

Auditory Perception: How Your Brain Interprets Sound

Authored by
mohammad looti

June 16, 2026

RECOMMENDED CITATION

mohammad looti (2026). *Auditory Perception: How Your Brain Interprets Sound*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=38130>

The neuronal encoding of sound is the representation of auditory sensation and perception in the nervous system.

This article explores the basic physiological principles of sound perception, and traces hearing mechanisms from sound as pressure waves in air to the transduction of these waves into electrical impulses (action potentials) along auditory nerve fibers, and further processing in the brain.

Introduction

The complexities of contemporary neuroscience are continually redefined. Thus what is known now of the auditory system has changed in the recent times and thus conceivably in the next two years or so, much of this will change.

This article is structured in a format that starts with a small exploration of what sound is followed by the general anatomy of the ear which in turn will finally give way to explaining the encoding mechanism of the engineering marvel that is the ear. This article traces the route that sound waves first take from generation at an unknown source to their integration and perception by the auditory cortex.

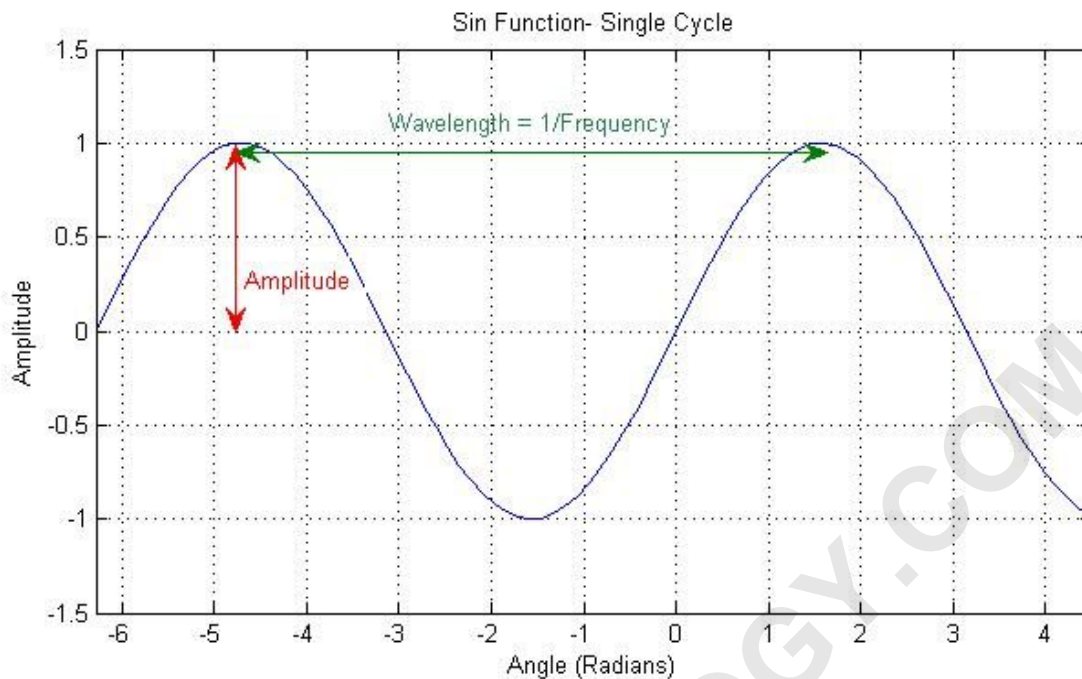
Basic physics of sound

Sound waves are what physicists call longitudinal waves, which consist of propagating regions of high pressure (compression) and corresponding regions of low pressure (rarefaction).

Waveform

Waveform is a description of the general shape of the sound wave. Waveforms are sometimes described by the sum of sinusoids, via Fourier analysis.

Amplitude



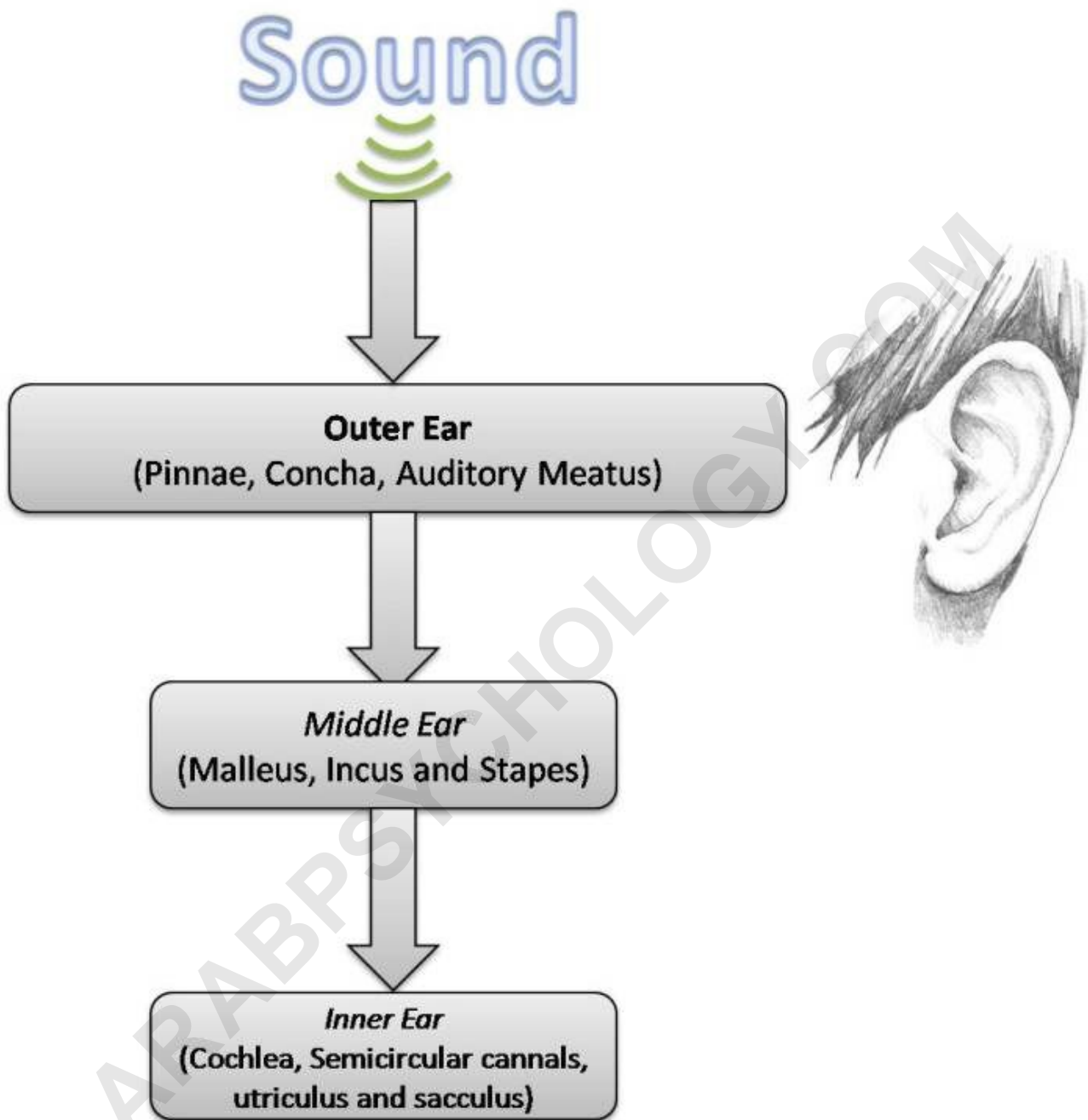
Graph of a simple sine wave

Amplitude is the size of the pressure variations in a sound wave, and primarily determines the loudness with which the sound is perceived. In a sinusoidal function such as $C \sin(2\pi ft)$, C represents the amplitude of the sound wave.

Frequency and wavelength

The frequency of a sound is defined as the number of repetitions of its waveform per second, and is measured in hertz; frequency is inversely proportional to wavelength (in a medium of uniform propagation velocity, such as sound in air). The wavelength of a sound is the distance between any two consecutive matching points on the waveform. The audible frequency range for young humans is about 20 Hz to 20 kHz. Hearing of higher frequencies decreases with age, limiting to about 16 kHz for adults, and even down to 3 kHz for elders.

Anatomy of the ear



Flowchart of sound passage - outer ear

Given the simple physics of sound, the anatomy and physiology of hearing can be studied in greater detail.

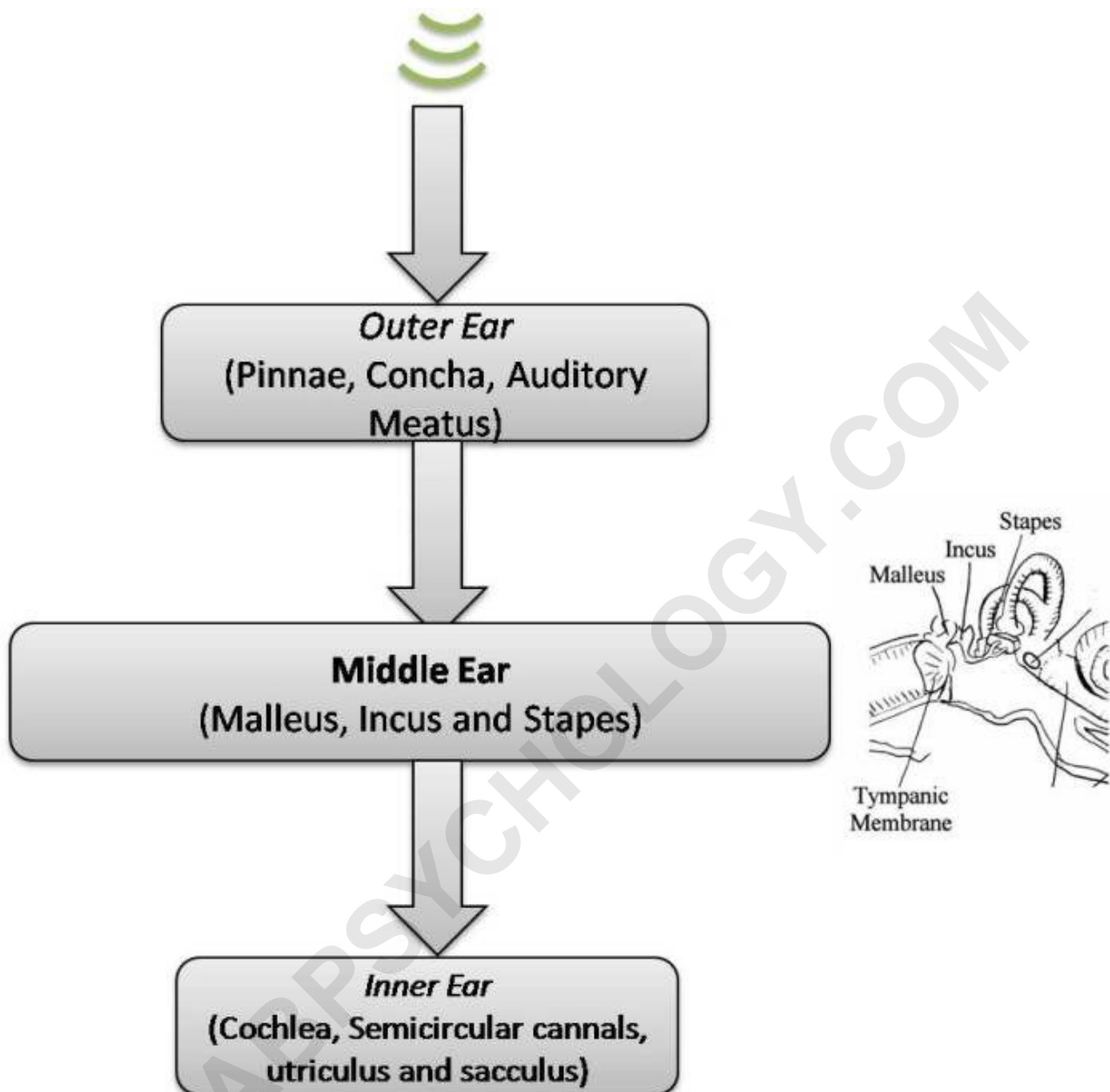
Outer ear

The Outer ear consists of the pinna or auricle (visible parts including ear lobes and concha), and the auditory meatus (the passageway for sound). The fundamental function of this part of the ear is to gather sound energy and deliver it to the eardrum. Resonances of the external ear selectively boost sound pressure with frequency in the range 2-5 kHz.

The pinna as a result of its asymmetrical structure is able to provide further cues about the elevation from which the sound originated. The vertical asymmetry of the pinna selectively amplifies sounds of higher frequency from high elevation thereby providing spatial information by virtue of its mechanical design.

Middle ear

ARABPSYCHOLOGY.COM

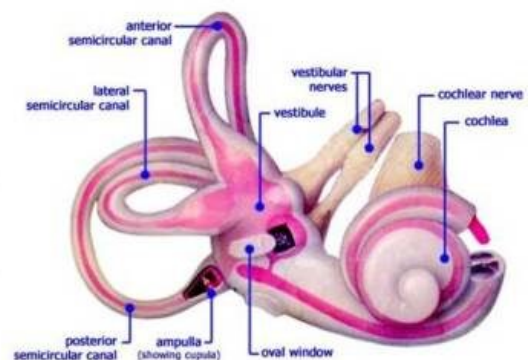
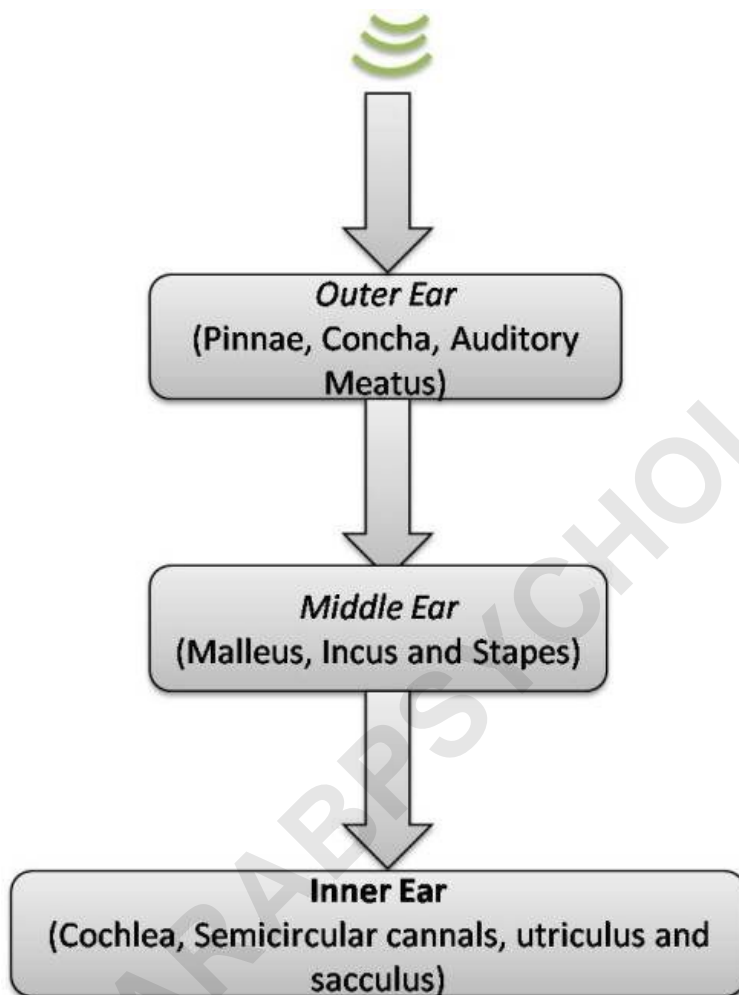


Flowchart of sound passage - middle ear

The middle ear plays a crucial role in the auditory process, as it essentially converts pressure variations in air to perturbations in the fluids of the inner ear. In other words, it is the mechanical transfer function that allows for efficient transfer of collected sound energy between two different media. The three small bones that are responsible for this complex process are the malleus, the incus, and the stapes, collectively known as the ear ossicles. The impedance matching is done through via lever ratios and the ratio of areas of the tympanic membrane and the footplate of the stapes, creating a transformer-like mechanism. Furthermore, the ossicles are arranged in such a manner as to resonate at 700-800 Hz while at the same time protecting the inner ear from

excessive energy. A certain degree of top-down control is present at the middle ear level primarily through two muscles present in this anatomical region: the tensor tympani and the stapedius. These two muscles can restrain the ossicles so as to reduce the amount of energy that is transmitted into the inner ear in loud surroundings.

Inner ear



Flowchart of sound passage - inner ear

The cochlea of the inner ear, a marvel of physiological engineering, acts as both a frequency analyzer and nonlinear acoustic amplifier. The cochlea has over 32,000 hair cells. Outer hair cells primarily provide amplification of traveling waves that are induced by sound energy, while inner hair cells detect the motion of those waves and excite the (Type I) neurons of the auditory nerve.

The basal end of the cochlea, where sounds enter from the middle ear, encodes the higher end of the audible frequency range while the apical end of the cochlea encodes the lower end of the frequency range. This tonotopy plays a crucial role in hearing, as it allows for spectral separation of sounds. A cross section of the cochlea will reveal an anatomical structure with three main chambers (scala vestibuli, scala media, and scala tympani). At the apical end of the cochlea, at an opening known as the helicotrema, the scala vestibuli merges with the scala tympani. The fluid found in these two cochlear chambers is perilymph, while scala media, or the cochlear duct, is filled with endolymph.

Transduction

Auditory hair cells

The auditory hair cells in the cochlea are at the core of the auditory system's special functionality (similar hair cells are located in the semicircular canals). Their primary function is mechanotransduction, or conversion between mechanical and neural signals. The relatively small number of the auditory hair cells is surprising when compared to other sensory cells such as the rods and cones of the visual system. Thus the loss of low number (in the order of thousands) of auditory hair cells can be devastating while the loss of a larger number of retinal cells (in the order to hundreds of thousands) will not be as bad from a sensory standpoint.

Cochlear hair cells are organized as inner hair cells and outer hair cells; inner and outer refer to relative position from the axis of the cochlear spiral. The inner hair cells are the primary sensory receptors and a significant amount of the sensory input to the auditory cortex occurs from these hair cells. Outer hair cells on the other hand boost the mechanical signal by using electromechanical feedback.

Mechanotransduction

The apical surface of each cochlear hair cell contains a hair bundle. Each hair bundle contains approximately 300 fine projections known as stereocilia, formed by actin cytoskeletal elements. The stereocilia in a hair bundle are arranged in multiple rows of different heights. In addition to the stereocilia, a true ciliary structure known as the kinocilium exists and is believed to play a role in hair cell degeneration that is caused by exposure to high frequencies.

A stereocilium is able to bend at its point of attachment to the apical surface of the hair cell. The actin filaments that form the core of a stereocilium are highly interlinked and cross linked with fibrin, and are therefore stiff and inflexible at positions other than the base. When stereocilia in the tallest row are deflected in the positive-stimulus direction, the shorter rows of stereocilia are also deflected. These simultaneous deflections occur due to filaments called tip links that attach the

side of each taller stereocilium to the top of the shorter stereocilium in the adjacent row. When the tallest stereocilia are deflected, tension is produced in the tip links and causes the stereocilia in the other rows to deflect as well. At the lower end of each tip link is one or more mechano-electrical transduction (MET) channels, which are opened by tension in the tip links. These MET channels are cation-selective transduction channels that allow potassium and calcium ions to enter the hair cell from the endolymph that bathes its apical end.

The influx of cations, particularly potassium, through the open MET channels causes the membrane potential of the hair cell to depolarize. This depolarization opens voltage-gated calcium channels to allow the further influx of calcium. This results in an increase in the calcium concentration, which triggers the exocytosis of neurotransmitter vesicles at ribbon synapses at the basolateral surface of the hair cell. The release of neurotransmitter at a ribbon synapse, in turn, generates an action potential in the connected auditory-nerve fiber. Hyperpolarization of the hair cell, which occurs when potassium leaves the cell, is also important, as it stops the influx of calcium and therefore stops the fusion of vesicles at the ribbon synapses. Thus, as elsewhere in the body, the transduction is dependent on the concentration and distribution of ions. The perilymph that is found in the scala tympani has a low potassium concentration, whereas the endolymph found in the scala media has a high potassium concentration and an electrical potential of about 80 millivolts compared to the perilymph. Mechanotransduction by stereocilia is highly sensitive and able to detect perturbations as small as fluid fluctuations of 0.3 nanometers, and can convert this mechanical stimulation into an electrical nerve impulse in about 10 microseconds.

Nerve fibers from the cochlea

There are two types of afferent neurons found in the cochlear nerve: Type I and Type II. Each type of neuron has specific cell selectivity within the cochlea. The mechanism that determines the selectivity of each type of neuron for a specific hair cell has been proposed by two diametrically opposed theories in neuroscience known as the peripheral instruction hypothesis and the cell autonomous instruction hypothesis. The peripheral instruction hypothesis states that phenotypic differentiation between the two neurons are not made until after these undifferentiated neurons attach to hair cells which in turn will dictate the differentiation pathway. The cell autonomous instruction hypothesis states that differentiation into Type I and Type II neurons occur following the last phase of mitotic division but preceding innervations. Both types of neuron participate in the encoding of sound for transmission to the brain.

Type I neurons

Type I neurons innervate inner hair cells. There is significantly greater convergence of this type of neuron towards the basal end in comparison with the apical end. A radial fiber bundle acts as an

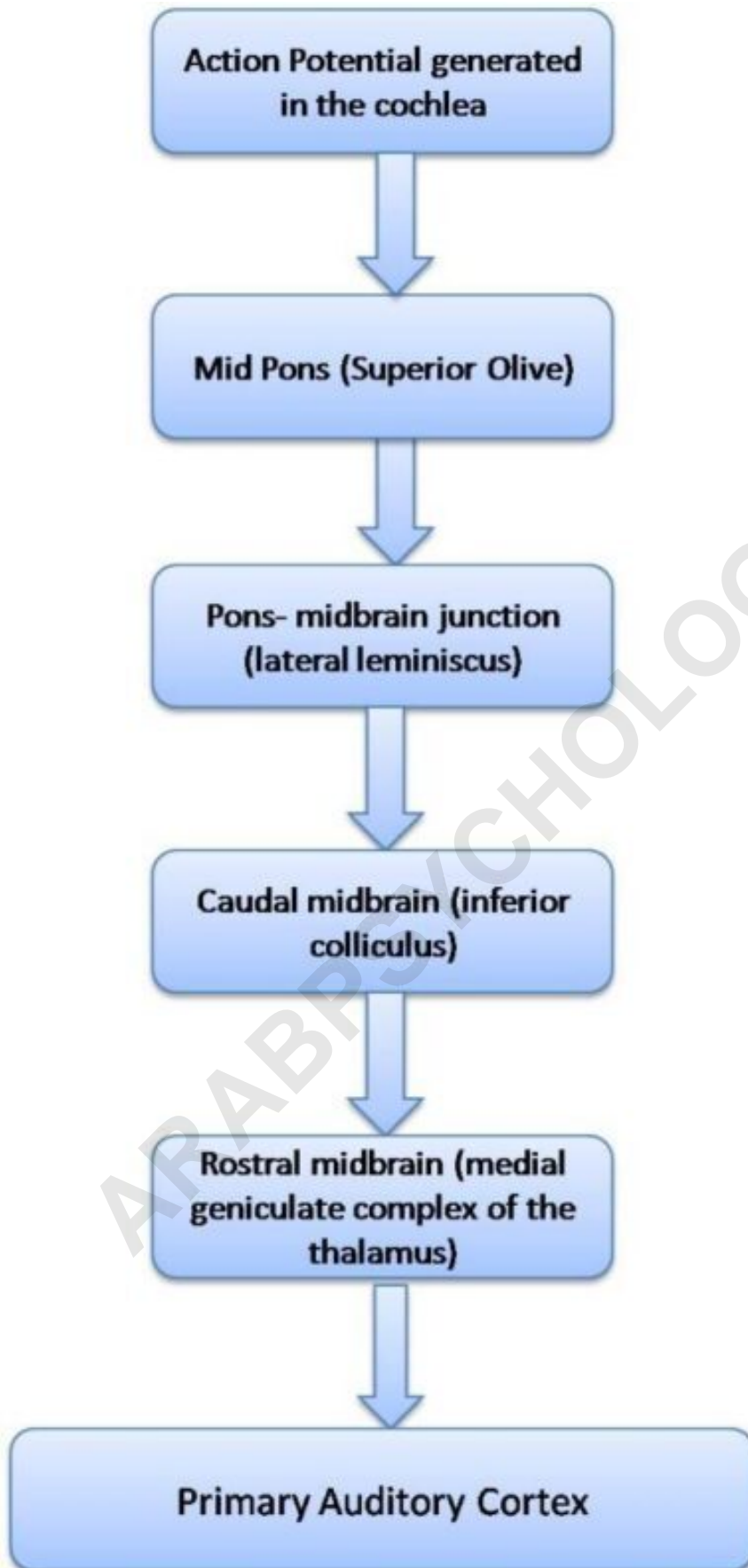
intermediary between Type I neurons and inner hair cells. The ratio of innervation that is seen between Type I neurons and inner hair cells is 1:1 which results in high signal transmission fidelity and resolution.

Type II neurons

Type II neurons on the other hand innervate outer hair cells. However, there is significantly greater convergence of this type of neuron towards the apex end in comparison with the basal end. A 1:30-60 ratio of innervation is seen between Type II neurons and outer hair cells which in turn make these neurons ideal for electromechanical feedback. Type II neurons can be physiologically manipulated to innervate inner hair cells provided outer hair cells have been destroyed either through mechanical damage or by chemical damage induced by drugs such as gentamicin.

Brainstem and midbrain

ARABPSYCHOLOGY.COM



Levels of transmission of neuronal auditory signals

The auditory nervous system includes many stages of information processing between the ear and cortex.

Auditory cortex

Primary auditory neurons carry action potentials from the cochlea into the transmission pathway shown in the adjacent image. Multiple relay stations act as integration and processing centers. The signals reach the first level of cortical processing at the primary auditory cortex (A1), in the superior temporal gyrus of the temporal lobe. Most areas up to and including A1 are tonotopically mapped (that is, frequencies are kept in an ordered arrangement). However, A1 participates in coding more complex and abstract aspects of auditory stimuli without coding well the frequency content, including the presence of a distinct sound or its echoes. Like lower regions, this region of the brain has combination-sensitive neurons that have nonlinear responses to stimuli.

Recent studies conducted in bats and other mammals have revealed that the ability to process and interpret modulation in frequencies primarily occurs in the superior and middle temporal gyri of the temporal lobe. Lateralization of brain function exists in the cortex, with the processing of speech in the left cerebral hemisphere and environmental sounds in the right hemisphere of the auditory cortex. Music, with its influence on emotions, is also processed in the right hemisphere of the auditory cortex. While the reason for such localization is not quite understood, lateralization in this instance does not imply exclusivity as both hemispheres do participate in the processing, but one hemisphere tends to play a more significant role than the other.

Recent ideas

Alternation in encoding mechanisms have been noticed as one progresses through the auditory cortex. Encoding shifts from synchronous responses in the cochlear nucleus and later becomes dependent on rate encoding in the inferior colliculus.

Despite advances in gene therapy that allow for the alteration of the expression of genes that affect audition, such as ATOH1, and the use of viral vectors for such end, the micro-mechanical and neuronal complexities that surrounds the inner ear hair cells, artificial regeneration in vitro remains a distant reality.

Recent studies suggest that the auditory cortex may not be as involved in top down processing as was previous thought. In studies conducted on primates for tasks that required the discrimination of acoustic flutter, Lemus found that the auditory cortex played only a sensory role and had nothing to do with the cognition of the task at hand.

Due to the presence of the tonotopic maps in the auditory cortex at an early age, it has been assumed that cortical reorganization had little to do with the establishment of these maps, but these maps are subject to plasticity. The cortex seems to perform a more complex processing than spectral analysis or even spectro-temporal analysis.

ARABPSYCHOLOGY.COM