Harvard-MIT Division of Health Sciences and Technology
HST.723: Neural Coding and Perception of Sound
Instructor: Bertrand Delgutte, and Andrew Oxenham

Hearing and the Auditory System: Overview

Bertrand Delgutte

Andrew Oxenham

The Auditory System

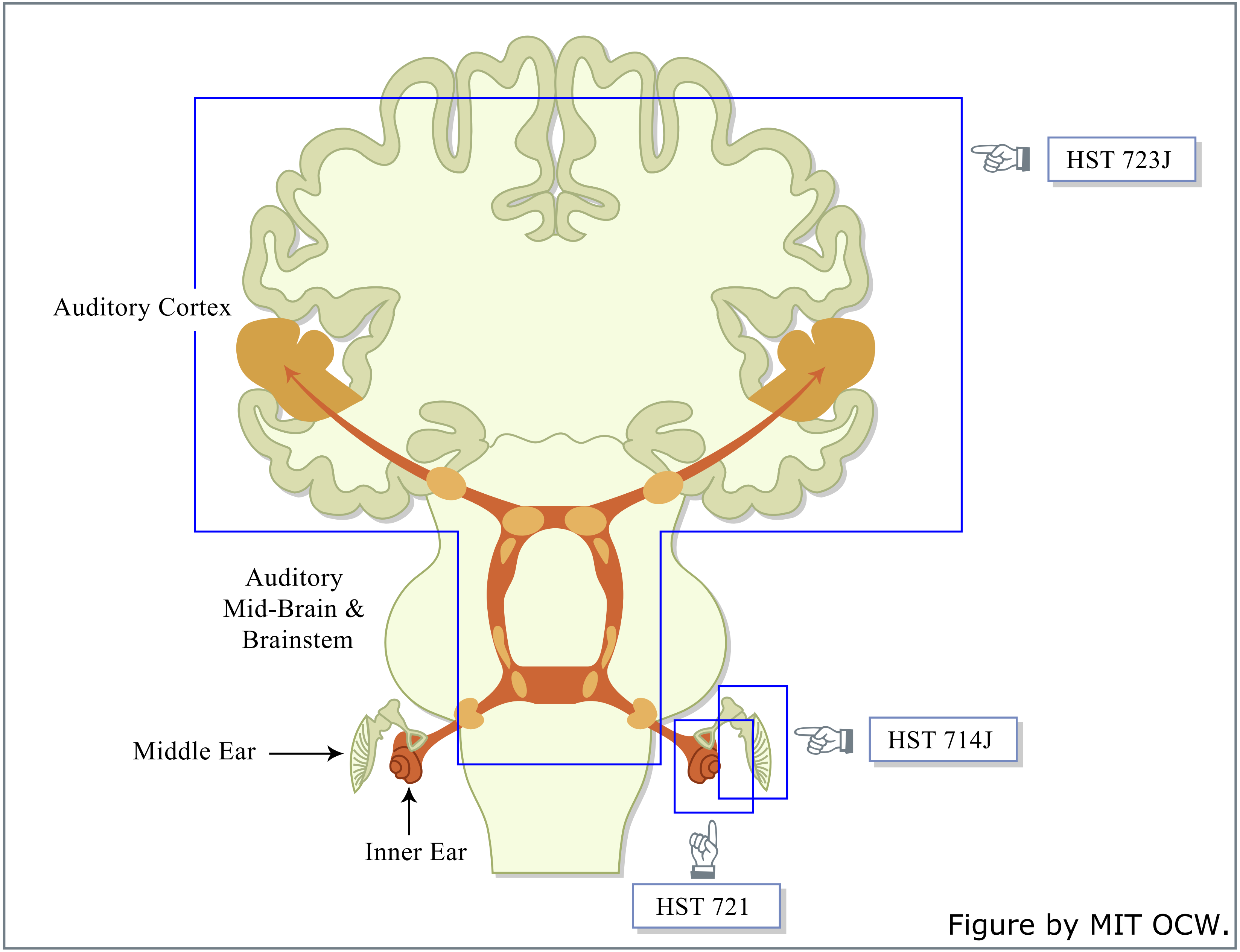


Figure by MIT OCW.

The Human Ear

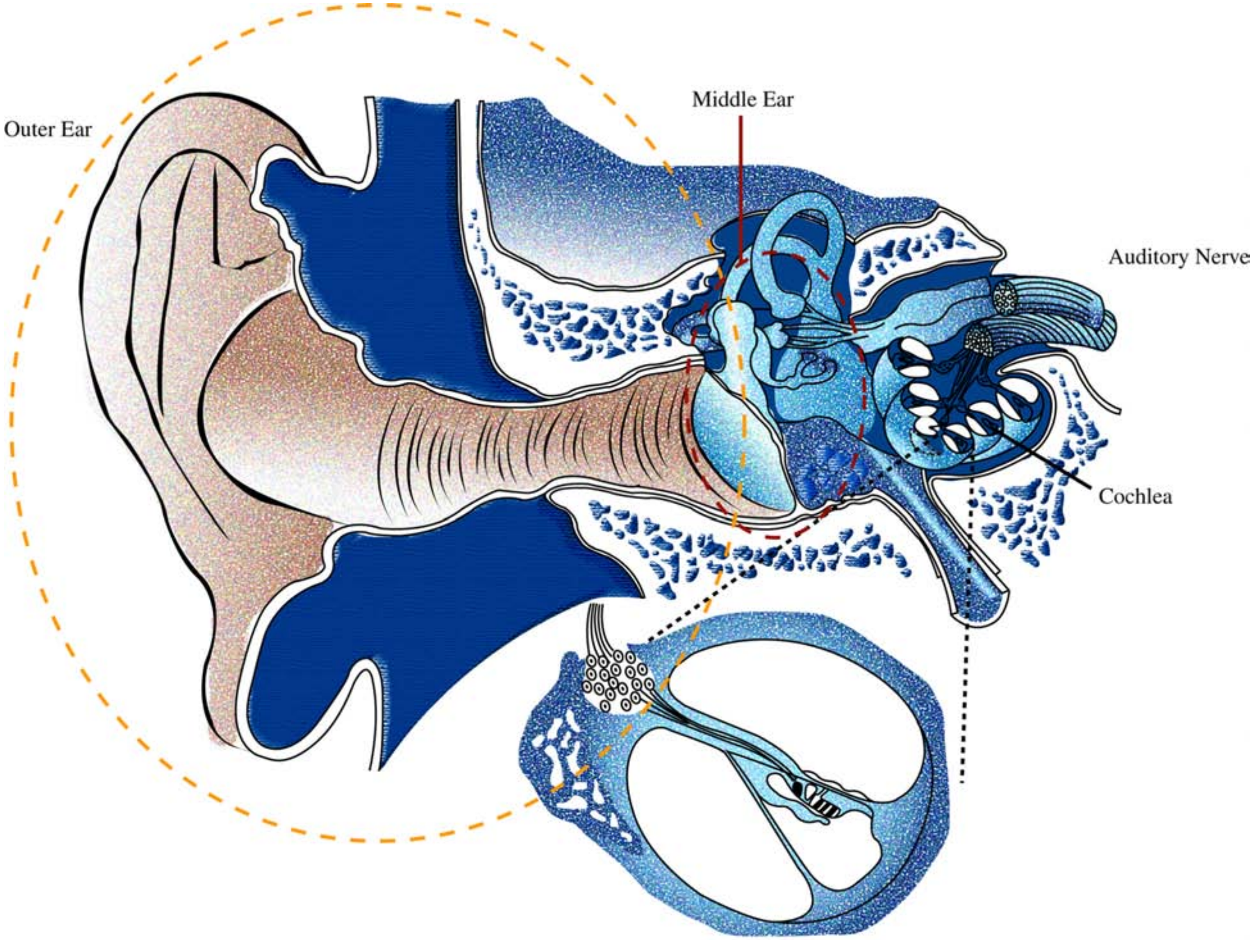


Figure by MIT OCW.

Cochlear mechanics: Frequency analysis

Figure removed due to
copyright reasons.

Apex: Low Frequency" and "Base: High Frequency

1. Most basic form: Cochlear is compartment filled with water-like fluid, with two openings (oval and round windows).
2. Stapes into oval window provides interface with air-borne sound. Receptor cells and innervation locate on cochlear partition. When pressure applied to the stapes, displacement of partition follows applied pressure. No mechanical frequency selectivity in gross cochlear mechanics. Because cochlear fluids are incompressible at sound frequencies, any inward volume displacement of the stapes results in an equal outward displacement of the round window. Similar to situation in frogs and some lizards.
3. Key step for mechanical frequency analysis. Cochlear is elongated, and mechanical properties of partition (stiffness, width) vary along the lengths of cochlea. Also helicotrema.
    - For very low frequencies, stapes displacement leads to fluid flow through the *helicotrema*, where scala vestibuli connects with scala tympani. However, at most sound frequencies, the mechanical impedance of the fluid mass in the cochlea is higher than that of the cochlear partition due to the stiffness of the basilar membrane, so that the cochlear partition is deflected.
    - Because the stiffness of the basilar membrane drops by two orders of magnitude from the base to the apex, the effective resonance frequency of the membrane decreases systematically from base to apex, and therefore the point of maximal displacement of the cochlear partition increases from base to apex as frequency is lowered.
    - Similar to situation in birds.
4. In mammals, cochlea coiled. Probably not essential to function. Makes CI more difficult. Cochlea is Latin for snail.

Hair Cells: Mechano-electric transduction

Figures removed due to copyright reasons.

- Hair cells transduce the mechanical energy in sound into electrical energy.
- Loss of hair cells due to environmental or genetic conditions is the leading cause of sensorineural hearing impairment.
- Hair cells have two poles that serve different functions:
  - The apical ends are covered with cilia (hairs) which contain transducer channels in their membranes. Mechanical displacement of the cilia in one direction leads to opening of these channels and current flow into the cell, resulting into depolarization.
  - The basal ends form synapses onto auditory nerve fibers. Depolarization of the hair cell resulting from cilia displacement leads to release of neurotransmitter and initiation of all-or-none action potentials in the postsynaptic nerve fiber.
- Hair cell stererocilia are connected together by thin *tip links* to which the transduction channels are attached. At rest, a small fraction of the channels are open. Displacement of the cilia towards the largest cilia stretches the tip links and opens more transduction channels, while displacement in the opposite direction closes the channels.
- Opening the transduction channels causes K+ ions in scala media to flow into the cell, resulting in a depolarization (increase towards 0 mV) of the normally negative (-60 mV) intracellular potential.

Auditory Nerve: Frequency selectivity

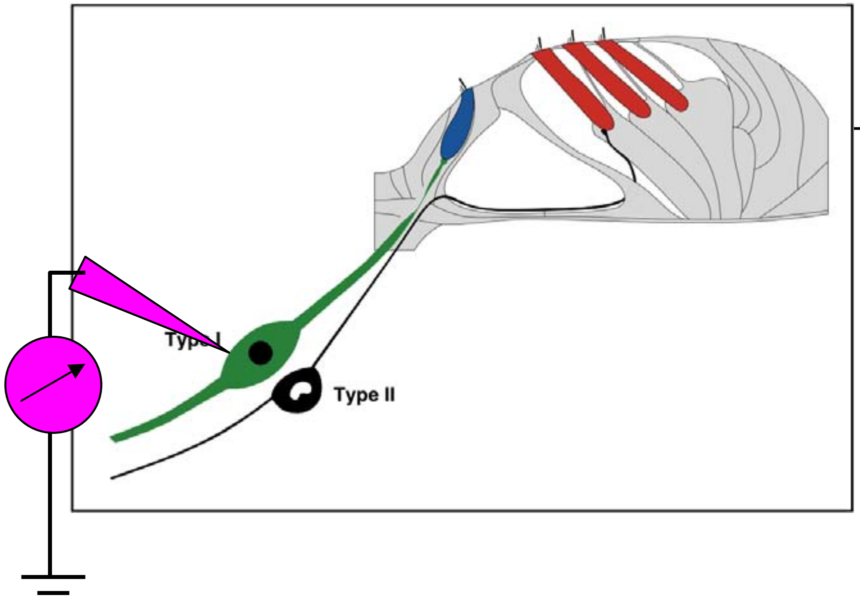
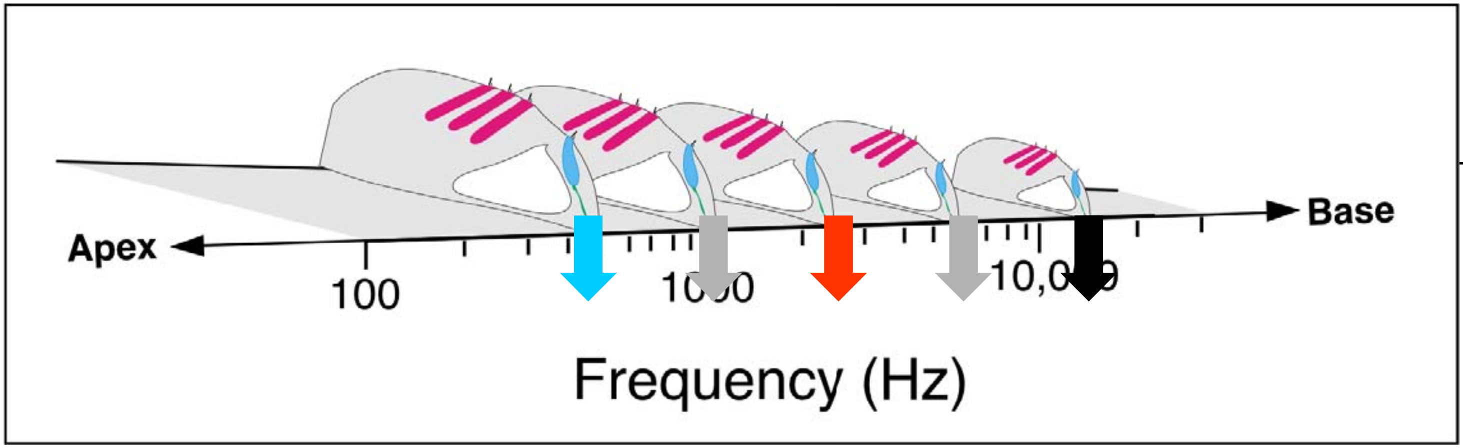
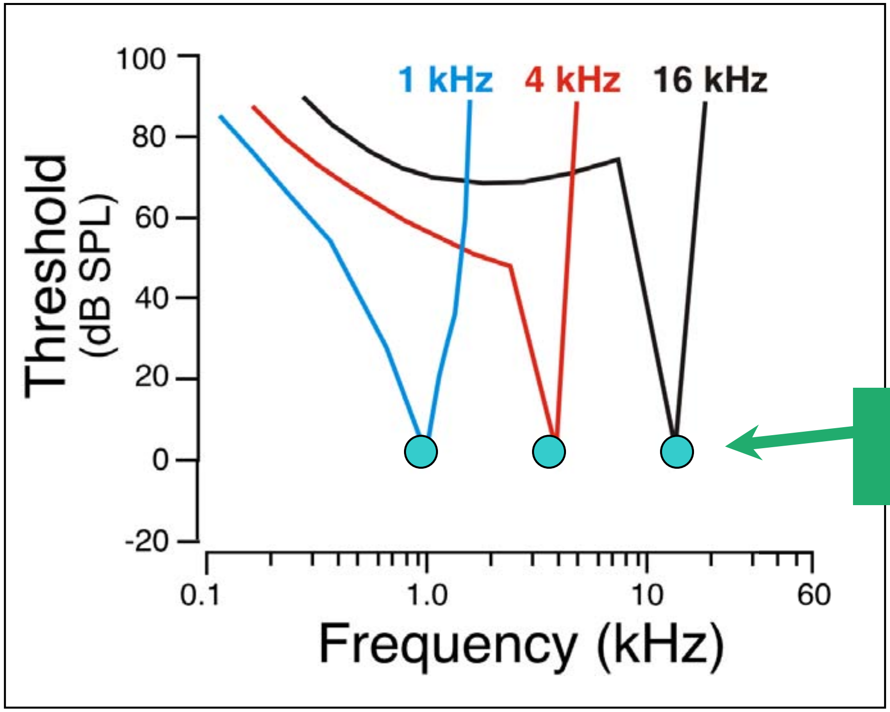


Figure removed due
to copyright reasons.

- Each vertical bar represents one spike (action potential) recorded from an AN fiber in response to a pure tone swept in frequency at different intensities.
- Spike discharges occur even in the absence of a sound: There is *spontaneous activity*.
- For low intensities, discharge rate increases above spontaneous only for a narrow range of frequencies. As intensity increases, so does the range of frequencies to which the fiber responds.
- The outline of the response area if the pure tone *tuning curve* or *frequency threshold curve*. The frequency for which threshold is minimum is the *characteristic frequency* (CF).

Auditory Nerve: Frequency map (tonotopy)

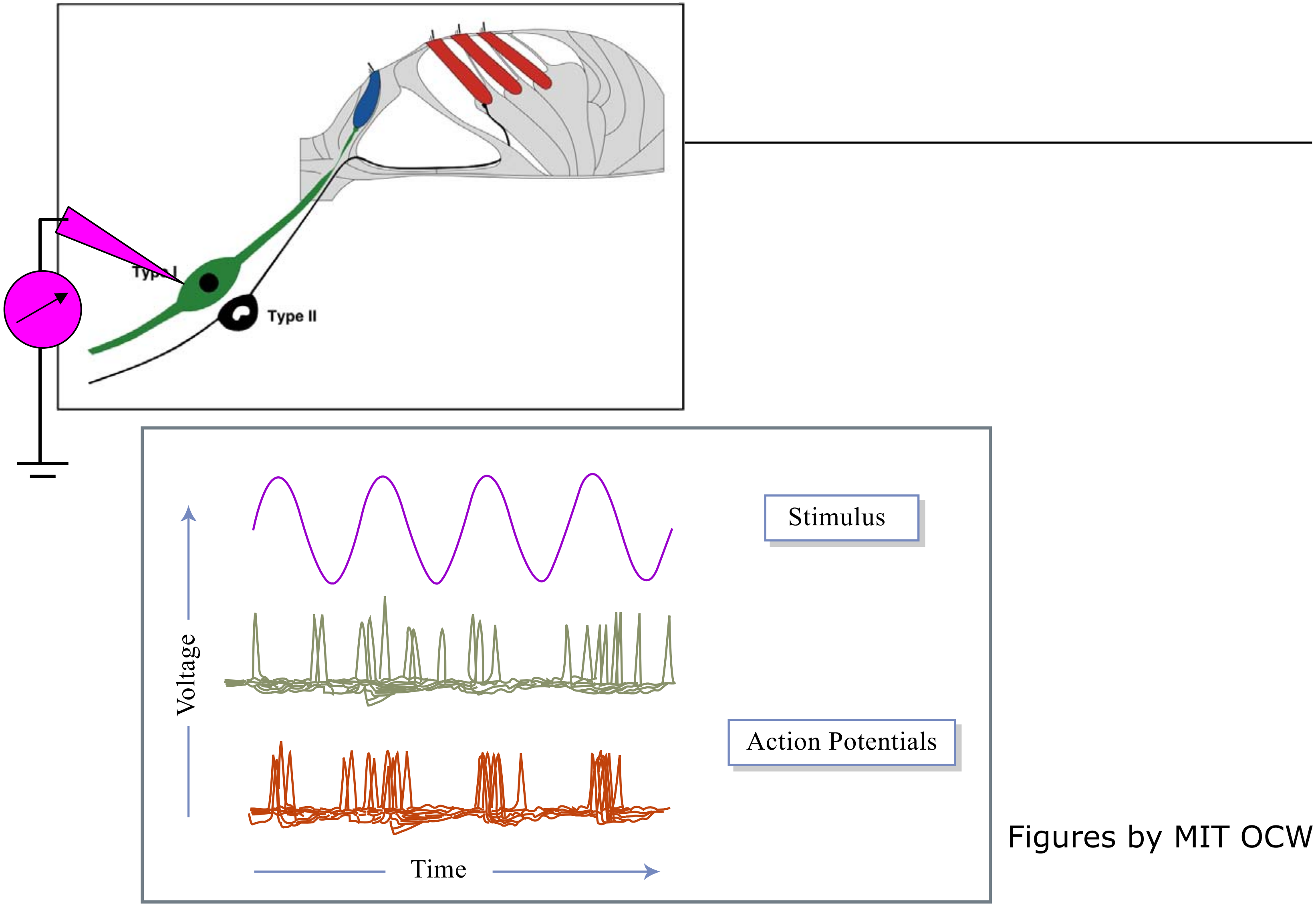




Characteristic Frequency (CF)

- The cochlear frequency map is the basis for a *place code* for sound frequency.
- **Theme 1** explores consequences of frequency selectivity and cochlear nonlinearities for auditory masking.

Auditory Nerve: Phase locking



Figures by MIT OCW.

- In response to low-frequency (< 5 kHz) pure tones, spike discharges tend to occur at a particular phase within the stimulus cycle. However, spikes do not always occur on every cycle, i.e. there can be 2, 3, or more cycles between consecutive spikes.
- Phase locking best seen with frequencies below 5 kHz.
- Phase locking is the basis for a temporal code for sound frequency.

The ascending auditory pathway

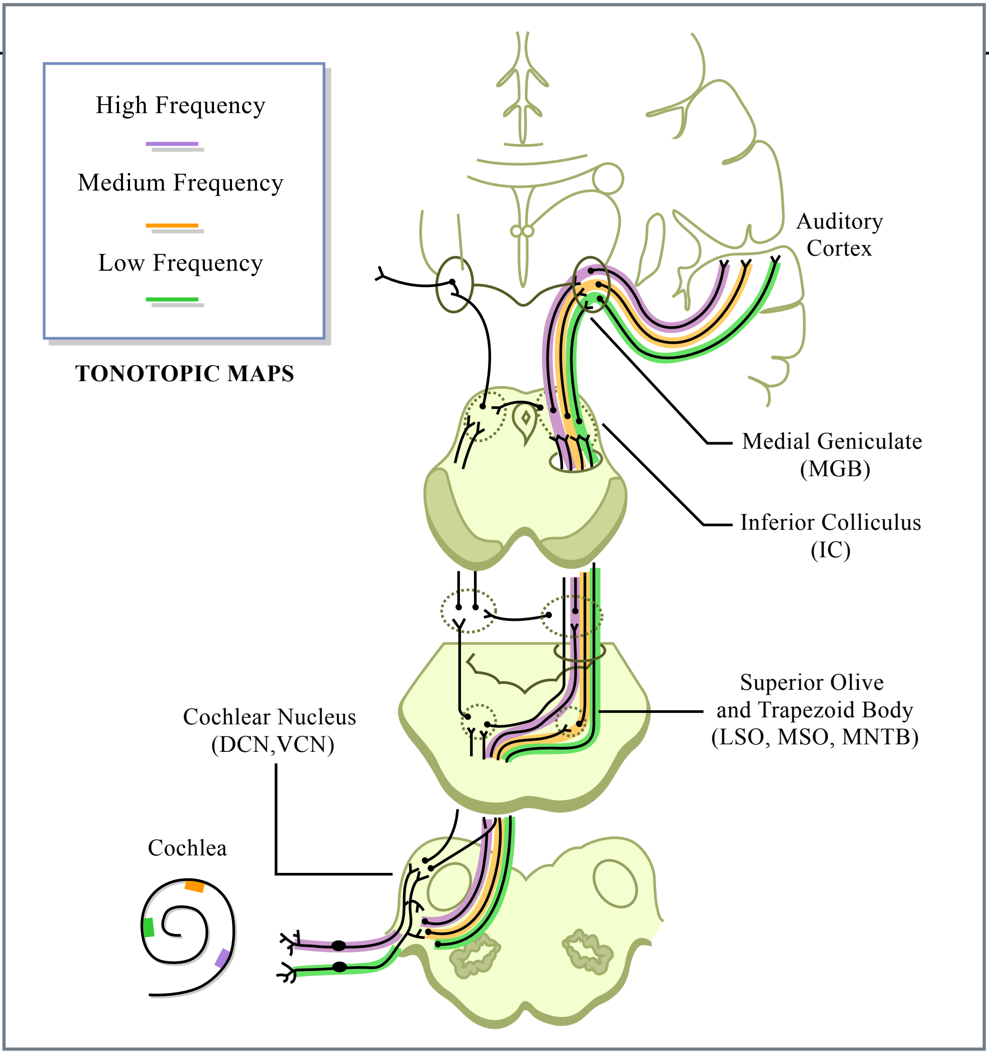


Figure by MIT OCW.

Descending pathway roughly parallels ascending pathway

Tonotopic organization maintained from cochlea to auditory cortex

Stages of processing – differs from visual system where almost direct input from retina to cortex

- Cochlear nucleus
- Superior olivary complex
- Inferior colliculus (midbrain)
- Medial geniculate (thalamus)
- Auditory cortex

Most auditory nuclei are located near dorsal surface of brainstem

Auditory Structures

- 8N: 8th Nerve
- CN: Cochlear Nucleus
- LL: Lateral Lemniscus
- IC: Inferior Colliculus
- SC: Superior Colliculus
- ICO: Commissure of IC
- BIC: Brachium of IC
- MGB: Medial Geniculate Body
- AI: Primary Auditory Cortex

Other Structures

- ICM: Cerebellum
- 5N; Trigeminal Nerve

Figure removed due to copyright reasons.

Aitkin (1986)

Cochlear Nucleus: Parallel processing paths

Tonotopy

Figure removed due to copyright reasons. Please see: Ryugo, D. K., and T. N. Parks. "Primary innervation of the avian and mammalian cochlear nucleus." *Brain Res Bull* 60, no. 5-6 (Jun 15, 2003): 435-56.

Cell types

Figure removed due to copyright reasons.

- CN has 3 major subdivisions: AVCN, PVCN and DCN
- Each subdivision contains distinct *cell types* differing in morphology, cytochemistry, patterns of inputs and outputs, and responses to sound

- CN has 3 major subdivisions: AVCN, PVCN and DCN. ANR is nerve root, where AN fibers enter CN
- Auditory nerve fibers bifurcate, forming an ascending branch innervating AVCN and descending branch innervating PVCN and then DCN.
- Each of the 3 regions contains tonotopic map, resulting form orderly arrangement and bifurcation of ANFs
- Each of the three subdivisions contains distinct cell types. Cell types differ by morphology, pattern of inputs, cytochemistry (membrane channels), and output projections. Also differ in response to sounds.
    - Pyramidal in DCN. Projects to IC. Pauser PSTH.
    - Octopus in PVCN. Projects to VNLL. Onset PSTH.
    - SBC in anterior AVCN. Projects to SOC. Primary-like response pattern.
    - GBC in posterior AVCN. Projects to SOC (different cell types than SBC). Pri-notch response pattern.
    - Multipolar in AVCN/PVCN. Subtypes exist with different projections. Chopper response pattern.
    - Because central projections of each of these cell types are distinct, **CN estaplishes parallel processing pathways**
- Association between anatomical cell types and physiological response types is now well understood for CN, and CN is the only place in auditory system for which this is true. Serves are model for studying cellular mechanisms throughout auditory system.
- CN also model for understanding trnasformation of stereotypical input into diverse outputs. **Theme 2**

Superior olivary complex: Binaural processing of time and level differences

ITD circuit:
Medial superior olive receives excitatory inputs from both sides. MSO cells are coincidence detectors

ILD circuit:
Lateral superior olive receives excitatory and inhibitory inputs from opposite sides

Figures removed due to copyright reasons.

Two separate circuits involving the SOC.
“ITD circuit” centered on MSO. MSO neurons receive excitatory inputs from both sides via SBCs in AVCN. MSO neurons act as coincidence detectors, only responding when inputs from both sides arrive simultabeously. Gives MSO neurons tuning to ITD.
“ILD circuit” centered on LSO. Receive excitatory inputs from one side via SBC and inhibitory inputs from other side via GBCs and inhibitory neurons in MNTB. Interplay of excitation and inhibition gives LSO neurons sensitivity to ILD

These are best examples of circuits whose function can be understood. Each circuit contains specializations adapted to its function. Good model for understanding how circuits can achieve certain signal processing function. **Theme 3.**

Inferior Colliculus: Convergence and reorganization

Figures removed due to
copyright reasons.
Please see: **Irvine (1992)**

- Parallel processing pathways originating in the cochlear nucleus converge onto the central nucleus of the IC.
- Some of the projections from CN to IC are direct (monosynaptic); others are via intervening synapses in the SOC and/or the nuclei of the lateral lemniscus.
- IC performs considerable processing: Its neuron responses are more complex than those of its inputs.

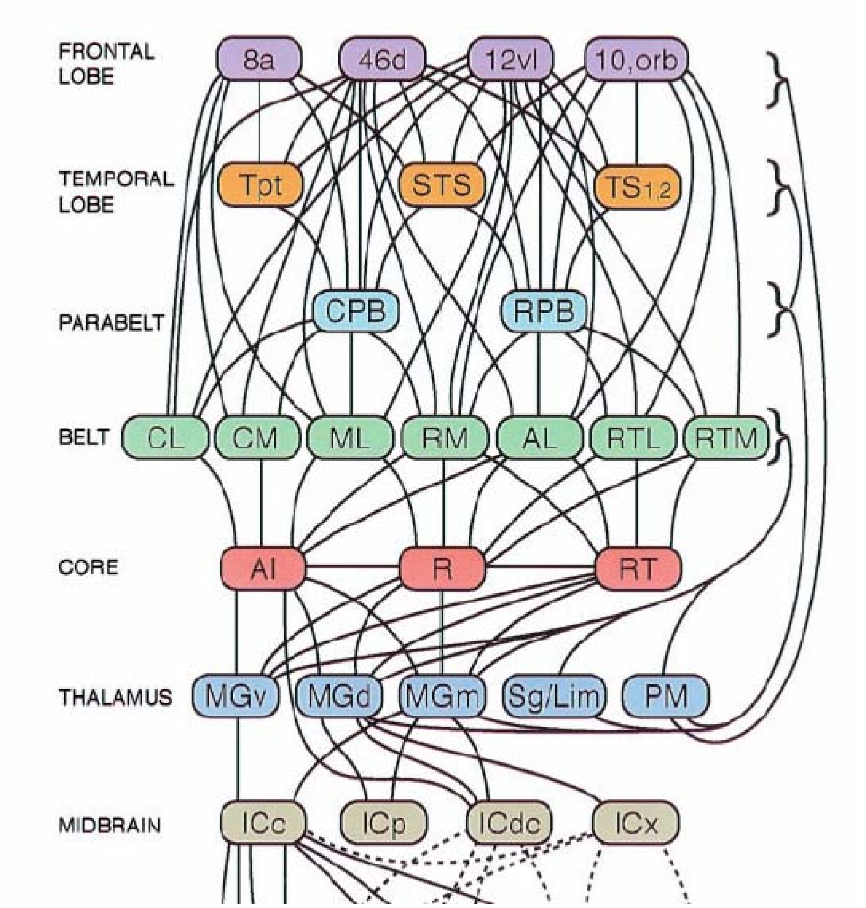
- Processing in IC includes temporal, spectral, and binaural aspects
- IC by itself not topic of any theme since its function is poorly understood.
- Some IC papers may be discussed in Themes 3 and 4 to address questions that are best addressed at this stage of processing (e.g.requiring large amount of data).

Auditory cortex:
Subdivisions and processing streams

Figure removed due to
copyright reasons.

Macaque

- In old world primates including humans, auditory cortex is located deep within the sylvian fissure. Part of overlying cortex of a macaque has been excised to visualize auditory cortex
- Not one auditory cortex, but hierarchically organized auditory cortices. A dozen different auditory areas identified in cortex of macaques and cats.
- *Core* is most closely connected to lower auditory nuclei and give more general and stable responses to sound. Contains primary auditory cortex A1.
- Responses in belt are more complex, more specific, and more dependent on state.
- Extensive pattern of hierarchical and reciprocal connections between core and belt, belt and parabelt, and belt/parabelt with association areas of temporal and frontal lobes.
- This organization is thought to provide anatomical substrate for processing streams. E.g streams specialized for sound localization or pitch processing or phonetic processing



(Figure from: Kaas, J. H., and T. A. Hackett. “Subdivisions of auditory cortex and processing streams in primates.” *Proc Natl Acad Sci U S A* 97, no. 22 (Oct 24, 2000): 11793-9. “Copyright (2000) National Academy of Sciences, U.S.A.”)

fMRI activation of the human auditory system

Figure removed due to
copyright reasons.

•Bilateral stimulation with white noise; measure cerebral blood flow with fMRI
- Not only cortex, but also brainstem auditory structures can be imaged using cardiac gating
- fMRI is beginning to give evidence for processing streams in humans
- Neuroimaging papers discussed in Themes 4-6

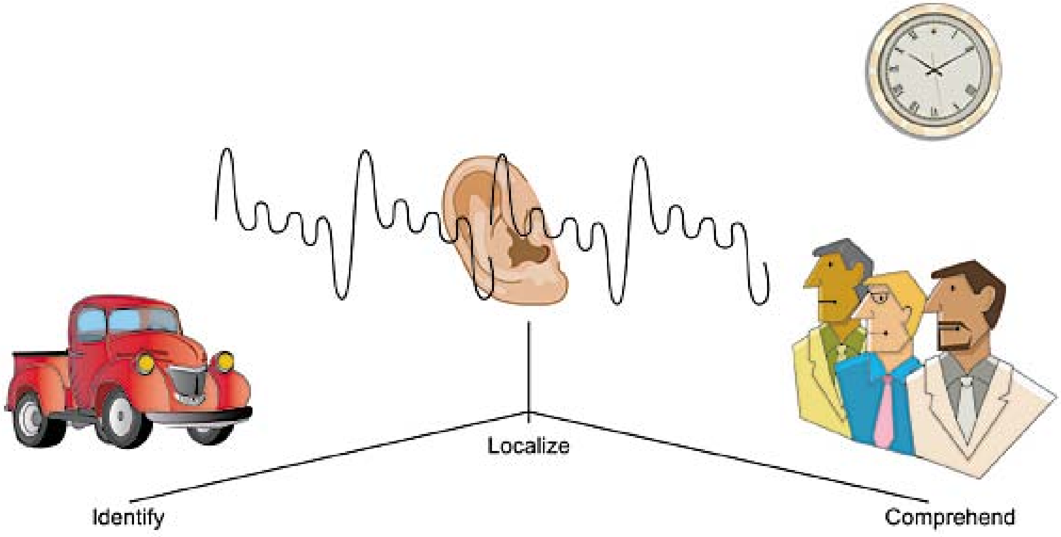
Auditory Cortex: Plasticity

Figure removed due to copyright reasons.
Please see: Kilgard, M. P., and M. M. Merzenich. "Cortical map reorganization enabled by nucleus basalis activity." *Science* 279, no. 5357 (Mar 13, 1998): 1714-8.

NBM: Nucleus basalis

- NB neurons are activated as a function of the behavioral significance of stimuli (*6*).
- Several forms of learning and memory are impaired by NB lesions (*7*).
- NB to auditory cortex is example of descending projection where higher level centers involved in control of behavior can influence activity of auditory system.
- Electric stimulation of NB in rat was paired with acoustic stimulation with a 9 kHz tone for 20-25 days.
- Unstimulated ("naïve rat) shows gradual tonopotic map where all frequencies from 1 to >40 khz are represented.
- After NB stimulation, 9kHz region is greatly expanded.
- Changes in cortical responses related to behavioral state can occur on much shorter time scales (minutes)

Functions of Hearing



Figures by MIT OCW.

- But how do we know what sounds belong to what source, if more than one sound is present?
- Imagine identifying all shipping traffic from the motion of two buoys…

How does the auditory system allow us to navigate complex acoustic environments?

How can we study hearing?

- Physics (Acoustics)
- Anatomy and Physiology
- Psychology

- Psychophysics: Establishing a quantitative relationship between physical and psychological variables (Fechner, 1801-1887).

Physical	Psychological
Sound	*Auditory sensation*
Intensity	Loudness
Frequency	Pitch
Spectral distribution	Timbre
Temporal characteristics	

Subjective and objective variables

- Subjective measures
  - Loudness estimation, pitch matching, judgments of “sound quality”, etc.
  - Require introspection
  - Subject to non-sensory biases
- Objective measures
  - Stimulus detection, discrimination
  - A measure of performance: There must be a correct answer.
  - Effects of bias can be separated from sensitivity using *Signal Detection Theory.*

Frequency Selectivity in Hearing

Any periodic waveform can be described in terms of a sum of simple sinusoids (Fourier).

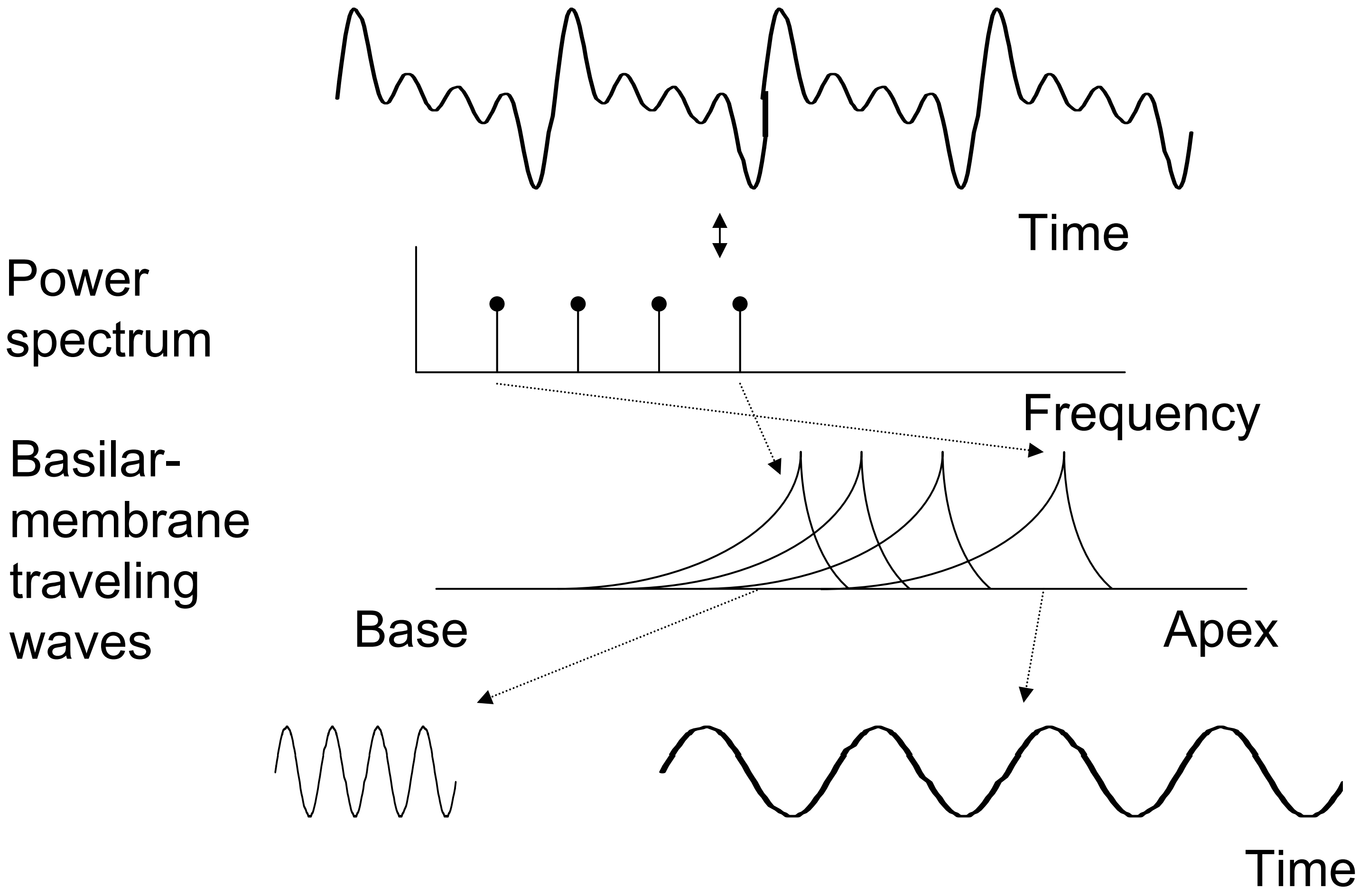
The ear operates in a similar manner: Sound is broken down into constituent frequency components.

Helmholtz (1863) confirmed that it is possible to "hear out" individual sinusoidal components within a complex tone. [Compare with vision.]

Demonstration I: Canceled harmonics

This is an example of synthetic and analytic listening. At first the complex is heard as a whole with a single pitch. Then, individual harmonics are emphasized by gating them on and off.

Frequency-to-place mapping in the cochlea enables the separation of different frequency components. [Cochlear traveling wave – von Bekesy (1960).]



Time

Power spectrum

Frequency

Basilar-membrane traveling waves

Base

Apex

Time

Frequency differences play a major role in the perceptual organization of concurrent and consecutive sounds.

- Demonstration II: Hear out a melody in the presence of distractors

M E L O D Y

D I S**M**T**E**R**L**A**O**C**D**T**Y**O R S

- Demonstration III: East African xylophone music. (Note that the pitches of both xylophones remain the same after filtering.)

Masking

- Masking is an everyday phenomenon which occurs when frequency selectivity fails, and we are no longer able to resolve, or hear out, one stimulus (the signal) in the presence of another (the masker).
- This phenomenon can be used to probe the limits of frequency and intensity resolution, using objective measures.

Issues of Masking and Nonlinearity

- How can we best use masking to measure frequency selectivity?
- What is the relationship between frequency selectivity measured physiologically and our own abilities, as measured in psychoacoustic tests?
- What are neural mechanisms underlying masking?

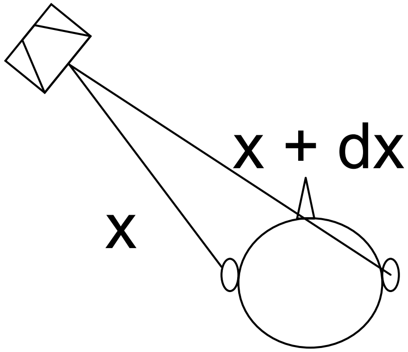
Binaural and Spatial Hearing

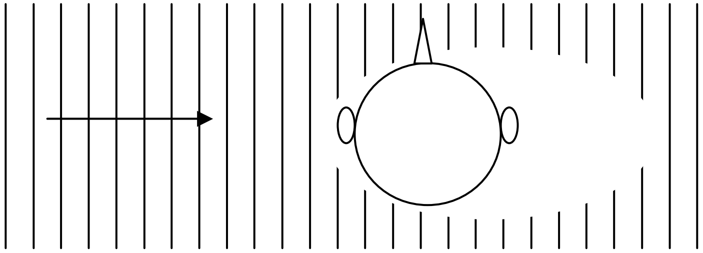
Two ears give us the potential to compare two different acoustic waveforms.

This comparison allows us to localize sounds in space, often with remarkable accuracy.

Differences between the ears:

- Interaural time differences (ITDs)
- Interaural level differences (ILDs)





Binaural and Spatial Hearing

- How are level and timing differences used to create the perception of source location?
- How do binaural cues help us to cope in noisy and reverberant environments?
- How do we “know” where a sound source is: are our percepts subject to change (plasticity)? Is there a critical age?
- Why does the ventriloquism effect work?

Pitch and Temporal Coding

- Most periodic sounds (and some aperiodic ones) elicit a sensation of pitch.
- Pitch forms the basis of most music and is also an important aspect of speech.
- The relationship between pitch and frequency is relatively simple for pure tones (sinusoids). What about complex stimuli with many frequencies?

Demonstration IV: Spectral and “virtual” pitch

Pitch and Temporal Coding

- How are temporal patterns coded in the auditory system?
- How is pitch extracted?
- How can we explain the phenomenon of the "missing fundamental"?
- What role does pitch play in the perceptual organization of sounds?

Auditory Object and Stream Formation

Sound elements combine over time and frequency to form “objects” (e.g., notes) and “streams” (e.g., melodies).

- How do we solve the problem of assigning sound elements to the correct object?
- To what extent is this a “top-down” or “bottom-up” process?
- Do pitch, spatial location, etc., determine grouping, or vice versa?

Auditory object and stream formation

Xylophone demonstration showed segregation by frequency, but many other parameters also influence percepts.

- Demonstration V (track 1):

Figures removed due to copyright reasons.

References

- Helmholtz, H. L. F. (**1885/1954**). On the Sensations of Tone (Dover, New York). Transl. A. J. Ellis.
- von Békésy, G. (**1960**). Experiments in Hearing (McGraw-Hill, New York).

CDs

- ASA/IPO. Auditory Demonstrations. (Available via the ASA website.)
- Bregman, A. Auditory Scene Analysis. (Available via Bregman's website at McGill University.

Recommended background reading:

- Moore, B. C. J. (**2003**). An Introduction to the Psychology of Hearing, 5th Edition (Academic Press, London).
- Blauert, J. (**1997**). Spatial Hearing: The Psychophysics of Human Sound Localization (MIT, Cambridge).

Plasticity and learning

- How is sound coded in higher levels of the auditory system – are there maps for higher-order percepts, such as pitch, loudness, and spatial location?
- Does learning or exposure to sound change brain organization?
- What techniques are available to study cortical processing in humans and animals?