar membrane is called the place or rate-place theory, to indicate that the central nervous system analyzes firing rate (or changes in firing rate) according to the place of the fiber (along the cochlear partition) and its CF. As will be seen, this theory has some difficulties accounting for our ability to discriminate differences in complex sounds containing many frequencies, such as vowels.

Suppression. A tone which by itself does not cause any measurable effect on the firing rate of a given auditory nerve fiber may nevertheless interfere with that fiber's response to tones to which it ordinarily does respond. This is most easily illustrated by considering a conventional frequency-threshold curve (filled circles, Fig. 17). The space above the curve indicates combinations of frequency and intensity to which the fiber will respond with an increase in discharge rate, and can be considered to be that fiber's excitatory "response area." Tones in the cross-hatched area (mostly outside the excitatory response area), presented simultaneously with a normally excitatory tone, suppress the response to the latter tone. Such suppression tends to enhance the contrasts among the different frequencies present in a complex sound since higher-intensity components would suppress the response to less-intense components. Unlike the superficially similar phenomenon of lateral inhibition in the visual system, two-tone suppression cannot be explained by neural inhibition (no synaptic circuitry appropriate for the task exists in the cochlea), but rather by a nonlinearity measurable in the mechanical responses of the basilar membrane (Rhode, 1977). Obvious-

ly, this factor must be important in all cochlear responses to complex stimuli, although its effects are almost impossible to predict for stimuli containing more than a few components.

Intensity Coding

For steady-state tones average firing rate increases as a function of stimulus intensity. Mammalian auditory nerve fibers show relatively steep increases in firing rate with sound intensity, and they reach their maximum firing rates (saturation) within about 40 dB of their thresholds. Figure 18 shows a series of such rate-intensity functions for a single auditory nerve fiber when it is stimulated with a tone at its CF (11.3 kHz) or at higher frequencies.

Above the saturation level, a fiber obviously cannot signal a change in stimulus intensity by a change in firing rate. Humans, however, are able to discriminate intensity differences from 0 to 120 dB, although it is clear that no single fiber can encode intensity over that range. An obvious solution would be to have fibers which operate at different stimulus levels; while each might only encode stimulus intensity over a 40 dB range, the ensemble of fibers would be able to encompass a much broader range (such a situation obtains for the rods and cones in the retina; see Fig. 19-18). Although the vast majority of auditory nerve fibers studied have very low thresholds (as in Fig. 18), a small population does have high thresholds and could encode stimulus intensity changes for loud sounds (they also tend to have less steep slopes to their rate-intensity functions). These high-threshold units have been

overlooked in many investigations because they have very low spontaneous rates and are, therefore, somewhat harder to locate (Liberman, 1978).

The wide range of dynamic intensity coding could also be accomplished by spread of excitation recruiting additional afferents. Consider the fiber in Fig. 18. Although its CF is 11.3 kHz, it also responds to 14.5 kHz if the sound is loud enough (i.e., > 50 dB). A loud 14.5 kHz tone, therefore, excites not only fibers whose CFs are 14.5 kHz but also nearby fibers with similar CFs. If, like the fiber in Fig. 18, the firing rate of the recruited neurons is not saturated, the firing rate could encode intensity over a considerable range (in this case, a range of > 50 dB). This mechanism alone, however, cannot solve the dynamic range problem, because dynamic range and intensity discrimination are nearly as good for white noise, which contains all audible frequencies and stimulates the entire cochlea, as for pure tones!

It should be noted that most auditory nerve recordings have been performed in anesthetized animals whose efferent cochlear innervation and acoustic middle ear muscle reflex are depressed. We now show that these systems may further improve the dynamic range and intensity-processing capabilities of the peripheral auditory apparatus.

The Efferent System. Efferent neurons which innervate the sensory cells of the inner ear have cell bodies in the superior olivary complex; each ear receives efferent fibers from both sides of the brainstem. These olivocochlear efferent fibers travel in the inferior vestibular branch of the 8th nerve and only join the auditory nerve just before it enters the cochlea. Most efferent fibers end on the cell bodies of outer

hair cells; a smaller number form axodendritic synapses on the distal processes of afferent fibers innervating the inner hair cells. Efferent fibers respond to sound presented to either ear, and display frequency tuning similar to that seen in afferent auditory neurons (Liberman and Brown, 1986).

Electrical stimulation of the crossed olivocochlear bundle causes about a 10 dB shift in rate-intensity functions of afferent nerve fibers (Wiederhold and Kiang, 1970). In other words, when the efferent fibers are activated, the afferent fiber does not begin to change its firing rate until the stimulus tone is about 10 dB more intense than the fiber's normal threshold. Similarly, the intensity saturation level is 10 dB higher during efferent activation. As mentioned above, efferent activity may decrease sensitivity by changing the mechanical properties of the cochlea. If sound-evoked activity of the efferent system can cause similar changes (this has not been demonstrated), it would clearly provide a mechanism to shift the operating range of individual neurons. The central connections of the efferent system will be considered further in Chapter 18.

Acoustic Reflex. Loud sounds presented to either ear elicit a brisk contraction of both middle ear muscles (primarily the stapedius in man), which tenses the ossicular chain. This acoustic reflex reduces sound transmission through the middle ear and, like the efferent responses described above, indirectly reduces the sensitivity of individual auditory fibers. As might be expected from earlier discussions of middle ear mechanical factors, the reduction in sensitivity is greatest for low

frequencies; in fact, sound transmission and afferent response for some high frequencies are even accentuated by the stiffening of the ossicular chain (see Figure 19).

The threshold sound intensity for eliciting the acoustic reflex is usually considered to be about 80 dB, but this reflex is quite sensitive to anesthesia. In awake cats, some modulation of stapedius tension, with concomitant effects on middle ear transmission, are present at much lower sound levels (Simmons, 1959). The neuroanatomical pathways mediating the stapedius reflex also will be discussed in the next chapter.

Adaptation. Another phenomenon which may be related to setting the dynamic range for discrimination of many complex sounds is adaptation. When a tone is switched on, the neural discharge is high initially, and then drops to lower steady-state rates. This adaptation is not demonstrable in intracellular voltage records made from inner hair cells but is seen in the excitatory post-synaptic potentials in tetrodotoxin-blocked afferent fibers (Furukawa et al., 1978). Thus, adaptation is a pre-synaptic phenomenon, probably attributable to the inability of the hair cell to release neurotransmitter fast enough to support sustained, high firing rates.

Rate-intensity functions such as those shown in Figure 18 usually reflect the steady-state, or adapted, firing rate. However, when the onset firing rate (or, more precisely, the probability of a spike within a few msec after tone onset) is plotted as a function of intensity, changes are seen over a much greater range than is seen for adapted fibers (Smith, 1979). In other words, the dynamic range is greater if

the initial rather than the adapted firing rate is considered. This observation has functional significance because most sounds of biological relevance, including speech, demand the encoding of rapid changes of amplitude rather than steady-state intensity levels.

Synchrony

Thus far, we have considered how frequency and intensity is encoded by changes in the absolute firing rates of afferent auditory neurons. However, a great deal of information is also encoded in the timing of discharges. For frequencies below about 5 kHz, auditory nerve fibers tend to synchronize their spike discharges to the period of a stimulus tone (phase-locking). As seen in Fig. 20, individual neurons almost never discharge on every cycle of a pure tone but rather every 2nd or 3rd cycle or so. When they do discharge, however, they fire only during that part of the cycle in which the basilar membrane is displaced upward (this direction of deflection leads to hair cell depolarization). If all action potentials occurred at exactly the same phase in the stimulus cycle, the intervals between successive spikes should always be integer multiples of the stimulus period. The interspike interval histograms for the fiber shown in Fig. 21 indicates that indeed most intervals cluster around multiples of the stimulus period (Rose et al., 1967). For each individual fiber responding to a low- to medium-frequency tone, the probability of firing is relatively fixed for each cycle of the stimulating tone, resulting in a Poisson-type distribution of interspike intervals.

Little is actually known about the way in which the central nervous system makes use of these temporal cues for frequency discrimination. However, an ensemble of 8th nerve fibers from the same region of the cochlea could converge on more central units to give an accurate measure of stimulus frequency up to about 1 kHz (Godfrey, et al., 1975). Brainstem units tuned to particular stimulus periods could in turn convert this temporal coding to a spike-rate representation, as has been demonstrated in the coding of interaural timing cues involved in binaural hearing which will be discussed in the following chapter.

Temporal information in auditory nerve firing patterns is undoubtedly responsible for frequency discrimination in some patients with prosthetic devices. Many totally deaf persons can have a rudimentary form of hearing restored by a cochlear implant, which electrically stimulates the auditory nerve directly. With a single stimulating electrode the same population of fibers is being activated, regardless of the stimulus frequency; therefore, any frequency discrimination must depend entirely on temporal cues in the discharge patterns of the auditory neurons. Patients with these implants can discriminate pulse trains of different frequencies up to about 1000 Hz by detecting differences in inter-pulse periods (Dobie and Dillier, 1985). This provides an unambiguous demonstration of at least a limited use of temporal cues by the central nervous system.

Population Studies

Recent technical advances have made it possible to measure the responses of large numbers of auditory nerve cells in the same animal

and to test the idea that all of the information needed to discriminate sound is provided by the combination of which auditory nerve fibers are active (place coding) and the absolute firing rate of these fibers. While such recordings make it easy to see how the frequencies of simple signals are encoded, it is difficult to understand how broad-spectrum stimuli such as speech sounds are transmitted. Sachs and Young (1979) have shown that the acoustic features which make different vowels distinguishable at normal conversational intensities, i.e., the relative concentration of sound in particular frequency regions (formants), are not distinguishable in the profile of firing rates across the population of auditory nerve fibers. In their experiment recordings were made from a large population of auditory nerve fibers in cats and the responses of each fiber to a variety of vowel sounds was determined. As shown in Figure 22, the firing rates of most fibers were saturated at ordinary conversational levels. However, when temporal coding was considered across the same array of auditory neurons, the vowel formants were easily distinguished in the response profile over a wide range of intensities (Young and Sachs, 1979) (Fig. 23). Therefore, the auditory system may make use of temporal encoding of discharges to discriminate among complex stimuli such as vowels.

It should be remembered that the acoustic reflex and efferent regulation of cochlear function are abnormal in these anesthetized animals. In addition, most of the fibers were probably high-spontaneous rate/low-threshold fibers; high-threshold fibers, on the other hand, may permit encoding of vowel formants without requiring temporal coding.

Actually, further studies have shown that for consonants, analyses of firing rates are about as good as synchrony (temporal) measures for detecting the acoustic features of a complex stimulus (Sinex and Geisler, 1983). Synchrony measures, however, do appear to be more resistant to the effects of background noise and to reflect more accurately the psychophysical effects of background noise (Degutte and Kiang, 1984).

CONCLUSION

In the foregoing discussion we have attempted to relate physiology to perception at every level of the auditory periphery, from the pinna to the auditory nerve. Perceptual abilities, after all, can be no better than the quality of information transmitted at each stage of processing, and attempts to reconcile exquisite psychoacoustic performance with physiological measures have motivated and directed a great deal of physiological research in recent years. The next chapter will relate both psychoacoustics and peripheral physiology to the anatomy and physiology of the central nervous system auditory pathways, and consider the special problem of integrating information from the two ears.

ANNOTATED REFERENCES

1. Pickles, J.O.: An introduction to the physiology of hearing. New York: Academic Press, 1982. This lucid and complete treatment of peripheral auditory physiology provides the interested reader with further details about issues discussed in this chapter.
2. Berlin, C.I. (ed.): Hearing Science. San Diego, College-Hill Press, 1984. An account of areas of rapid progress in the periphery, including inner and outer hair cell function, temporal coding, and cochlear fluid dynamics.
3. Pickles, J.O.: Recent advances in cochlear physiology. Prog. Neurobiol 24:1-42, 1985. A well-written, somewhat technical, review article summarizing recent advances in cochlear physiology.
4. von Bekesy, G.: Experiments in hearing (research articles from 1928 to 1958), New York, McGraw-Hill Book Co., 1960. A classic, especially valuable for students interested in the history of auditory science.
5. Wever, E. and Lawrence, M.: Physiological Acoustics. Princeton University Press, Princeton, NJ, 1954. A very readable account of acoustics and outer and middle ear function, areas in which the ideas have not changed much since 1954.

REFERENCES

1. Bobbin, R.P.: Glutamate and aspartate mimic the afferent transmitter in the cochlea. Exp. Brain Res. 34:385-393, 1979.
2. Brownell, W.E.: Observations on a motile response in isolated outer hair cells. In: Mechanisms of Hearing, W.R. Webster and L.M. Aitkin (eds.), Monash University Press, Clayton, Victoria, 1983.
3. Bruns, V.: Structural adaptation in the cochlea of the horseshoe bat for the analysis of the long CF-FM echolocating signals. In R.G. Busnel and J.F. Fish (Eds.) Animal Sonar Systems, pp. 867-869, New York, Plenum, 1980.
4. Crawford, A.C. and Fettiplace, R.: The frequency selectivity of auditory nerve fibres and hair cells in the cochlea of the turtle. J. Physiol. 306:79-125, 1980.
5. Davis, H.: A mechano-electrical theory of cochlear action. Ann. Otol. Rhinol. Laryngol. 67:789-801, 1958.
6. Delgutte, B., and Kiang, N.Y.S.: Speech coding in the auditory nerve: V. Vowels in background noise. J. Acoust. Soc. Am. 75:908-918, 1984.
7. Dobie, R.A. and Dillier, N.: Some aspects of temporal coding for single-channel electrical stimulation of the cochlea. Hearing Research 18:41-55, 1985.
8. Drescher, M.S., Drescher, D.G., and Medina, S.E.: Effect of sound stimulation at several levels on concentration of primary amines,

including neurotransmitter candidates. J. Neurochem. 41:309-320, 1983.

9. Fettiplace, R. and Crawford, A.C.: The coding of sound pressure and frequency in cochlear hair cells of the terrapin. Proc R. Soc. Lond. B. 203:209-218, 1978.
10. Flock, A., Flock, B., and Ulfendahl, M.: Mechanisms of movement in outer hair cells and a possible structural basis. Arch. Otorhinolaryngol. 24:83-90, 1986.
11. Frishkopf, L.S. and DeRosier, D.J.: Mechanical tuning of free-standing stereociliary bundles and frequency analysis in the alligator lizard cochlea. Hearing Res. 12:393-404, 1983.
12. Furakawa, T., Hayashida, Y., and Masuura, S.: Quantal analysis of the size of excitatory post-synaptic potentials at synapses between hair cells and afferent nerve fibers in goldfish. J. Physiol. 276: 211-226, 1978.
13. Gitter, A.H., Zenner, H.P., and Fromter, E.: Membrane potential and ion channels in isolated outer hair cells of guinea pig cochlea. ORL 48:68-75, 1986.
14. Godfrey, D.A., Kiang, N.Y.-S., and Norris, B.E.: Single unit activity in the posterovental cochlear nucleus of the cat. J. Comp. Neurol. 162:247-268, 1975.
15. Goodman, D.A., Smith, R.L., and Chamberlain, S.C.: Intracellular and extracellular responses in the organ of Corti of the gerbil. Hearing Res. 7:161-179, 1982.

16. Hudspeth, A.J.: Mechanoelectrical transduction by hair cells in the acousticolateralis sensory system. Ann. Rev. Neurosci. 6:187-215, 1983.
17. Hudspeth, A.J. and Corey, D.P.: Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli. Proc. Nat. Acad. Sci (USA) 74:2407- 2411, 1977.
18. Hunter-Duvar, I., Landolt, I., and Cameron, R.: X-ray microanalysis of fluid spaces in the frozen cochlea. Arch. Otorhinolaryngol. 230: 245-249, 1981.
19. Khanna, S.M. and Leonard, D.G.B.: Laser interferometric measurements of basilar membrane vibrations in cats. Science 215:305-306, 1982.
20. Kim, D.O.: Functional roles of the inner- and outer-hair-cell subsystems in the cochlea and brainstem. In C.I. Berlin (ed.), Hearing Science, College-Hill Press, San Diego, 1986.
21. Leonard, D.G.B. and Khanna, S.M.: Histological evaluation of damage in cat cochleas used for measurement of basilar membrane mechanics. J. Acoust. Soc. Am. 75:515-527, 1984.
22. Lewis, R.S. and Hudspeth, A.J.: Voltage and ion dependent conductances in solitary vertebrate hair cells. Nature 304: 538-541, 1983.
23. Liberman, M.C.: Auditory-nerve response from cats raised in a low-noise chamber. J. Acoust. Soc. Am. 63:442-455, 1978.

24. Liberman, M.C. and Brown, M.C.: Physiology and anatomy of single olivocochlear neurons in the cat. Hearing Res., in press, 1986.

25. Lim, D.J.: Cochlear anatomy related to cochlear micromechanics. A review. J. Acoust. Soc. Am. 67:1686-1695, 1980.

26. Mountain, D.C.: Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanics. Science 210:71-72, 1980.

27. Mulroy, M.J.: Cochlear anatomy of the alligator lizard. Brain Behav. Evol. 10:69-87, 1974.

28. Peake, W.T. and Ling, A.L.: Basilar-membrane motion in the alligator lizard: Its relation to tonotopic organization and frequency selectivity. J. Acoust. Soc. Am. 67:1736-1745, 1980.

29. Popper, A.N. and Fay, R.R.: Comparative studies of hearing in vertebrates. New York: Springer-Verlag, 1980.

30. Rasmussen, G.L.: Studies of the eighth cranial nerve of man. Laryngoscope 50:67-83, 1940.

31. Rhode, W.S.: Some observations on two-tone interaction measured with the Mossbauer effect. In E.F. Evans and J.P. Wilson (eds.), Psychophysics and Physiology of Hearing, London: Academic Press, 1977.

32. Robertson, D.: Horseradish peroxidase injection of physiologically characterized afferent and efferent neurones in the guinea pig spiral ganglion. Hearing Res. 15:113-121, 1984.

33. Rose, J.E., Brugge, J.L., Anderson, D.J., and Hind, J.E.: Phase-locked response to low-frequency tones in single auditory-nerve fibers of the squirrel monkey. J. Neurophysiol. 30:769-793, 1967.

34. Rosowski, J.J., Carney, L.H., Lynch, T.J., III and Peake, W.T.: The effectiveness of external and middle ears in coupling power into the cochlea. In: Peripheral Auditory Mechanisms, Lecture Notes in Mathematics, Vol. 64, J.B. Allen, J.L. Hall, A. Hubbard, S.T. Neely, and A. Tubis (eds.), New York: Springer-Verlag, 1986 (Appendix Insert #3).

35. Rubel, E.W and Lippe, W.R. and Ryals, B.M.: Development of the place principle. Ann Otol Rhin Laryngol 93(6):609-615, 1984.

36. Russell, I.J. and Sellick, P.M.: Intracellular studies of hair cells in the mammalian cochlea. J. Physiol. 284:261-290, 1978.

37. Ryan, A. and Dallos, P.: Effect of absence of cochlear outer hair cells on behavioral auditory thresholds. Nature (London) 253:44-46, 1975.

38. Sachs, M.B. and Young, E.D.: Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate. J. Acoust. Soc. Am. 66:470-479, 1979.

39. Sellick, P.M., Patuzzi, R., and Johnstone, B.M.: Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique. J. Acoust. Soc. Am. 72:131-141, 1982.

40. Simmons, F.B.: Middle ear muscle activity at moderate sound levels. Ann. Otol. Rhinol. Laryngol. 68:1126-1143, 1959.

41. Sinex, D.G. and Geisler, C.D.: Responses of auditory-nerve fibers to consonant-vowel syllables. J. Acoust. Soc. Am. 73:602-615, 1983.

42. Smith, R.L.: Adaptation, saturation and physiological masking in single auditory-nerve fibers. J. Acoust. Soc. Am. 65:166-178, 1979.

43. Spoendlin, H.: Neural connections of the outer hair cell system. Acta. Otolaryngol. 87:381-387, 1979.

44. Strelioff, D. and Flock, A.: Mechanical properties of hair bundles of receptor cells in the guinea pig cochlea. Soc. Neurosci., Abstract 8, 1982.

45. Thalmann, I., Thallinger, G., Comegys, T.H., and Thalmann, R.: Collagen - the predominant protein of the tectorial membrane. ORL 48:107-115, 1986.

46. Tilney, L.G. and Saunders, J.C.: Actin filaments, stereocilia, and hair cells of the bird cochlea: The length, number, width, and distribution of stereocilia of each hair cell is related to the position of the hair cell on the cochlea. J. Cell Biol. 86:244-259, 1982.

47. Webster, D.M., and Webster, M.: The specialized auditory system of kangaroo rats. In W.D. Neff (ed.), Contributions to Sensory Physiology, vol. 8 (pp. 161-196), New York, Academic Press, 1984.

48. Wiederhold, M.L. and Kiang, N.Y.S.: Effects of electric stimulation of the crossed olivocochlear bundle on single auditory-nerve fibers in the cat. J. Acoust. Soc. Am. 48:950-965, 1970.

49. Young, E.D. and Sachs, M.B.: Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J. Acoust. Soc. Am. 66:1381, 1979.

FIGURE LEGENDS

Fig. 1 This schematic drawing combines a coronal section through the ear canal and middle ear with a more diagrammatic representation of the eustachian tube, cochlea, and internal auditory canal. The middle ear muscles are not shown.

Fig. 2 A section through the modiolus, or axis, of the cochlea shows its spiral orientation, with the three cochlear fluid compartments in each turn.

Fig. 3 The lower curve shows the minimum audible sound pressure level (re 20 Pa) for human subjects, as a function of frequency. The upper curve shows the upper limits of dynamic range: the intensity at which sounds are felt or cause discomfort.

Fig. 4 The steady-state spectra of three vowels (as sounded in the words "beet," "bet," and "bat") are shown here. Each has a fundamental frequency component of about 125 Hz (voice pitch), produced by the vibrations of the vocal cords, and a series of overtones whose frequencies are integer multiples of the fundamental. The positions of vocal tract structures, especially the tongue, vary to make different cavity resonances in the mouth and throat, enhancing some frequencies at the expense of others. The peaks in the spectra (called "formants") are different for each vowel.

Fig. 5 A tiny microphone placed near the tympanic membrane records different sound levels (for many frequencies), in response to a

constant sound source, than a microphone placed in the same point in space, but with the person absent. This effect, due mainly to ear canal and pinna resonance, is demonstrated here; the difference between the two microphone readings is plotted as a function of frequency.

Fig. 6 Both the difference in length between malleus handle (L_1) and incus long process (L_2), and the much larger ratio of areas of tympanic membrane (A_1) and stapes footplate (A_2) are shown here.

Fig. 7 Cross-section of cochlear duct. The boundaries of scala media with scala vestibuli and scala tympani are Reissner's membrane and the basilar membrane, respectively.

Fig. 8 For ease of illustration, the cochlea is drawn "uncoiled." Inward movements of the stapes footplate cause compensatory outward movements of the round window membrane. For static pressure changes and very low frequencies, the pressure is transmitted from scala vestibuli to scala tympani via the helicotrema. For audible frequencies, the cochlear partition is displaced as indicated by the dotted line, in different places for different frequencies.

Fig. 9 Each solid curve indicates the displacement of the basilar membrane at a particular point in time, in response to a 200 Hz tone. The darker curves occur later in time and show the progression of the traveling wave from base to apex of the cochlea. The dotted line indicates the envelope of displace-

ment for this tone, i.e., the maximum displacement for each point along the basilar membrane. The actual excursions are many times smaller (relative to the length of the basilar membrane) than illustrated here.

Fig. 10 Each curve shows the response for a given point along the basilar membrane to tones of varying frequency. The curve farthest to the right is for a point near the midpoint of the cochlea and shows maximum response to about 2.5 kHz, gradually decreasing response to lower tones, and sharply reduced response to higher tones. The curves to the left are for progressively more apical locations.

Fig. 11 The solid lines show isoresponse curves for basilar membrane displacement (x) and velocity (o) for a point near the base of the cochlea. For each frequency, the sound pressure level needed to obtain a criterion response (3.5 A and 0.04 mm/sec, respectively) is plotted. The dashed line shows an isoresponse curve obtained from an auditory nerve fiber with a similar best frequency (about 18 kHz); the sound level needed for a small increase over resting spike rate is plotted. Neural frequency threshold curves are discussed later in this chapter (AUDITORY NERVE). All curves are from guinea pigs.

Fig. 12 This diagram shows how an "upward" (toward scala vestibuli) displacement of the cochlear partition can create a shearing force tending to bend outer hair cell stereocilia in an excitatory direction.

Fig. 13 Deflection of the hair bundle toward the tallest row of stereocilia opens poorly-selective cationic channels near the stereocilia tips. Influx of potassium depolarizes the cell. Voltage-sensitive calcium channels open in turn, permitting neurotransmitter release across the synapse to the afferent neuron.

Fig. 14 The lower trace shows the waveform of a 5 kHz tone burst stimulus, while the upper trace displays an electrical potential response recorded by an electrode on the round window of a rat's cochlea. "N_1" and "N_2" are compound action potentials from synchronous auditory nerve activity; "CM" is the cochlear microphonic, or AC cochlear potential; the summating potential (SP), or DC cochlear potential is not labeled but is clearly visible as the elevation of the entire response above the baseline level seen after the tone burst ends.

Fig. 15 The afferent innervation of the cochlea is shown diagrammatically. The more numerous Type I neurons converge on inner hair cells, while each Type II neuron branches extensively to supply several outer hair cells after running spirally along the organ of Corti.

Fig. 16 Each of these frequency threshold curves (FTC's) plots sound intensity required to produce a minimal increase in spike rate (over spontaneous rate), as a function of stimulus frequency for a single auditory nerve fiber. Each neuron has a characteristic frequency to which it is most sensitive. Threshold is

plotted in arbitrary units of dB attenuation from a standard dB attenuation from a standard (high-intensity) sound; thus, "-100" represents a sound 100 dB less intense than the reference level.

Fig. 17 The curve joining filled circles shows a typical frequency threshold curve. The open triangle depicts an excitatory tone at the fiber's characteristic frequency, about 10 dB above threshold (plotted here as dB re stapes displacement, rather than in sound pressure level). Tones in the shaded area, although mostly outside the primary response area and thus unable to elicit any fiber response on their own, are able to inhibit the fiber's response when presented simultaneously with the excitatory tone.

Fig. 18 A single auditory nerve fiber's response to sounds of different frequencies is shown here. For its characteristic frequency (11.3 kHz), the spontaneous spike rate (about 40/sec.) is exceeded for tones above about 10 dB SPL, but the neuron's dynamic range is only about 40 dB.

Fig. 19 The effects of stapedius muscle contraction on the acoustic admittance (inverse of impedance) and on sound transmission through the middle ear is plotted as a function of frequency. Both admittance and transmission are decreased most for low frequencies.

Fig. 20 Spike discharges of a single auditory fiber are shown (upper trace) in response to a 300 Hz tone (lower trace).

Fig. 21 These interval histograms show the distribution of interspike intervals seen for different tones, in the responses of a single nerve fiber. For 599 Hz, the modal interval between spikes was about 1.67 msec, the period of the tone, the next-most-common interval was twice the period, and so on. Integer multiples of stimulus tone period are shown as dots along the ordinate.

Fig. 22 The upper family of curves depicts responses of a large number of auditory nerve fibers (each point along the frequency axis represents a fiber with that characteristic frequency) to the vowel /I/, as in "beet." For moderate conversational intensities (64-84 dB SPL), most fibers fire near their saturation rates (normalized rate = 1.0), and formant peaks (see Fig. 4) are indistinguishable. The lower family of curves shows the same phenomenon for the vowel /a/ as in "bat."

Fig. 23 As in Fig. 22, the ordinate arrays auditory nerve fibers according to characteristic frequencies, and the families of curves indicate responses of this population of neurons to vowel sounds at different intensities. However, the abscissa here is not spike rate, but "average localized synchronized rate," a measure of the degree to which each neuron's response was phase-locked to the spectral components of the vowel stimulus. The formant peaks which characterize the stimuli are now apparent in the response profiles, across a range of intensities.

FIGURE 1

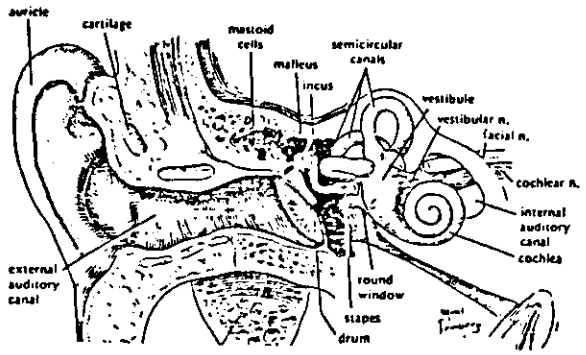


FIGURE 2

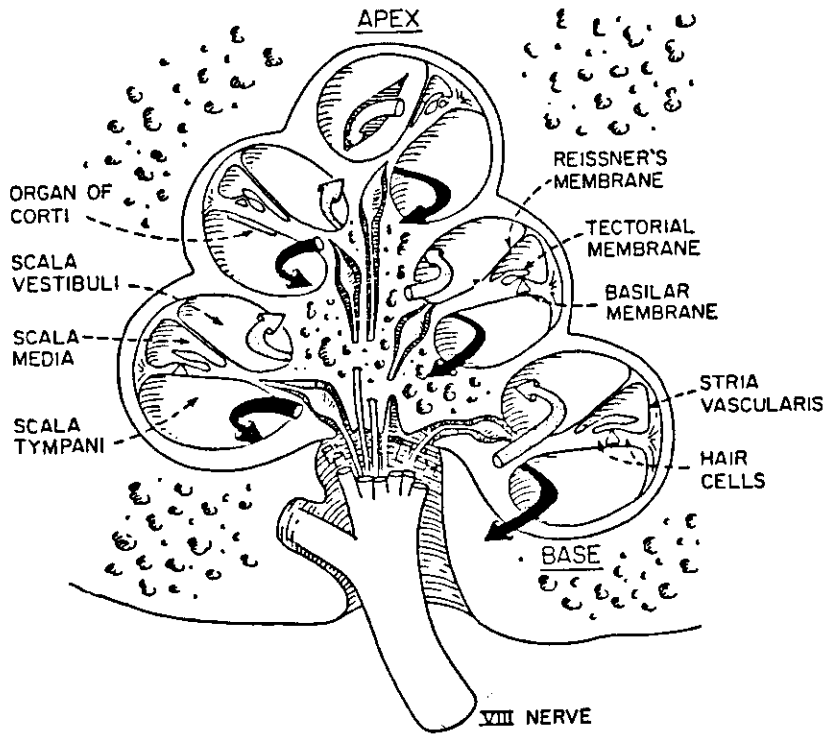


FIGURE 3

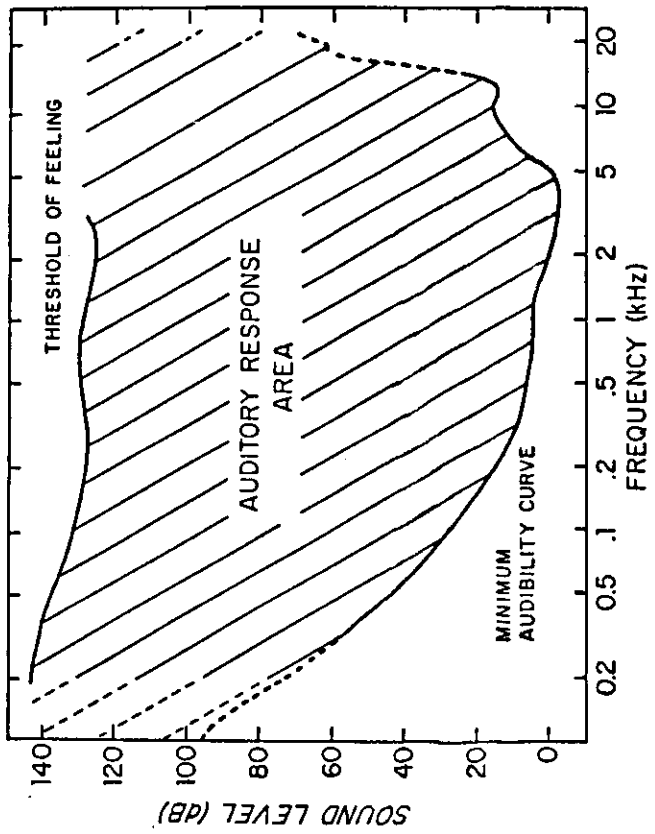


FIGURE 4

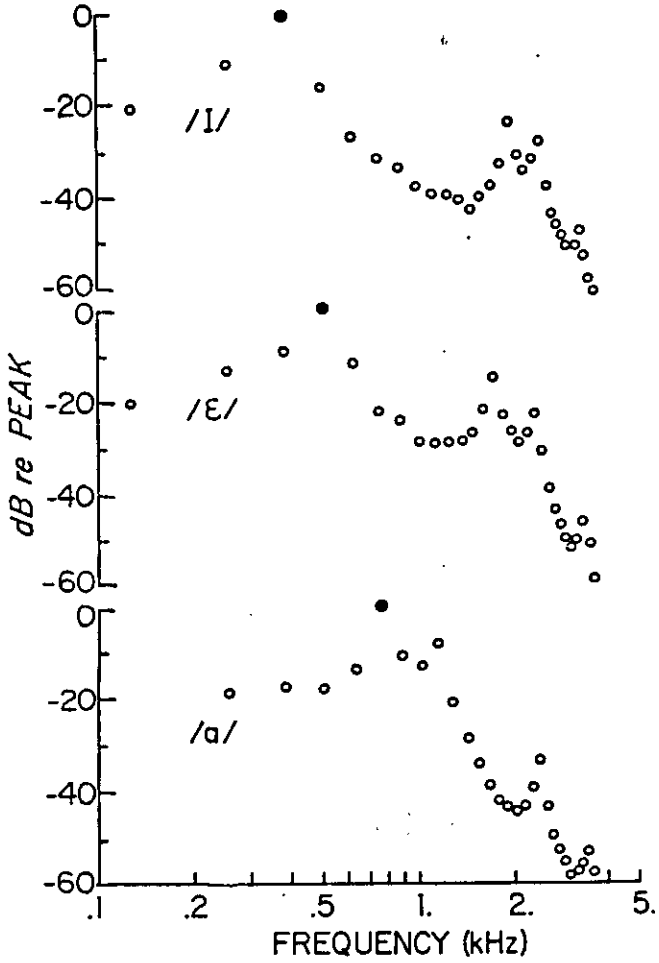


FIGURE 5

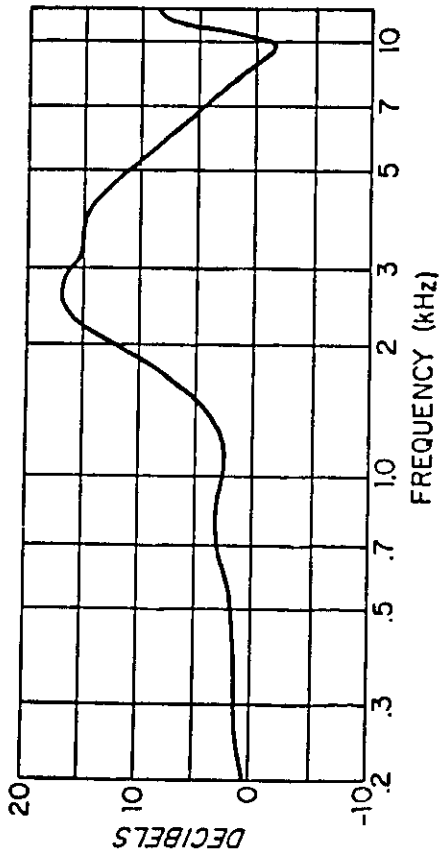


FIGURE 6

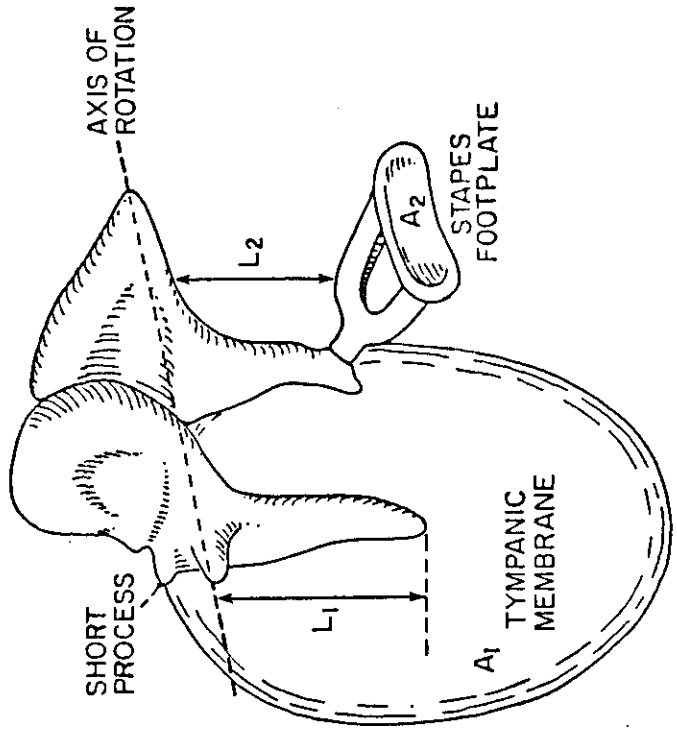
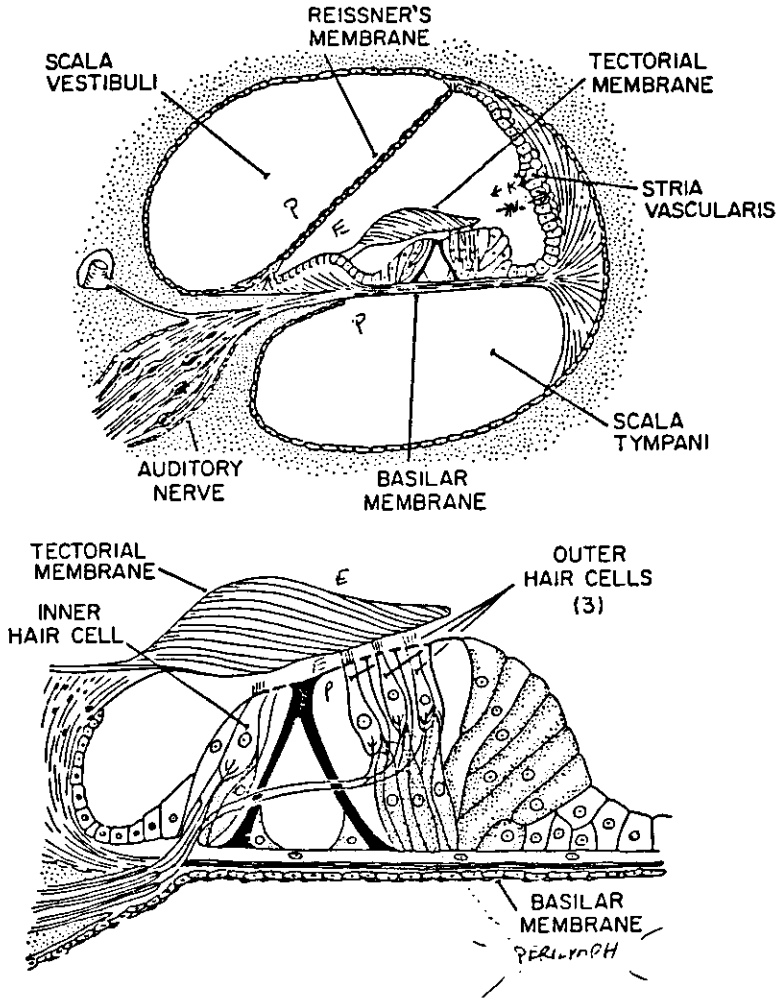
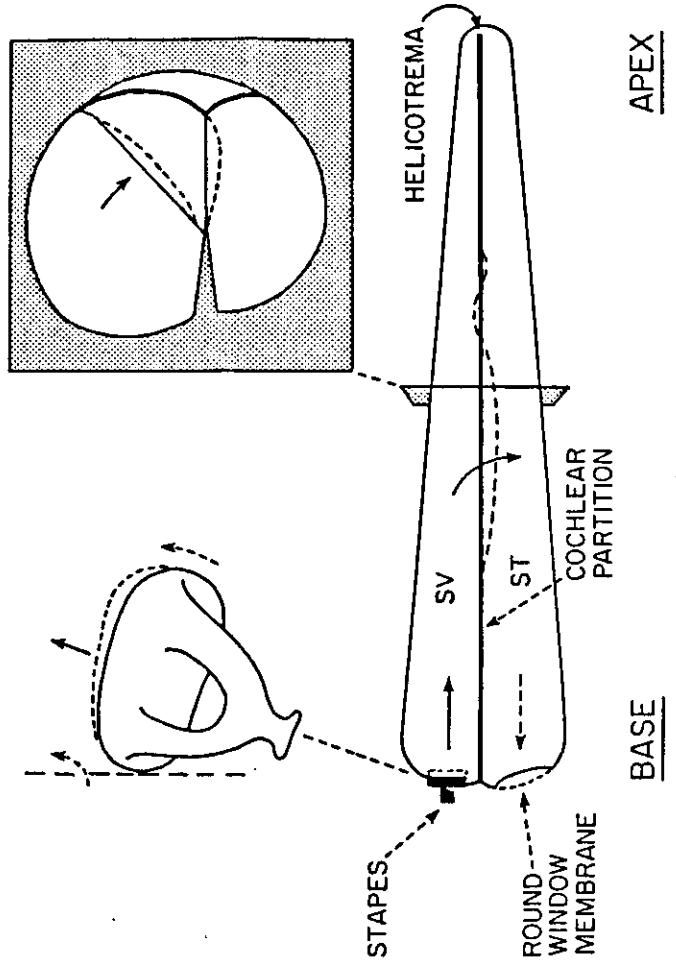


FIGURE 7



FIGUR



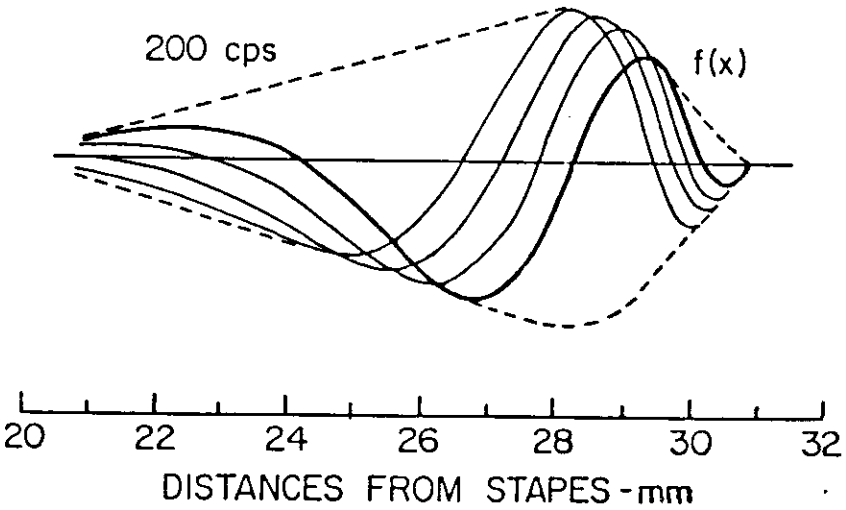


FIGURE 9

17-10

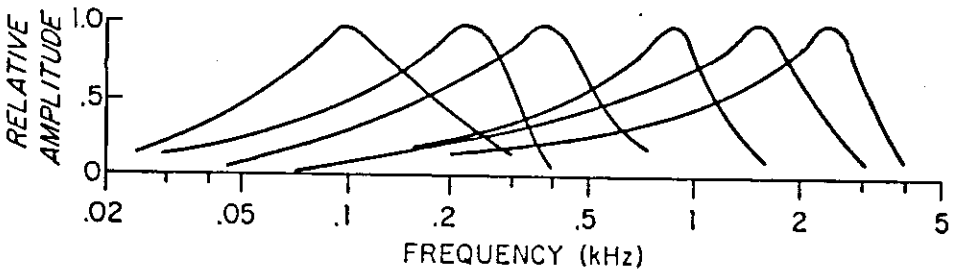


FIGURE 10

FIGURE 11

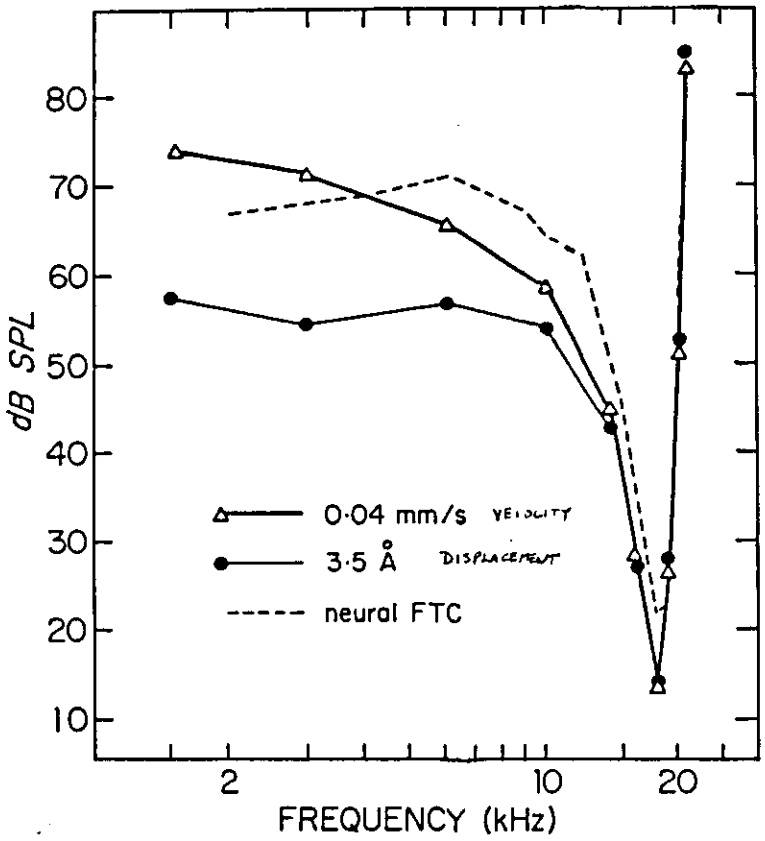
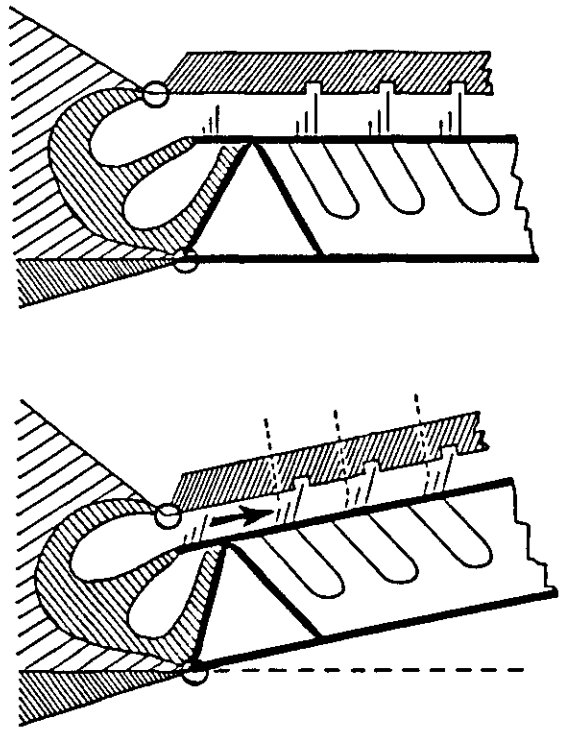


FIGURE 12



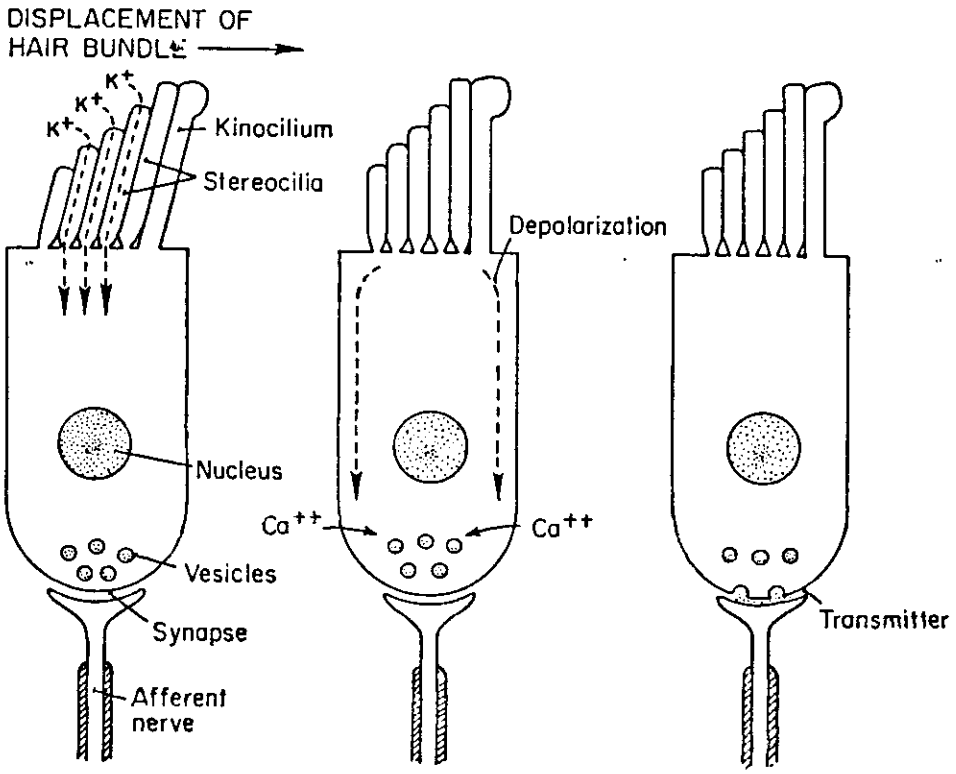


FIGURE 13

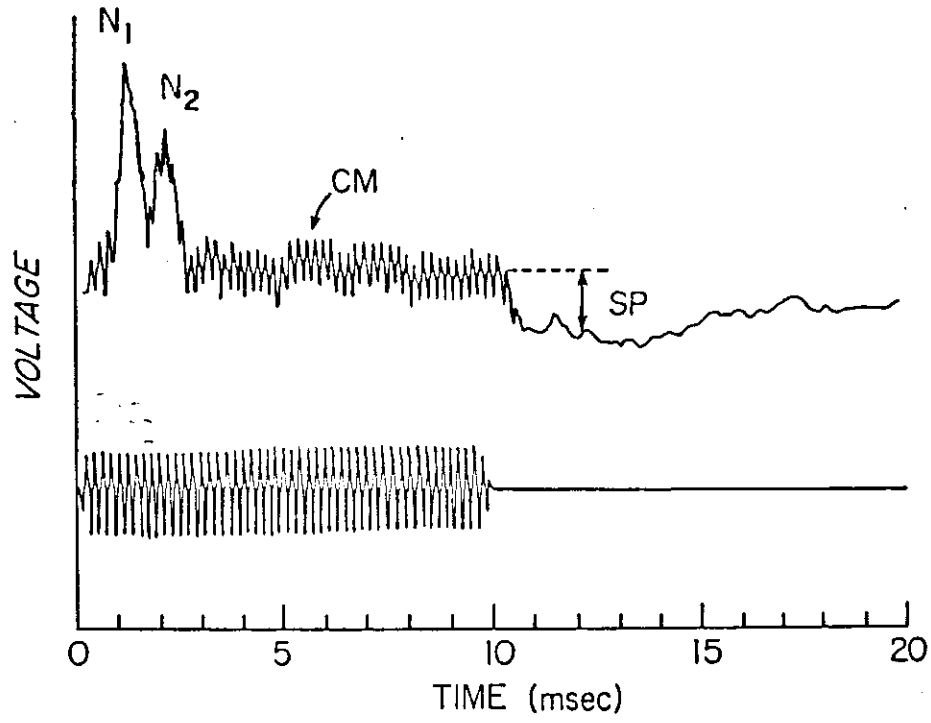


FIGURE 14

FIGURE 15

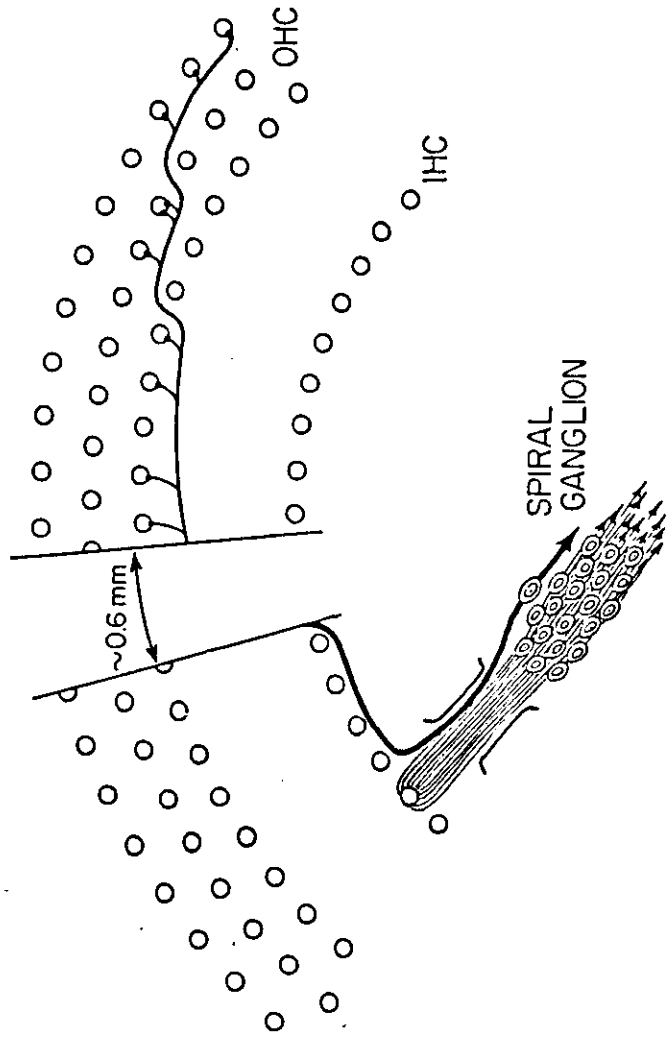


FIGURE 16

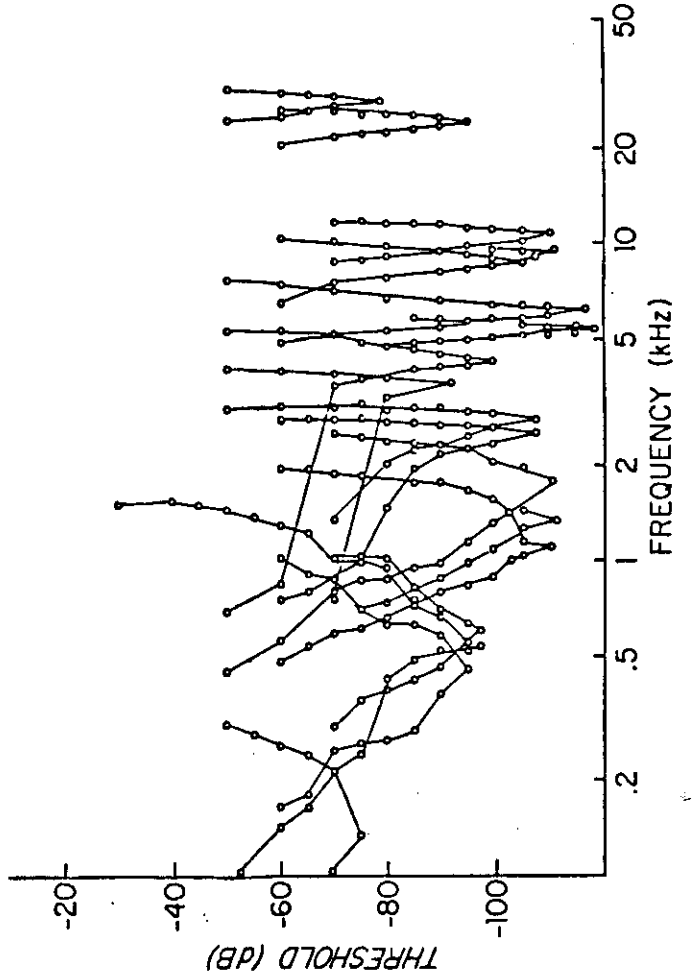


FIGURE 17

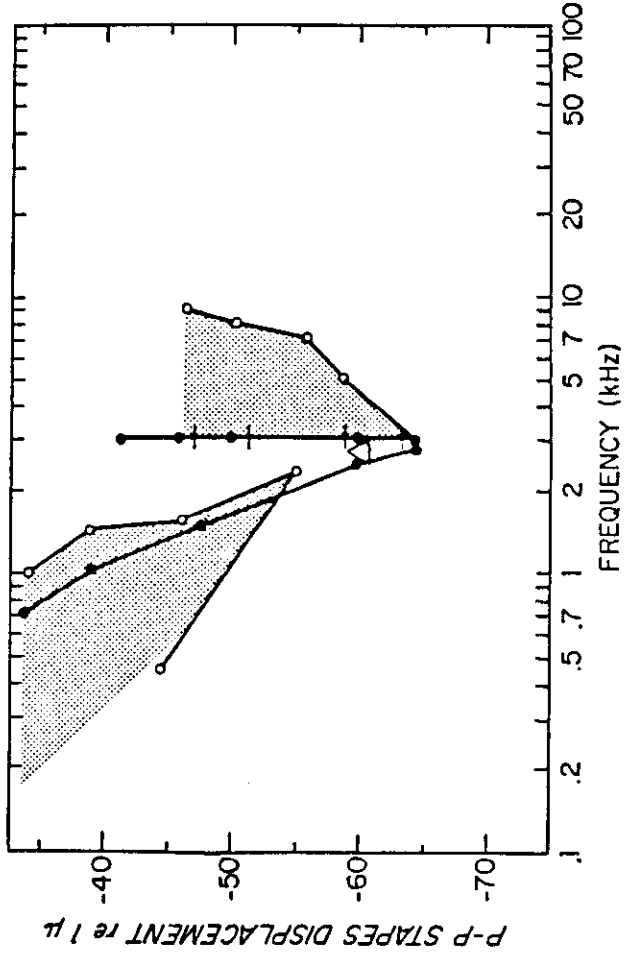
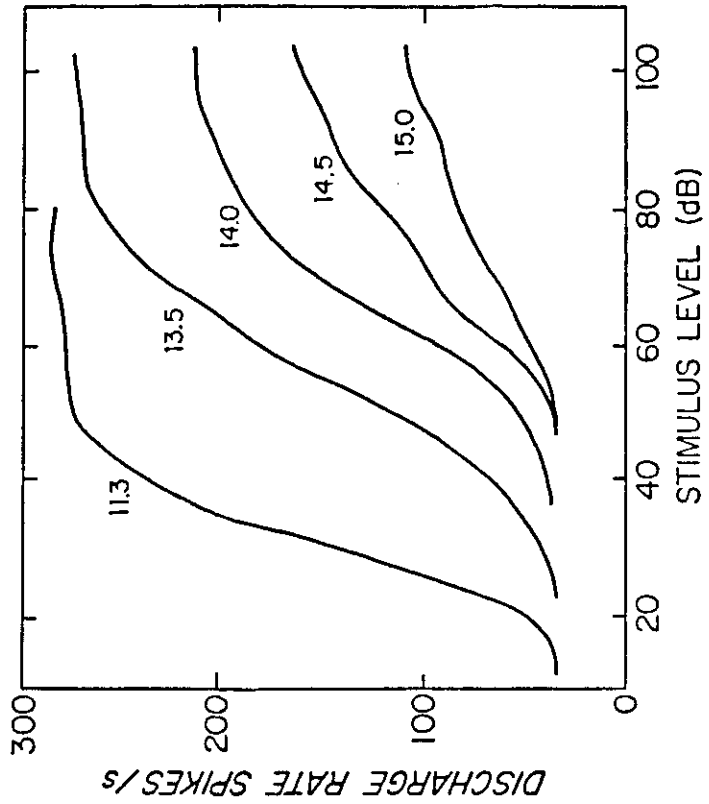


FIGURE 18



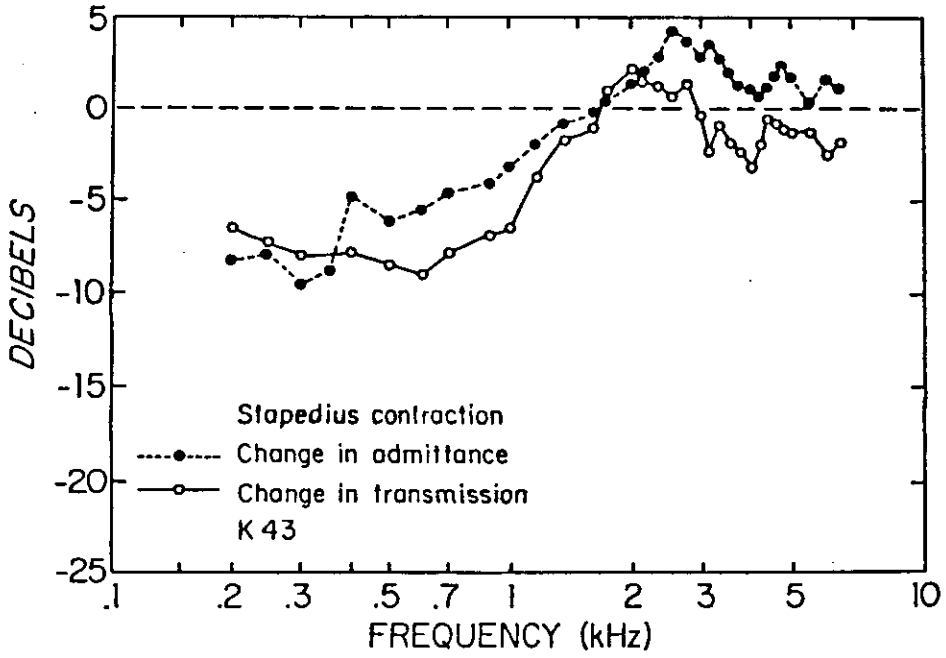


FIGURE 19

TOP

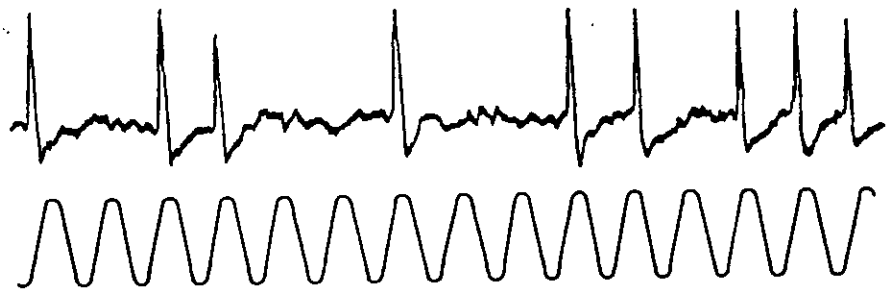


FIGURE 2(

FIGURE 21

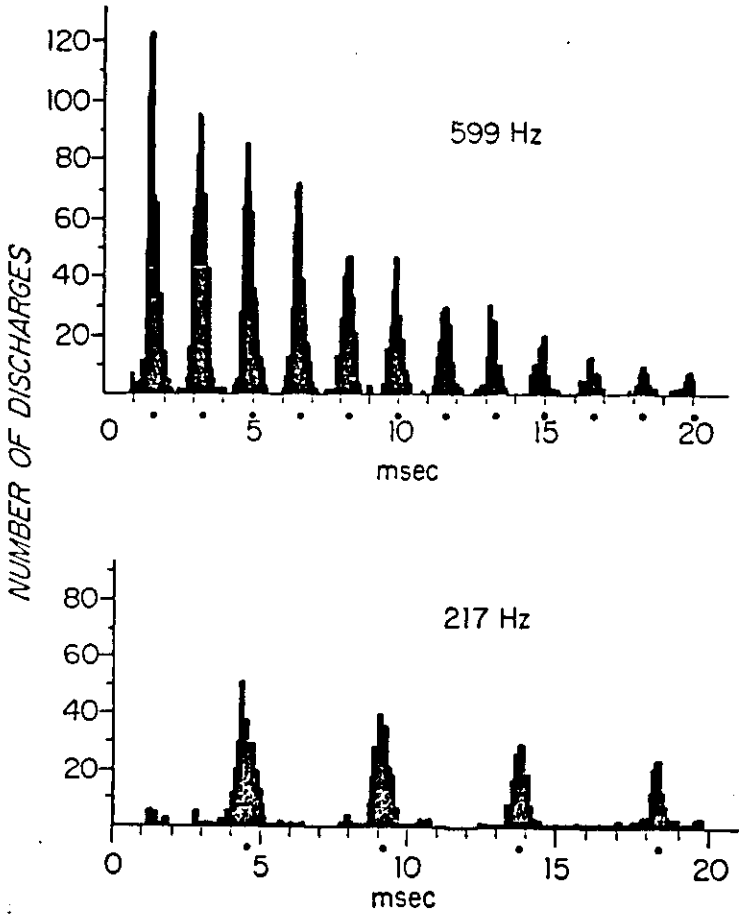


FIGURE 22

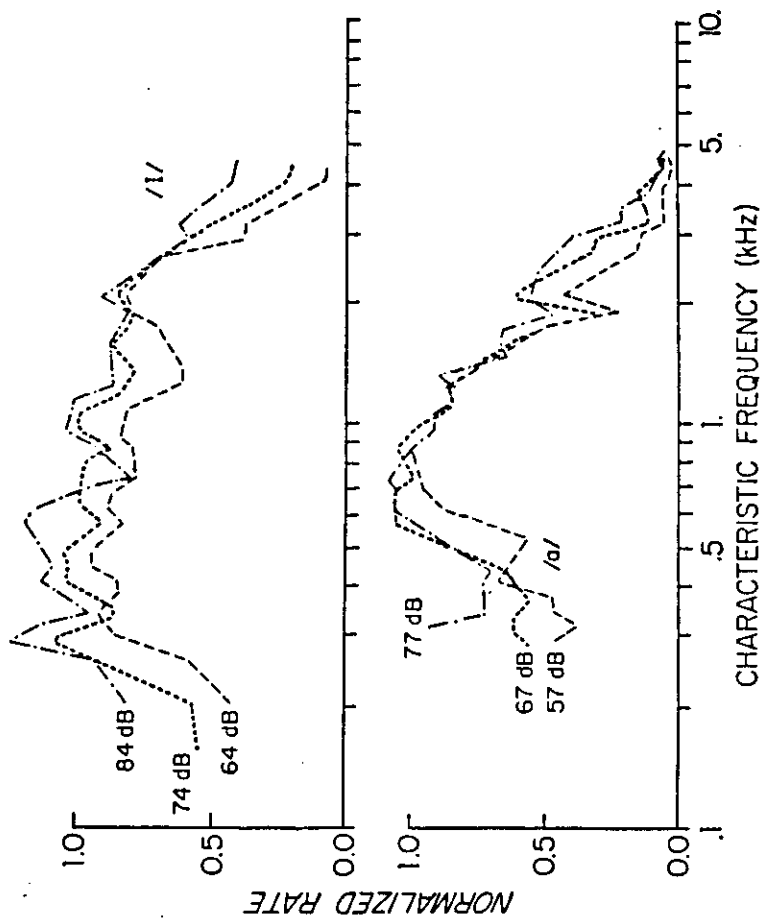
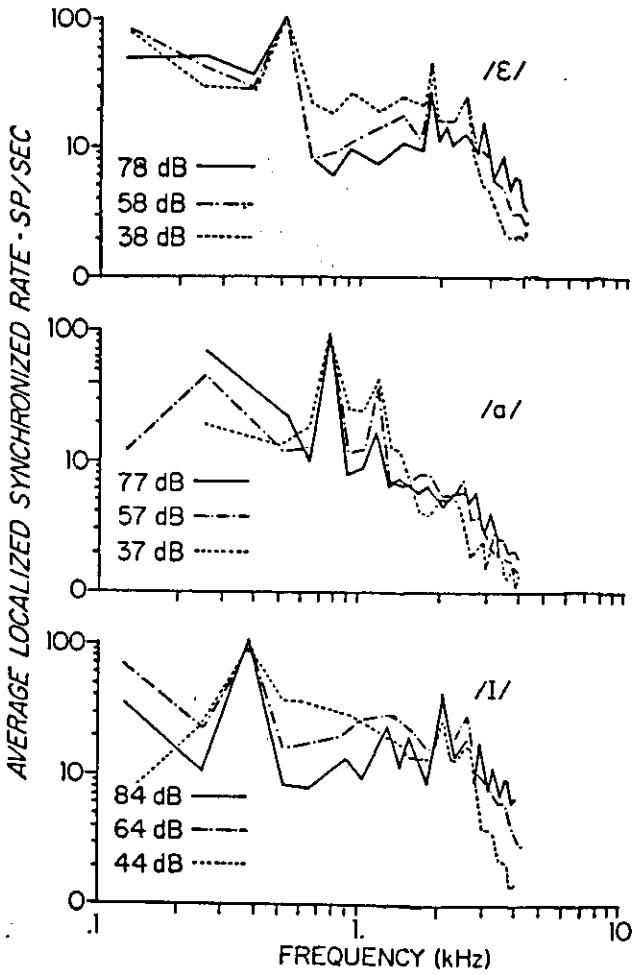


FIGURE 23


AVERAGE LOCALIZED SYNCHRONIZED RATE - SP/SEC
100
10
0
/ɛ/
78 dB
58 dB
38 dB
/a/
77 dB
57 dB
37 dB
/I/
84 dB
64 dB
44 dB
.1
1.
10
FREQUENCY (kHz)