

Chapter 2

Physical and Biological Bases of Hearing

Hearing relates to the sense responsible for translating a series of pressure variations in the air into an action potential, i.e., something that the brain can recognize. Before describing the biological bases of hearing, it is first necessary to understand what the brain needs to recognize.

2.1 Physical Characteristics of a Simple Sound Wave

Sounds are produced because something vibrates in the environment. These vibrations are disturbances and their propagation is possible only because it happens in a material medium. This medium is usually air, but it could also be, for example, water or any other substance. If you are underwater and try to talk to someone, you will find that this is possible, but the carried message is far from being as clear as it is usually. In short, a body which vibrates produces sound, provided that the vibrations do not occur in a vacuum where nothing is transmitted.

More specifically, the vibrations cause a series of compressions and rarefactions of the molecules in the environment. The normal pressure in the air is successively increased or decreased. As discussed below, the characteristics of these variations can be represented using a simple sine wave (for pure sound).

2.1.1 Frequency and Phase

A key thing to consider in the analysis of sound is the speed of variations ranging from compressions to rarefactions to compressions and so on. These changes occur more or less rapidly. This speed of state changes is called the frequency, i.e., the number of cycles (“compression-rarefaction”) completed during a given period. It

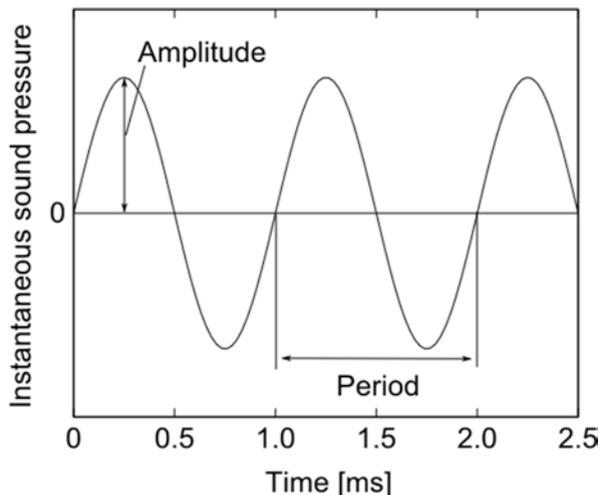


Fig. 2.1 Illustration of a sound (sinusoidal) wave for a pure tone of 1000 Hz (or 1 kHz)

was agreed to express this frequency in a number of cycles completed in 1 s. One cycle per second is 1 hertz (Hz), the unit used to express frequency and named after the German physicist Heinrich Hertz.

The time taken to complete one cycle of the sine wave is called the period (Fig. 2.1). As for the circular motion, a period (or a complete cycle) involves 360° (360 degrees). The beginning of the cycle is 0° , whereas the maximum compression and the maximum rarefaction occur at 90° and 270° , respectively. Also, the relative position of two sounds over time is called phase. If two pure tones arrive at a given point in time with a difference of $1/8$ of a cycle, they will be described as being 45° out of phase.

If a sound has a frequency of 1 Hz when a cycle is completed in 1 s, a sound completing 500 cycles in 1 s has a 500-Hz frequency. If a cycle takes only 1 ms to be completed, that is to say, 1000 cycles are completed in a second, it will be a 1000-Hz, or 1-kHz, sound (pronounce kHz “kilohertz”).

Sometimes, to express the idea of frequency, we use the notion of wavelength. This is denoted by the Greek letter lambda (λ) and consists of the linear distance between two successive compressions. Of course, the fewer cycles traveled in a given time, the longer the wave. However, this length is also determined by the propagation speed of the wave. Determined by the environment in which the wave is generated, the speed is greater in a denser medium. The speed is, for example, 340 m/s in the air and 1500 m/s in water. Thus, two waves having the same frequency in the air and water do not have the same length.

The span of audible frequencies by the human ear ranges from about 20 Hz to 20 kHz. In fact, toward the ends of this range, the detection threshold is much higher; in other words, to be heard, a sound of 20 Hz must be much louder than a 5000-Hz sound. Also, most often, conversations remain in a range of frequencies

extending from about 100 Hz to 10 kHz. Note also that the hearing abilities vary with age; thus, it becomes difficult with age to hear sounds above 15 kHz. In fact, some people, even young people, are unable to hear such sounds. Humans therefore can deal with a wide range of audible frequencies. However, this capability of hearing high frequencies does not compare at all to that of, for instance, mice (up to 90 kHz), bats (over 100 kHz), or dolphins (up to 200 kHz), which are therefore able to hear ultrasounds. In the next chapter (Fig. 3.6), you will return to this notion of frequency ranges emphasizing the ones covered by some musical instruments and by the human voices. Note that the animals who are able to hear ultrasounds will be unable to hear, for example, frequencies below 1000 Hz in the case of mice or 3000 Hz in the case of bats. Elephants, however, hear low-frequency sounds (up to 17 Hz), but cannot hear sounds above 10 kHz.

2.1.2 Amplitude

A second physical characteristic for describing sounds is called amplitude or intensity (Fig. 2.2). This feature refers to the fact that pressure variations may be more or less pronounced. It was agreed to express this magnitude with a unit called the decibel (dB—the name “bel” given in honor of Alexander Graham Bell). Indeed, this unit is issued from a pressure ratio between that exerted by a given sound and that exerted by a reference sound. In such a case, we refer more specifically to dB SPL (SPL for *sound pressure level*).

A pressure measurement is expressed in terms of force per unit of area. Thus, the sound pressure used as a reference is, by convention, 0.0002 dyne/cm^2 , a “dyne” corresponding to the force required to give a mass of 1 g an acceleration of 1 cm/s^2 . It is also possible to express the pressure with a unit called pascal (named after the scientist and philosopher Blaise Pascal), the reference sound being equal to $20 \text{ } \mu\text{Pa}$ (micropascal).

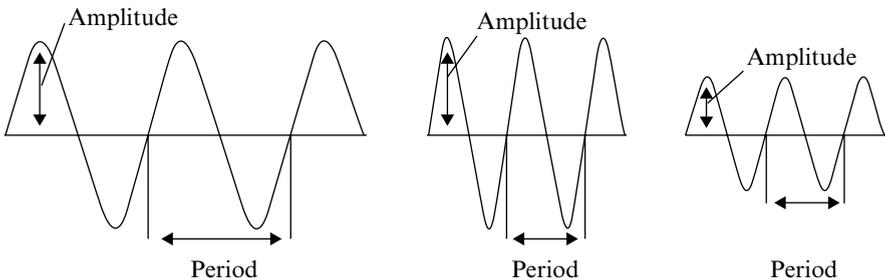


Fig. 2.2 While the wave of the *left* and the one in the *center* have the same amplitude but different frequencies, the wave in the *center* and the one on the *right* have the same frequency but different magnitudes

More specifically, to avoid having to deal with very high numbers, it was agreed to express amplitude as a logarithmic scale. Thus, the number, N , of decibels produced by a sound can be calculated as follows:

$$N \text{ dB} = 20 \log \frac{Pr_{\text{sound}}}{Pr_{\text{ref}}}$$

where Pr_{sound} is the sound pressure that is being measured and Pr_{ref} is the pressure of the reference sound (20 μPa). So, we can easily calculate the amplitude of a sound in dB once we know the pressure this sound exerts. Thus, if a sound creates a pressure that is 100,000 times greater than that of the reference sound, its amplitude is 20 times the log of 100,000, that is to say, $20 \times \log(10^5)$. The log of 10^5 is equal to 5. Accordingly, the sound has an amplitude of “100 dB” (20×5).

The constant “20” used in the calculation of the number of dB is due indeed to two constants: multiplied by 2 and multiplied by 10. The 10 stands for the decision made to use decibels rather than bels; this avoids having to work with decimals. The source of the 2 is a bit more subtle. The bel is indeed a measure of power and not a measure of pressure. Since it is agreed to express the amplitude of sound in terms of pressure ratio, it is necessary to consider what the relationship between power and pressure is. The acoustic power (P_o) is equivalent to the acoustic pressure (Pr) squared:

$$P_u = Pr^2 = 2 \log(Pr), \text{ which explains where the 2 comes from.}$$

In order to have some idea of what some sound intensities represent, here are some examples drawn from everyday life. A simple whisper or rustling leaves reaches a loudness of about 20 dB. A library is never really completely silent, and ambient sound may approach 40 dB, which is still well below the 60–70 dB observed in a work office. In fact, the intensity level of normal speech is around 60 dB. A heavy car traffic creates an amplitude level of about 80 dB, a level that reaches up to about 90 dB with the presence of a large truck or even up to 100 dB with some motorbikes. This remains a little weaker than the 100 dB of a jackhammer or 110 dB (and even more) of certain night clubs, at least near to one of the sound sources. You will understand why workers taking care of the luggage near a large airplane wear helmets to cover their ears, now that you know that large airplanes produce sound intensities of more than 130 dB, which might cause pain. Noises provoked by firing a gun or a pistol can reach more than 160 dB.

2.2 Physical Characteristics of a Complex Sound Wave

Usually, the waves heard in the environment are not pure tones like those described early in the previous section. Pure tones can be created easily in a laboratory, with a tuning fork or with some electronic instruments. Most often, what is heard, whether

it is noise, voice, or musical instruments, are complex sounds. While pure tones consist of only a single frequency, complex sounds result from the mixing of two or more waves of different frequencies.

Complex sounds can be periodic or aperiodic. They are periodic when their components are integer multiples of the lowest frequency. The lowest frequency in a sound is called the fundamental frequency (often abbreviated as F_0). It is also called the first harmonic. A periodic sound is said harmonic when it contains all the other harmonics. The vowels produced by the voice and the sounds of musical instruments, except percussion, belong in this category (harmonic sounds). If one or a few of these frequencies are missing, the sound is referred to as inharmonic. If a sound is composed of different frequencies that are not multiples of the fundamental frequency, then it is an aperiodic sound.partials, rather than harmonics, are used to describe the composition of an aperiodic sound.

Thus, the fundamental frequency is the lowest note generated by something vibrating. All the frequencies generated are specific to the properties of what is vibrating. What distinguishes one sound from another is not only the frequency and amplitude, as seen above. The distinction may also depend on what might be called the complexity, i.e., the harmonic series that the sound contains, including its fundamental frequency. Why two sounds having the same fundamental frequency and the same intensity would sound differently is because they have different harmonics.

For understanding the nuances about the complexity of sounds, one approach consists of asking the following question: why do two “Cs” on the piano are “C”? There are two elements of response to this question. On the one hand, two pure tones separated by an octave seem identical. This quality is called chroma. On the other hand, these two “Cs” share harmonics that are not shared by other notes. A 32.70-Hz C and a 65.41-Hz C will both have in their harmonics a C of 130.81 Hz; there exists such a pattern for any other note (D, F, ...). Note, however, that a C of 32.70 Hz has in its harmonics a frequency of 65.41 Hz, but the latter C (65.41) does not comprise the C of 32.70 Hz, the lowest frequency of 65.41-Hz C being actually 65.41 Hz.

Equally crucial is this second question: why, since it has the same fundamental and the same harmonics, does a C of 32.70 Hz sound differently when played on a piano rather than on a guitar? These same “Cs” differ because the relative importance of each harmonic is not the same for both instruments. The relative contributions of each harmonic depend on the vibrating properties of the instruments. The use of an oscilloscope allows to seeing that both identical “Cs” played sometimes on guitar, sometimes on the piano, have a same frequency, but the wave drawn is not the same for each instrument. In each case, however, the configuration is more complicated than that of a pure tone (simple sine wave).

There is a way to know the relative importance of the harmonics of a complex periodic sound. This could be done with a Fourier analysis, named after Jean Fourier, a French physicist of the early nineteenth century. Such an analysis allows describing quantitatively any complex wave into a series of simple components (sine waves). It is interesting to note, as stipulated by the acoustic law of Ohm, that

the ear can somehow act as a Fourier analyzer. Thus, if a few notes are played together, the auditory system can extract and hear each of the simple sounds contained in the complex sound that was produced.

White noises enter in the category of aperiodic complex sounds. These sounds are made of the mixture of all frequencies. This name, white noise, is given by analogy to white light which means, as it will be discussed in Chap. 5, not the absence of wavelengths that would allow to observe a color, but the presence of all wavelengths. White noise gives a sound similar to the one we sometimes hear when trying to tune a frequency on a radio, moving through different frequencies without being able to capture a station correctly.

It is possible to create sounds by using a filter that let pass only frequencies included within a certain range. This range is called bandwidth and can be more or less narrow. One can also use high-pass filters that allow the passage of frequencies above a certain value or low-pass filters that allow the passage of frequencies below a certain value.

Another phenomenon, called masking, refers to the incapacity to hear a sound normally audible because of the presence, simultaneously or nearly, of another sound (mask). For example, if two sounds are presented simultaneously, it is possible that both are heard. In some circumstances, i.e., according to their relative frequency and intensity, it is possible that a sound be heard and the other not. Most often, a loud sound will mask a weaker sound; also, a sound will mask sounds of equal frequencies or of higher frequencies. The frequency range that may be masked by a given sound is called the critical band. The mask does not need to be presented simultaneously to exert its influence. It can be shifted in time, but its influence will be greater if it is presented shortly before rather than shortly after the sound that is to be masked.

It should be noted that when a pure tone is produced in a laboratory, this sound may not be clear at the beginning (onset) and end (offset). In order to ensure that the transitions are not too steep, a gradual rise of intensity to reach the targeted intensity, and a gradual fall at the end of the sound, may be used. This shaping of a sound is called the envelope. The sound will be softened even with a rise and fall lasting only a few milliseconds each. Furthermore, if the sound is presented to each ear, we speak of a binaural presentation, as opposed to a monaural presentation if the sound is sent only to one ear.

2.3 Subjective Characteristics of Sounds

The impressions left by the sounds, especially when emitted by human voices or music instruments, are numerous and diverse. But before evoking their potential emotional connotation, it is first relevant to distinguish the broad categories of psychological impressions produced by the sounds that could be linked quite directly to the physical reality.

2.3.1 Pitch, Loudness, and Timbre

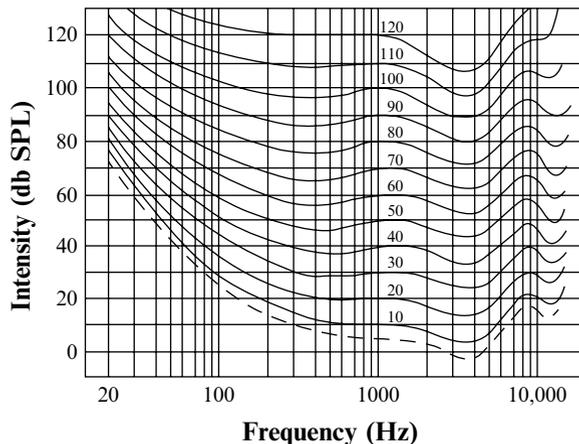
One of the subjective characteristics closely related to a physical characteristic is pitch (Hartmann, 1996; Yost, 2009). Pitch refers to the impression that the sound seems low or high. While high-pitch sounds are composed of high frequencies, the low-pitch sounds are made of low frequencies. Therefore, there is a close and direct correspondence between pitch and frequency. However, the pitch is not perfectly correlated with frequency. The intensity, for example, may exert some influence on the pitch.

It is difficult to measure directly a subjective dimension such as pitch. S. S. Stevens (see Chap. 1) addressed this problem based on the responses of observers and working on a new unit of measurement, operationally defined. Stevens thus developed the concept of mel, 1000 mels corresponding to the pitch of a 1000-Hz sound at 40 dB SPL.

A second fundamental subjective dimension of auditory perception is called loudness. This quality mainly refers to the sound intensity, that is to say, the impression that sound seems to be soft or loud. Of course, a high-amplitude sound appears louder than a sound of low amplitude, but this impression may vary depending on the frequency of the sound heard. Just as he developed the mel, Stevens also developed a unit of loudness, the sone, which is the loudness of a 1000-Hz sound at 40 dB SPL.

The fact that loudness depends not only on the intensity of sounds but also on the frequency has been highlighted by many psychophysical experiments that have led to the development of *equal-loudness contours*. These lines, called *phons* and reported in Fig. 2.3, are built on the basis of a 1-kHz standard sound. If the frequencies would exert no influence on loudness, the lines would remain flat. What the figure reveals is the fact, for example, that the loudness of a sound of 200 Hz and 60 dB SPL will be the same (about 50 sones) to that of a 2-kHz sound at 50 dB SPL. Note in conclusion that the impression of loudness is also dependent on the

Fig. 2.3 Equal-loudness contours, each expressed as phons (Fletcher & Munson, 1933)



presentation duration of the sound: with very short sounds (<200 ms), the intensity needs to be increased for generating the impression that this sound is as loud as a sound of longer duration.

A third subjective dimension of the auditory experience closely related to physical reality is called timbre. As reported above, two sounds may have the same fundamental frequency and the same amplitude, but they may nevertheless be different perceptually. What causes this difference is their timbre, which depends on the composition of each sound, i.e., on their respective harmonic arrangements.

2.3.2 Other Subjective Characteristics

Sounds can create many other subjective impressions. For example, we will have the impression that space is more or less filled by a sound. In such a case, we refer to volume (not to be confused with intensity). Of course, if we increase the intensity, the impression of volume is increased; the volume also appears greater if the pitch of a sound is low rather than high. Another subjective impression is related to the fact that a sound may seem more or less compact, or more or less hard. This quality of a sound is referred to as density, a loud sound dominated by high frequencies appearing denser than a weaker sound dominated by low frequencies.

In fact, the subjective impression caused by a sound can often be associated with its spectral composition. Already in the nineteenth century, Helmholtz reported that a sound composed only of its fundamental seems rather soft, but with a less intense fundamental and more intense harmonics, the sound rather appears hollow. You can also notice that some voices seem nasal and other sounds seem strident. Also, two notes played together seem dissonant and melodious, depending on distance (in Hz) between them.

In closing, it should be recalled that the pleasure afforded by the sounds of music can also depend on cultural habits and factors associated with learning. Complex music (e.g., symphonies or operas) are more difficult to enjoy, but with repeated exposure to a certain piece (some learning), it becomes more accessible (see Chapter 3).

2.4 Biological Bases

Between the arrival of a sound wave to the ear and the capture by the brain of an intelligible and revealing message, a long path is traveled. The waves of compressions and rarefactions included within the initial stimulus are translated through various stages that constitute the path from the external ear to the inner ear, via the middle ear.

2.4.1 Outer, Middle, and Inner Ear

The outer ear includes essentially two parts, the pinna and the auditory canal (Fig. 2.4). The function of the pinna is to collect sound waves and to direct them into the auditory canal. However, the role of the pinna, if we consider its lack of mobility, is much less important in humans than in some other vertebrates. Nevertheless, it serves to amplify sounds, especially those falling within a range of 1.5–7 kHz and, to a certain extent, contributes to locating the direction of sounds (Chap. 3).

The ear canal is a passageway that extends for about 2.5–3 cm from pinna to the eardrum. Throughout this duct, which has a diameter of about 0.75 cm, there are glands that secrete a wax, technically known as cerumen, which serves as a barrier for protecting the inner ear from dust and dirt.

Between the outer ear and the middle ear, there is a thin membrane, the eardrum, covering a surface of approximately 70 mm². The function of the middle ear is to ensure the transmission of the air movement from the eardrum to the inner ear. This transmission takes place via three tiny bones, called the ossicles: the malleus (hammer), the incus (anvil), and the stapes (stirrup). The malleus is attached to the eardrum; the incus is connected to the malleus and the stapes, and the stapes is attached to a small structure, the oval window (or vestibular window), which is the gateway through which the air vibrations are transmitted to the inner ear. The base of the stapes has a surface area of only 3 mm².

The inner ear contains an important amount of liquid. For transmitting the wave from an air medium to a liquid medium, it is necessary to overcome a certain amount

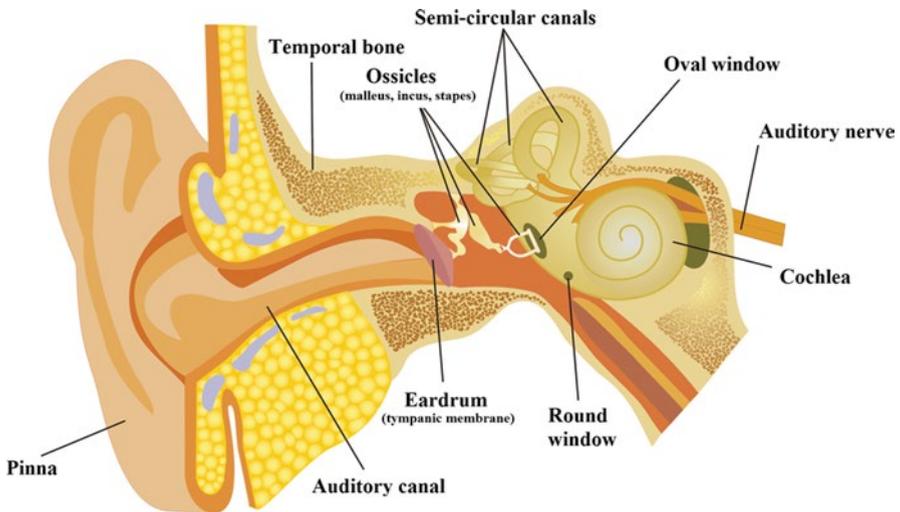


Fig. 2.4 General sketch of the outer ear, middle ear, and inner ear; illustrated here with the semi-circular canals, which are parts of the inner ear, but serve a function other than hearing, namely, the sense of balance (Figure by Leila Aazari)

of resistance. The fact of transmitting the vibrations from a large area, that of the eardrum, to a small area, that is the one at the base of the stapes, results in a significant increase of the pressure and allows to transmit effectively the information provided by the air vibrations. The main role of the middle ear is therefore to contribute to the production of this pressure when the vibrations are entering into the inner ear through the oval window.

Just below the oval window is the round window (or cochlear window). This is part of the inner ear, but its function is closely linked to the activity of the oval window. With the fluid in the ear being incompressible, any pressure on the oval window has to be absorbed elsewhere, which is made possible by the round window which is actually an elastic membrane.

Other structures that are parts of the middle ear contribute directly to the functioning of hearing. A structure called the Eustachian tube (or internal auditory meatus) connects the middle ear to the pharynx and to the nose or mouth. Its role is to make the air pressure in the middle ear equal to that existing in the ear canal. It is possible to notice the need to equilibrate that pressure when climbing in altitude or traveling in an airplane, which is made possible by swallowing or yawning. The air may then be sent from the pharynx to the middle ear, which enables the eardrum to vibrate normally.

Two muscles also have a key role in modulating the transmission of sound energy to the inner ear. One is called the tensor tympani muscle and the other is the stapedius muscle. They allow the release of the stapes from the oval window. The function of these two muscles is to protect the auditory system when sounds are too intense. Thus, while the middle ear is built so as to overcome the resistance of the liquid medium of the inner ear by increasing the pressure, it also has a security system, when sounds are too loud, for reducing the transmission of these sounds. The contraction of these two muscles is reflex activity, namely, the acoustic reflex or attenuation reflex.

The inner ear, also called the labyrinth, contains a bone structure, the bony labyrinth. Inside the bony labyrinth, there is a structure, the membranous labyrinth, immersed in a liquid called the perilymph. In the inner ear, there are three main structures. The first, the cochlea, has a crucial role in hearing which is described in the next section and later in the chapter. The other two structures, the vestibule and semicircular canals, have a key role in balance, but this topic will not be covered in this book.

Note that it is possible to hear without using the normal path of the outer ear and the middle ear. Vibrations can be conducted by the bones of the skull, a phenomenon called bone conduction. For experiencing the effect of conduction, just make a sound continuously and then cover your ears (while maintaining the sound). You will hear that the pitch of the sound is shifting. In fact, you will keep hearing the sound, even with plugged ears, but through bone conduction. This explains why we often feel we do not recognize our own voice on a recording. When we speak, we are hearing both the sounds that are transmitted via the outer ear and the middle ear and sound transmitted by bone conduction. The sound transmitted through bone conduction is not present when you hear a recording of your voice.

2.4.2 The Cochlea

The cochlea has the shape of a spiral, a kind of snail which is completing about two and a half turns. It contains a long membranous tube, the cochlear duct in which flows a liquid called endolymph.

Essentially, the cochlea is divided into three parts by two membranes (Fig. 2.5). Above the cochlear duct is the vestibular canal separated from the cochlear duct by a thin membrane called Reissner's membrane (or vestibular membrane). Below the basilar membrane, there is the tympanic canal in which flows, as is the case for the vestibular canal, the perilymph. Both canals communicate with each other through a narrow channel, the helicotrema.

When there are sound vibrations, they are transmitted to the perilymph. The fluid movement thus transmitted travels along the vestibular canal and returns to the tympanic canal. This movement then generates an oscillation of the basilar membrane which thus undergoes different deformations. The basilar membrane is narrower and stiffer at the base, close to the oval window and where the sound signals reach the cochlea, than on its apex.

It is actually on this basilar membrane that we find the spiral organ, also called the organ of Corti. In particular, it contains receptor cells that convert sound waves into action potentials. The organ of Corti is composed of thousands of hair cells. These cells, lying on supporting cells called Deiters' cells, each contain dozens of stereocilia. There are two types of hair cells, inner and outer. There are about 3500 inner hair cells in each ear, arranged on the same row, and more than 10,000 outer hair cells, arranged in three rows. Yet more than 90% of the 30,000 afferent fibers of the auditory nerve are connected with inner hair cells, whereas approximately 500 efferent fibers (from the brain) of the auditory nerve are connected to the outer hair cells. When the basilar membrane oscillates, it is the contact of the stereocilia with the tectorial membrane, which is located just above the sensory cells, that is the

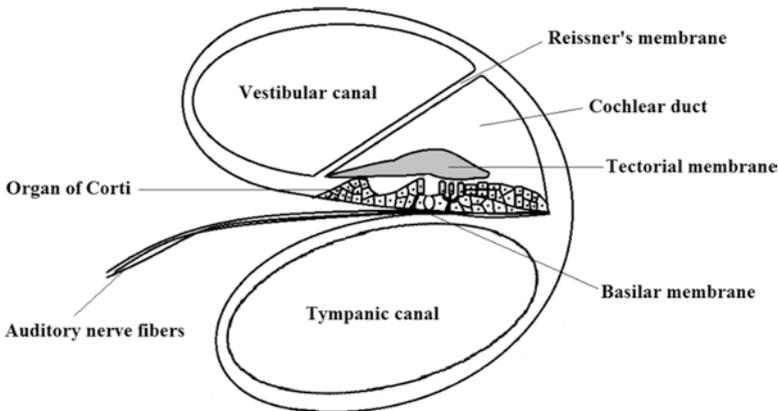


Fig. 2.5 Cross section of the cochlea

basis of hearing. It is at this point that all the mechanical vibration (first in air and then in the liquid medium of the inner ear) is converted into an electrical signal, nerve impulse, that the brain can recognize.

2.4.3 Central Mechanisms

Because they involve many crossings and relays, pathways bringing information from the cochlea to the auditory cortex are relatively complex. Auditory information enters into the brain at the bulb level (see Appendix C). The nerve impulses travel from the spiral ganglia to the brain structures by the vestibulocochlear nerves (the eighth cranial nerve) which split into two branches. In a given ear, the information is routed to the ventral and dorsal parts of the cochlear nucleus. From the cochlear nucleus, different routes can be followed. The neurons of the ventral part will make connection with the superior olivary nucleus, one-half traveling in the opposite half of the brain (contralateral side) and the other half remaining in the ipsilateral side. Early in the auditory system, at the olivary level (at the level of the medulla oblongata), there is a representation of the activity of both ears on each side of the brain.

The axons of the neurons in the dorsal cochlear nucleus all reach the *inferior colliculus* (at the level of the midbrain) on the contralateral side. Information from the superior olivary structure reaching the inferior colliculus originates from both the left ventral cochlear nucleus and the right ventral cochlear nucleus. Note that some fibers from the contralateral superior olivary structure and some fibers from the dorsal cochlear nucleus will transit through the nucleus of the medial lemniscus before reaching the inferior colliculus; moreover, at the level of this latter structure, many nerve fibers are crossing.

The nerve impulse is then routed to the thalamus, more specifically at the *median geniculate nucleus*, for eventually arriving at the *primary auditory cortex*, or A1, in the temporal lobe. Note that there are some relays between the inferior colliculus and the superior colliculus where would be processed the information about the location of a sound source, along with information from other sensory modalities. Finally, it should be noted that in A1, there is a *tonotopic organization*, i.e., a spatial representation of the different sound frequencies. In fact, this organization exists at all stages of auditory information processing described above.

2.5 Theories of Hearing

The previous section allowed to learn the role of different biological structures in the path of the sound wave from the pinna to the auditory cortex. However, the question remains as to how these waves can afford to hear with so many nuances. Researchers have long addressed this yet simple question: how can we perceive

pitch? What is happening exactly on the basilar membrane, in the organ of Corti? The next subsections provide an overview of the main answers revealed by research in the field of hearing.

2.5.1 Frequency Theory

The initial theoretical explanation based on the idea of the frequency was proposed by the English physiologist William Rutherford. In the past, telephones were built with a diaphragm, and it was the vibrations of this device, caused by the voice, that were converted into electrical signals. Once reaching the acoustic of another telephone, the signals were reproduced. Rutherford tried to draw a parallel between the basilar membrane and the diaphragm. According to him, the basilar membrane would serve to reproduce the pressure variations transmitted by the stapes. From such a perspective, the auditory nerve serves as a transmission cable, and the role of the brain is to interpret the frequency.

This formulation of the frequency theory was not going to hold the road. The basilar membrane is not like the diaphragm of a telephone was. The basilar membrane is not of the same width throughout and rigidity changes from one place to another. An even more serious objection to the original frequency theory is the simple fact that the ear is sensitive to frequencies ranging up to 20 kHz. This implies that a nerve fiber would have to be able to send 20,000 impulses per second. In fact, even the transmission of sound of 1000 Hz is problematic because a nerve cell cannot produce 1000 impulses per second. In short, this theory cannot account for the perception of all pitches associated with the audible frequency range. In other words, understanding the perception of high frequencies causes problem.

One solution to this problem was proposed by Wever and Bray (1937). This solution, which is based on the idea of cooperation between the nerve fibers, is called the *volley principle*. According to this principle, the neural activity associated with each of the different cycles of a sound is distributed via a series of fibers. Each fiber does not have to respond to every cycle of a sound wave. After a response to a cycle, a fiber has a recovery period and another fiber responds to the next cycle (Fig. 2.6). Indeed, a large number of fibers share the work. It is the grouped activity on a set of fibers that captures all cycles of a given sound wave. Finally, the volley principle accounts not only for the perception of pitch but also for that of loudness. The perception of loudness is accounted by the combined activity of more than one fiber for each cycle of the sound wave.

In fact, we now know that the activity of an auditory nerve fiber is generated when, in a given cycle, the wave is at its highest pressure level. So there is synchronization between the pressure change caused by a stimulus and the beginning of nerve activity. This phenomenon is called *phase locking*. Moreover, a neuron does not have to trigger its activity in each cycle, but when it does, it always happens at the same point in the cycle. This phenomenon also means that there is in the auditory nerve fiber a temporal code related to a sound wave. Due to the refractory

period required for each fiber of the auditory nerve, the temporal coding begins to be a little less reliable for frequencies above 1000 Hz and becomes virtually useless with frequencies above 5000 Hz.

2.5.2 Theories Based on Location

The idea of associating the processing of auditory information with a particular place on the basilar membrane is not new. Already in the nineteenth century, the German physiologist Hermann von Helmholtz proposed a theory of “the place of resonance” to explain the perception of pitch. Knowing that the width of the basilar membrane is not the same everywhere, he believed that, at a given location, the membrane, due to its width, would give a sound of a particular pitch, just like the strings of a piano, being of different lengths, give different notes. The analogy with the piano was proved to be incorrect, but the idea of linking the pitch to a specific place on the basilar membrane remains relevant. It is the basis of the place theory: there is indeed a tonotopic organization of hair cells in the organ of Corti. In other words, there is a spatial coding of frequency. Some frequencies are processed at specific locations on the basilar membrane.

Nobel Prize laureate in Physiology and Medicine in 1961, physicist Georg von Békésy described the mechanics inside the cochlea that underlies this spatial encoding. As we have seen earlier, the basilar membrane is narrow and rigid at the base of the cochlea, and, closer to its apex, it gradually widens and becomes less rigid (Fig. 2.6). Thus, when the stapes transmits the vibrations within the inner ear, this causes a hydrodynamic movement. The sound wave is thus propagated from one end of the basilar membrane to the other. This wave motion along the membrane constitutes the traveling wave.

The maximum point of displacement of the wave depends on its frequency. Indeed, this maximum point, that is the point where the basilar membrane is the more curved (Fig. 2.7), is nearest to the helicotrema if the frequency is low. The



Fig. 2.6 Illustration of the volley principle, with all fibers (A–F) combined on the *bottom curve*

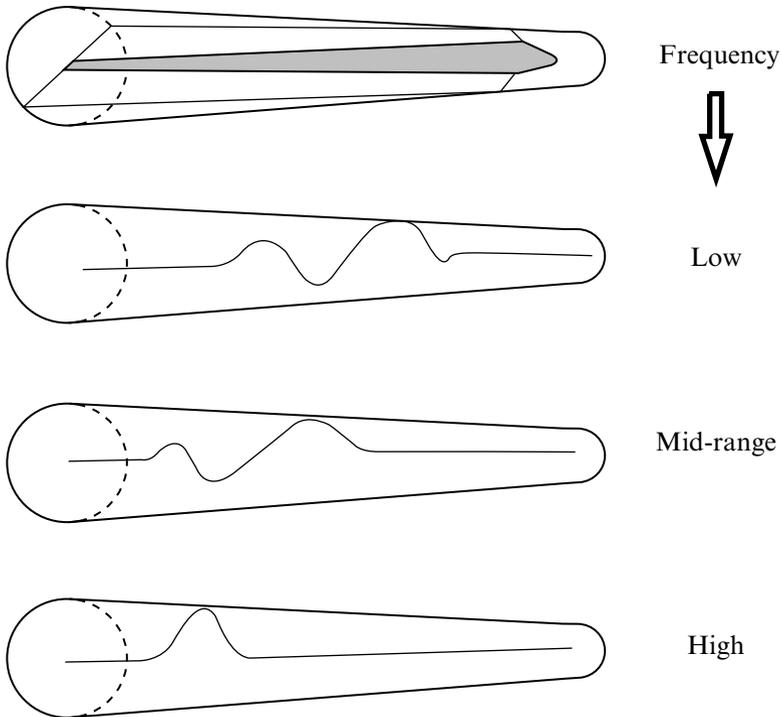


Fig. 2.7 On *top*, a representation of the basilar membrane (in *gray*) when the cochlea is unrolled; *bottom figures* illustrate the different points of maximum displacement of the traveling wave as a function of the sound frequencies

wave rapidly reaches its maximum amplitude and then quickly disappears. Conversely, the maximum point of displacement is reached farther away on the basilar membrane if the frequency is high. It is where the membrane is the most curved that the hair cells are the most displaced and generate the highest stimulation. The waves of different frequencies therefore will have their maximum impact on different parts of the basilar membrane, and the auditory nerve fibers stimulated will transmit their specific information to the auditory cortex.

This explanation of von Békésy based on the idea of a traveling wave allows not only to understand the perception of pitch but also the perception of loudness. Loudness would indeed depend on the magnitude of the traveling wave. Greater sound intensity provokes larger movement amplitudes on the basilar membrane. Larger amplitude affects more hair cells and produces greater inclination of these cells; therefore, larger amplitude results in more neural activity.

Note in closing this section that the frequency theory (volley principle) and the place theory (traveling wave) are both accepted. It is generally recognized that for low frequencies, the volley principle applies (frequency coding), and for high frequencies, the traveling wave hypothesis applies (spatial coding).

2.6 Clinical Aspects

Some breaks in the transmission sequence of the sound wave from the eardrum to the auditory cortex can cause hearing damage. Besides the fact that some diseases can be caused by damage to central auditory pathways or different regions of the auditory cortex—sometimes called *central deafness*—we generally distinguish two categories of hearing loss, depending on the place where the deficit is caused.

A first category of hearing loss is related to *transmission* problems (or *conduction*). Essentially, this type of disorder is mechanical, i.e., the sound wave is not transmitted efficiently to the cochlea. The causes of such a condition are therefore located at the outer ear or in the middle ear. These causes range from an excessive accumulation of wax to the deterioration of the ossicles. Similarly, throat infections, connected by the Eustachian tube to the middle ear, may interfere with the pressure balance in the middle ear and thus reduce the quality of transmission of the sound wave.

The second type of hearing loss is referred to as *sensorineural* (or *perceptive deafness*). This problem is caused by deterioration of the cochlea or of the auditory nerve. This deterioration occurs for various reasons such as metabolic problems or trauma. Some medications with toxic properties may also cause this kind of disorder.

Still about sensorineural hearing loss, it is most relevant to know that this deficit may occur as a result of deterioration of the hair cells located on the organ of Corti in the cochlea. Such deterioration is irreversible and can be caused by exposure to sounds of high intensity. The stronger the sounds—especially if you are close to the sound source—the less exposure time it takes to incur permanent damage. Therefore, there is a high price to pay when we offer ourselves this wonderful luxury of listening to loud music, often directly from the source using headphones!

If you are exposed, for instance, to sounds of approximately 85 dB, about 8 h per day, at some point you will affect your hearing. Exposure to loudness causes hearing fatigue, i.e., a shift of the detection threshold for a given period. The effects are the same, for example, (1) with an exposure to 88-dB sounds for 4 h per day or (2) with an exposure to 100-dB noise for 15 min per day. However, repeated exposure to even louder sounds may cause permanent threshold shift. Note that a loud sound might sound weaker after a few minutes of exposure. This phenomenon is called auditory adaptation.

The hearing abilities change with age. Indeed, age causes a loss of hearing called *presbycusis*. In particular, as we get older, the detection threshold for high frequencies becomes much higher. Consequently, it is possible for young persons to receive an audible signal indicating the arrival of a text message on their phone without an adult of a certain age (e.g., a teacher!) hearing it. It is unlikely that an adult over 40 years will hear a sound above 15 kHz or an adult over 50 years will hear a sound above 12 kHz. High frequencies have even been used to get rid of noisy teenagers loitering in a schoolyard.

Finally, among the quite severe disorders connected somehow to hearing, there is *tinnitus*. This problem consists of an impression that a sound or noise is present,

even in the absence of any external auditory stimulation. Tinnitus can sound like whistling or rustling and can be caused by several factors. The sound may be continuous or intermittent and is usually rather acute. Tinnitus can indicate the presence of a hearing disorder caused by damage to the cochlea, for example, or occur after noise trauma or during an infection.