Anatomy of the Human Ear

Antti Hepo-oja
Aalto Universtiy
Master's Programme AAT

`antti.hepo-oja@aalto.fi`

Abstract

The human hearing is a collaboration of several complex structures. The human auditory system is able to identify individual tones in a mix of sounds, and it has a unique system that can actively amplify sound signals with a low input level. The human ear can be roughly divided into three parts: the outer (or external) ear, the middle ear, and the inner ear. The role of the outer ear is passive as the pinna funnels the incoming sound waves into the ear canal. The pinna also helps with source localization together with binaural hearing. The middle ear has an important function at impedance matching as the high velocity sound waves are transmitted into liquid-filled inner ear through the collaboration of the tympanic membrane, the ossicles, and the oval window. The inner ear houses the most complex system of the human ear: the cochlea. The cochlea powers the frequency selectivity and active boosting of the ear. The basilar membrane vibrates as the result of incoming vibrations from the middle ear, and the vibrations are transmitted to hair cells. Finally, the hair cells encode the vibrations into neural signals that are processed into hearing sensation.

Keywords — Pinna, Ear canal, Tympanic membrane, Ossicles, Cochlea, Basilar membrane, Inner hair cells, Outer hair cells

1 Introduction

The human auditory system monitors a surprisingly large amount of information from the environment. It is an impressive collection of functionalities that are able to encode complex sounds into narrow-band spectrum.

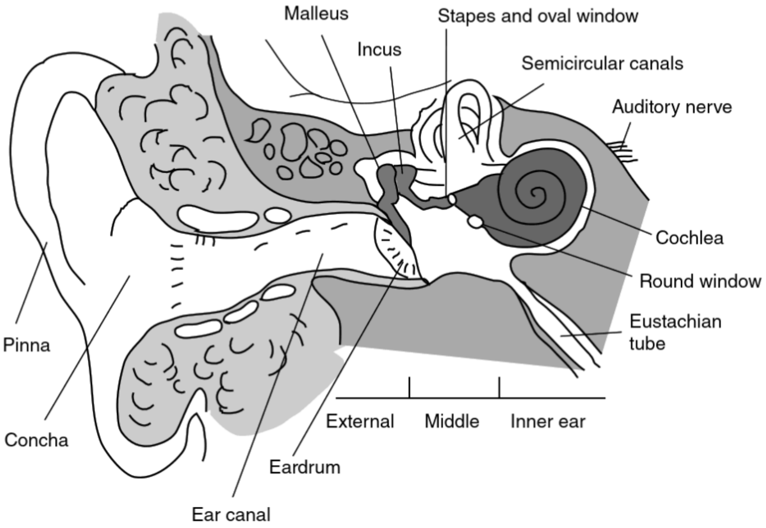

Figure 1: An overview of the human ear [4]. The three sections of the ear are visible: the external, the middle, and the inner ear.

This spectrum holds real-time information about the position of the sound source, and enables us to distinguish tones within a mix of sounds.

Binaural hearing allows us to quite accurately localize sound sources within three dimensional space. With the help of the interaural time difference and the interaural level difference, we are able to resolve the position of the source on the horizontal plane. In addition, the external ear (together with the head) causes the head shadowing effect that allows us to resolve the elevation of the source. Interestingly enough, people with monaural hearing have been observed to resolve the position on the horizontal plane solely with the head shadowing effect. [6]

In general, the human hearing ranges from 20 Hz to 20 kHz, and it is able to distinguish different frequencies at a frequency resolution of about 3.6 Hz [4]. The human hearing has a surprisingly large dynamic range of up to approximately 160 dB [3]. At the same time, we are able to hear very faint sounds, as the cochlea is able to amplify the sound signals at a factor of several hundreds. [4]

This literature study aims to describe the anatomy of the ear in a clear and comprehensive manner. The most important parts of the ear are included in the study, such as the function of the pinna, the ossicles, the cochlea, and the hair cells. In addition, the relevance to headphones is considered where appropriate i.e. in the external ear. Although the functionalities of the cochlea are discussed more extensively, the role of the brain and the nerve system in the human hearing has been left out on purpose as it is too large a topic to discuss.

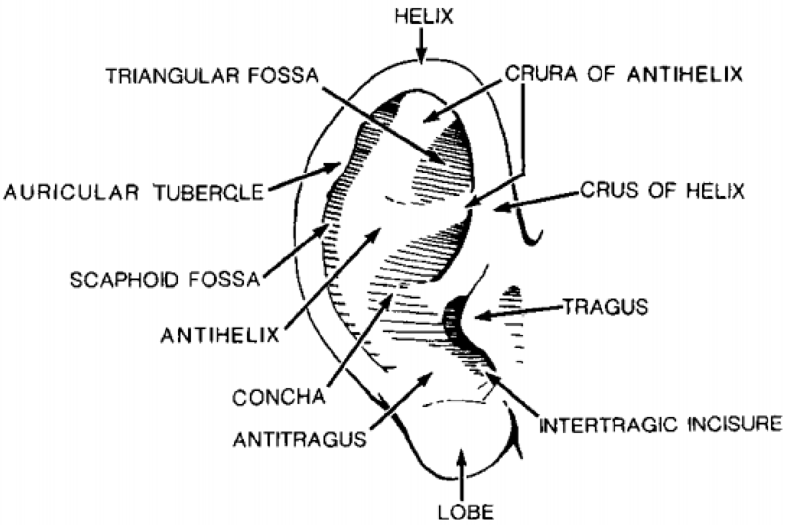

Figure 2: The structure of the pinna [1]

2 Outer ear

The outer ear is considered to include the *pinna*, or the auricle, and the ear canal, or the external auditory meatus. The pinna is the visible part of the ear that is oriented at an approximate angle of 30 degrees [6], whereas the ear canal connects the external ear with the middle ear. The middle ear and the external ear are separated by the tympanic membrane, or the ear drum. Although the outer layer of the tympanic membrane is considered to be a part of the external ear, the anatomy and the function of the eardrum is discussed in next section.

The function of the outer ear is purely passive; the active and non-linear properties of human hearing are manifested within the inner ear [4].

2.1 Pinna

The pinna is a system of several functional pats that are primarily formed of cartilage. These parts include the *helix* and the Y-shaped *antihelix* with its superior and inferior *crura*, the lobe, the *tragus* and the *antitragus*, the *concha* that is located below the antihelix, and the hollow *scapha* between the helix and the antihelix [1]. The structure of the pinna is depicted in Figure 2.

The pinna serves an important function in determining the elevation of the sound source. The structure of the pinna imposes changes to the sound spectrum at high frequencies according to the direction of the sound source [4, 6]. These changes, or cues, are encoded information about the position of the sound source. It has also been shown that while changes to the pinna immediately deteriorate the elevation localization, humans are able to adapt to these changes. [6] In addition to determining the elevation, the external ear emphasizes frequencies ranging from 1 kHz to 5 kHz [4].

Generally, sound localization on the horizontal plane relies on the interaural time difference (ITD) and the interaural level difference (ILD), and the azimuth localization is unaffected by the changes to the pinna. However, pinnae can also be used for the localization of the sound source on the azimuth, or the horizontal plane. Although these cues are usually ignored in presence of ITD and ILD, it has been proposed that people with monaural hearing are able to locate sound sources on the horizontal plane with the aid of the head shadowing effect (HSE). Quite expectedly, the performance of azimuth localization with the HSE is worse than with the ITD and the ILD as the HSE is an ambiguous signal on the horizontal plane. [6]

The lack of binaural cues (i.e. azimuth cues) are also an issue that needs to be considered with headphones. There is practically no cross-talk between the two ears, and the spatial effect of the listening environment is lost. With headphones, this has been resolved with stereo-widening: a technique where the interaural level difference and the interaural time difference have been implemented artificially by adding cross-talk with appropriate gains. [5]

The lack of azimuth cues is not the only problem. With ideal loudspeakers, the frequency response is flat when measured next to the speaker. When measuring the same signal in the ear, a distinct curve can be seen in the frequency response. This effect is called the head-related transfer function. Consider the typical circum-aural or supra-aural headphones; the audio elements are next to the ear canal, and there is no head shadowing effect present from the pinnae or the head. As a result, the head-related transfer function, and subsequently the elevation cues, are lost. As the flat frequency response does not sound natural, the headphones are tuned to imitate the head-related transfer function at the eardrum. [5]

2.2 Ear canal

The ear canal is partly formed of cartilage that connects the pinna to the ear canal. The rest of the ear canal, approximately two-thirds, is formed of bone [1]. The length of the ear canal in total is approximately 22.5 mm with a diameter of 7.5 mm [4]. It is also connected to sweat glands that produce anti-bacterial cerumen that acts as a lubricant to the skin. The consistency and the shape of the ear canal changes throughout the life, and for the adults, the ear canal has formed into a shape with two bends. When doing any research or measurements that include the ear canal, it is important to understand that the skin within it is prone to injuries due to its thinness and the lack of flexible fatty layer. [1]

Due to the open-ended acoustics, the ear canal manifests properties typical to quarter-wavelength resonators. The ear canal increases noticeably the level of sound signals at frequencies 3-4 kHz and slightly above 10 kHz. In contrast, it decreases the sound level at the frequency range of 7-8 kHz. [4] Consider the issue of losing the head-related transfer function with the circum-aural headphones. Even if the head-related transfer function is imitated in the design of the headphones, the natural resonance imposed by the earcanal is still intact. However, with in-ear headphones, this natural quarter-wavelength resonance is lost as the plugged earcanal becomes a half-wavelength resonator. Instead, a half-wavelength resonance is imposed to the frequency response. [5]

3 Middle ear

The middle ear is the part of the human ear filled with air that is located between the eardrum and the inner ear. It acts as a transducer that converts the mechanical vibrations of the eardrum, to the vibrations of the ossicles of the middle ear, and finally to the pressure changes in the fluid within the inner ear. In addition, one end of the *Eustachian tube* is located in the middle ear, and the other end of the tube is connected to the nasopharynx above the oral cavity. The function of the Eustachian tube is to match the pressure within the middle ear with the environment. A mismatch between the pressures causes reduced hearing as the eardrum is displaced. [4]

3.1 Eardrum

The eardrum is an elliptic, thin membrane with an area of approximately 60-70 mm2. It is oriented at an approximate angle of 40 degrees. It has a very low mass (approximately 14 mg), and hence moves very easily. The eardrum itself can be divided into several parts: the *pars flaccida* and the *pars tensa*. True to its name, the pars flaccida, or *Shrapnell's membrane* is a flaccid membrane that lies on the lateral process of the malleus. Below the pars flaccida lies the malleal folds, and in the center of the eardrum is the *manubrium* of the malleus that is attached to the eardrum. The rest of the eardrum has coined pars tensa which is thicker than the pars flaccida. The bony tympanic ring surrounds the eardrum with the notch of Rivinus that surrounds the pars flaccida. [1] The structure of the eardrum is depicted in Figure 3.

Although the structure of the eardrum may seem quite complex, its function is quite straightforward. The eardrum vibrates as a result of local pressure changes from sound waves in the ear canal [4]. As the eardrum is connected to the *malleus*, which is one of the three ossicles [6], the ossicles will vibrate simultaneously with the eardrum.

The eardrums reflect a part of the sound energy back to the ear canal. The headphone industry has found a use for these reflections. The reflections together with the otoacoustic emissions can be recorded and analyzed, and an individual target curve can be created for the headphones. As a result, the pair of headphones will automatically adapt to the listener's ears, making portable studio quality sound available to professionals. [2]

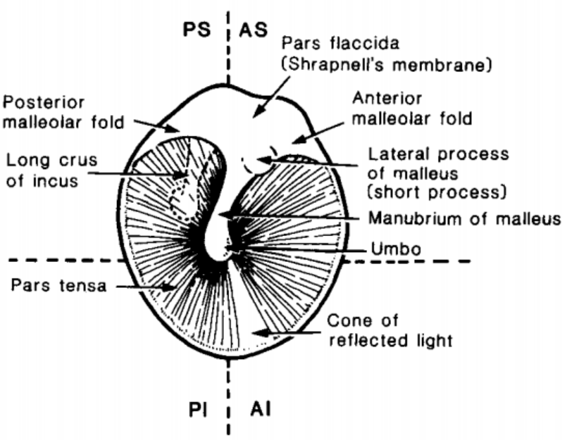

Figure 3: The structure of the middle eardrum [1]

3.2 Ossicles

The ossicles are three small bones within the middle ear that are used for impedance matching between the air within the ear canal and the liquid within the inner ear. These ossicles are the malleus, the *incus*, and the *stapes*, also called as the hammer, the anvil, and the stirrup, respectively. The links between the ossicles also contribute to improving the transmission of sound energy, but the majority of the impedance matching is done at the interface between the stapes and the oval window that connects to the fluid in the inner ear. At this interface, the small pressure and high velocity of the air is transformed into the high pressure and slow velocity of the liquid in the inner ear. [4] The structure of the middle ear is depicted in Figure 4.

Without the presence of the ossicles, a large amount of the energy associated to sound waves is reflected back due to the high difference in the impedances of the two media [4]. The middle ear also has the ability to protect, to some extent, the inner ear from loud sounds. At high input levels, the muscles around ossicles contract and, subsequently, the transmission efficiency of the ossicles is reduced. Unfortunately, this effect, coined as acoustic reflex, is not fast enough to counter impulsive sounds. [4]

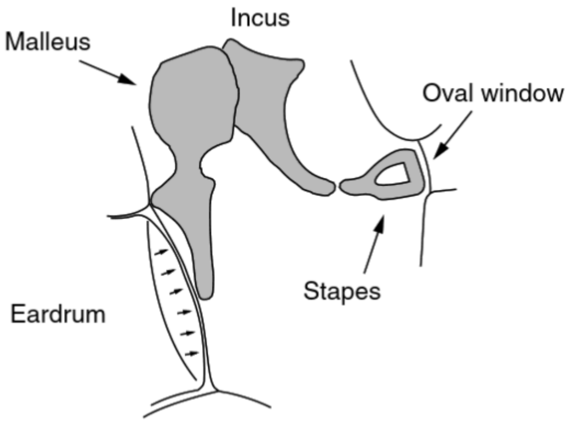

Figure 4: The structure of the middle ear [4]

4 Inner ear

The most important and definitely the most complex structure of the ear is the cochlea, located within the inner ear. The cochlea is a spiral structure which also hosts the cochlear hair cells that are located in the organ of Corti which is extended along the cochlea [3]. In addition, the fluid-filled semicircular ducts are located within the inner ear. These canals detect rotary motion in three dimensions i.e. they are used to induce a sense of balance. [4]

4.1 Cochlea

The structure of the cochlea is shown in Figure 5. The cochlea is filled with liquid partitioned by the bony shelf into two ducts: the *scala vestibuli* and the *scala tympani*. These ducts are connected to the middle ear by the oval and round windows, respectively. The oval window is the interface that connects the staple with the inner ear. The two ducts connect at the apex of the basilar membrane through the *helicotrema*. [4]

Along the cochlea, the *Reissner's membrane* separates the liquid-filled scala media from the scala vestibuli. In contrast to the scala vestibuli and the scala tympani, the liquid within the scala media has a higher ion concentration due to excess amount of potassium ions. The basilar membrane is located between the scala media and the scala tympani, with a potential difference due to different ion concentrations. [4]

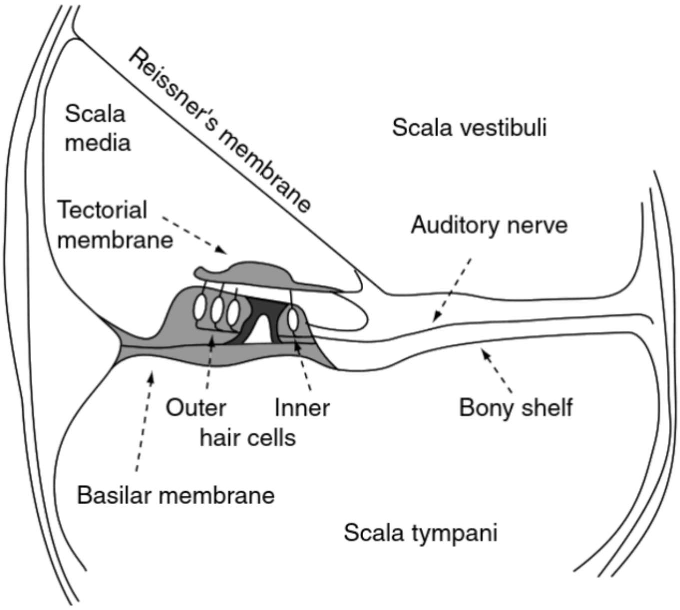

Figure 5: The structure of the cochlea [4]

4.2 Function of the basilar membrane

As the stimuli from the sound waves are transmitted into the fluid in the cochlea, the basilar membrane will vibrate as a result of the pressure difference between the two ducts. The vibrations of the basilar membrane are non-linear, meaning that some frequencies resonate more efficiently than others. [4] On the contrast, the vibrational energy associated to higher frequencies are dissipated more efficiently by the viscosity of the surrounding fluid than at lower frequencies [3].

One interesting feature of the cochlea is the existence of cochlear traveling waves as the stiffness of the basilar membrane changes along the membrane [3]. This also means that the characteristic frequency of the membrane also changes along the membrane. The vibrations that enter the oval window propagate towards the round window through the basilar membrane. However, due to changes in the characteristic frequency, the place where the vibrations cross the basilar membrane also varies. [4]

As a result, the basilar membrane does not vibrate as a whole, but seemingly traveling waves are formed at the base of the cochlea [3, 4]. In other words, these waves decrease in wavelength and increase in amplitude as they progress. The location of maximum amplitude is determined by the frequency of the input signal; the position of the highest frequencies are near to the base of the cochlea. Therefore, each frequency has their relative position on the basilar membrane, and this information is also encoded real-time by their appropriate hair cells to neural activity. [3]

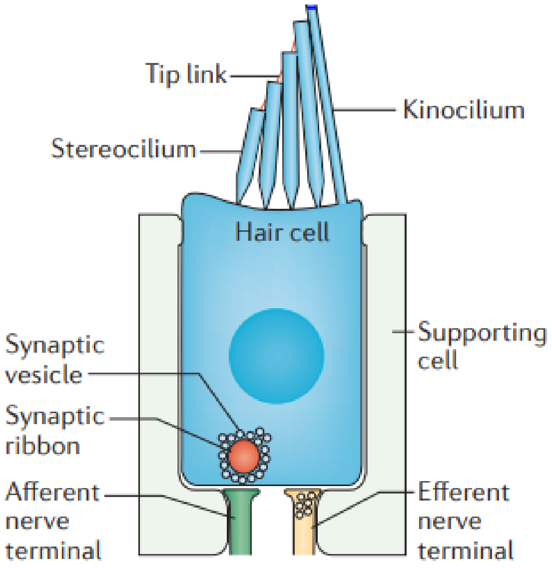

Figure 6: The structure of a hair cell [3]

4.3 Hair cells

The organ of Corti that hosts the hair cells is located on the basilar membrane within the cochlea [4]. The structure of a hair cell is depicted in Figure 6. *Stereocilia*, which are related to microvilli, are located on top of each hair cell. Each stereocilium is linked to adjacent stereocilia through a *tip link*. As the stereocilia bend due to shearing forces, potassium ions may enter the stereocilium. [3]

The cochlear hair cells can be divided into two groups: the inner and the outer hair cells.There are approximately 3500 inner hair cells in one row and 12000 outer hair cells distributed over three rows. The inner hair cells are connected to several (10-30) afferent cochlear nerves. [3] The neural information is routed through auditory pathway and passed to the auditory cortex [4].

4.3.1 Inner hair cells

The inner hair cells are responsible for encoding the mechanical vibrations into neural signals. As the basilar membrane vibrates, the hair cells (which are located on the organ of Corti) start bending. The resulting voltage difference caused by potassium ions within the stereocilia triggers the release of glutamate that serves as a neurotransmitter. [3, 4] Due to the existence of cochlear traveling waves, different frequencies are encoded by different hair cells to their respective afferent cochlear nerves [4], ultimately giving rise to the frequency selectivity of the human hearing.

The cochlear hair cells have quite a unique function as a part of the human hearing. They do not only passively encode the vibrations of the basilar membrane into neural activity, but they are also able to actively boost the vibrations of the basilar membrane. This active boosting can be captured as otoacoustic emissions from the ear. In addition, hair cells are able to boost certain frequencies and to dynamically change the level of amplification based on the input level. Due to the active properties of the human hearing, we are able to distinguish different frequencies from a mix of sounds. [3] The gain caused by this amplification is more prominent at low input levels and at frequencies around 8 kHz. The gain is decreased as the input level is increased, and, simultaneously, the frequency where the largest gain is achieved is shifted to lower frequencies. [4]

4.3.2 Outer hair cells

This active amplification of the vibrations of the basilar membrane is mainly done by the outer hair cells [4]. When the viscosity dissipates the mechanical energy associated to the vibrations of the basilar membrane, the outer hair cells can actively boost the vibrations by oscillating their length. The oscillation is caused by the piezoelectric proteins that change the potential within the hair cell through a complex process. The oscillations are maintained with the transmembrane electrical field. The level of boosting provided by the outer hair cells is changed dynamically to maintain stability, but the nature of this functionality is not known. [3]

5 Conclusion

The human ear has several individual functions. The anatomy of the ear is divided into three main parts: the outer ear, the middle ear, and the inner ear. The pinna of the outer ear serves the purpose to determine the elevation, i.e. the up-down and the front-back position, of the sound source. The pinna acts as a funnel for the sound waves thanks to its ridges and grooves that efficiently direct sound waves of different frequencies to the ear canal. The ear canal is a quarter-wavelength resonator that connects the pinna to the tympanic membrane, or the eardrum.

The middle ear has a sophisticated system to ensure that the sound energy is transmitted efficiently into the inner ear through a mechanism called impedance matching. As the sound waves cause vibrations in the tympanic membrane, the three small bones, called ossicles, vibrate simultaneously. The muscles around the ossicles may contract at higher input sound levels, and the efficiency of the transmission is reduced, which subsequently protects the ear. As the last of the three linked ossicles, the staples, vibrates at the oval window, the high velocity vibrations are transmitted into high pressure vibrations in the fluid-filled cochlea.

In the cochlea, the sound waves propagate through the scala vestibular. The basilar membrane changes in stiffness, and subsequently in characteristic frequency, and sound waves with different frequencies will pass the basilar membrane at different positions. A seemingly travelling wave will occur on the basilar membrane where sound waves with different frequencies are translated into the oscillation of a certain part of the basilar membrane. These vibrations are encoded by their respective inner hair cells to their respective afferent cochlear nerves as the potassium ions enter the hair cells, triggering the release of neurotransmitter.

In addition, the inner ear is able to actively change the length of the hair cells to boost the vibrations of the basilar membrane. This boosting is done in particular by the outer hair cells. As a result, low input levels are amplified significantly. This boosting is reduced as the input level is increased to maintain stability, but the exact mechanisms of this functionality are unknown.

References

[1] Alvord, L. S., and Farmer, B. L. Anatomy and orientation of the human external ear, Dec 1, 1997.

[2] Bolaños, J. G. Lecture slides for the course ELEC-E5630 AAT Seminar: Headphone acoustic coupling to the ear and equalization methods for binaural reproduction, 2019.

[3] Hudspeth, A. J. Integrating the active process of hair cells with cochlear function. *Nature reviews. Neuroscience 15*, 9 (Sep 2014), 600–614.

[4] Pulkki, V., and Karjalainen, M. *Communication acoustics: an introduction to speech, audio, and psychoacoustics.* Wiley, Chichester, West Sussex, United Kingdom, 2015.

[5] Rämö, J. Lecture slides for the course ELEC-E5630 AAT Seminar: Headphones, 2019.

[6] van Wanrooij, M., and van Opstal, A. J. Contribution of head shadow and pinna cues to chronic monaural sound localization. *The Journal of Neuroscience 24* (2004), 4163–4171.