

# Chapter 11: Sound, The Auditory System, and Pitch Perception

## Für Elise

L. van Beethoven (1770-1827)

*Poco moto*  
*pp*

*pp* *mf* *p* *dim. and rit.* *pp* *legato*

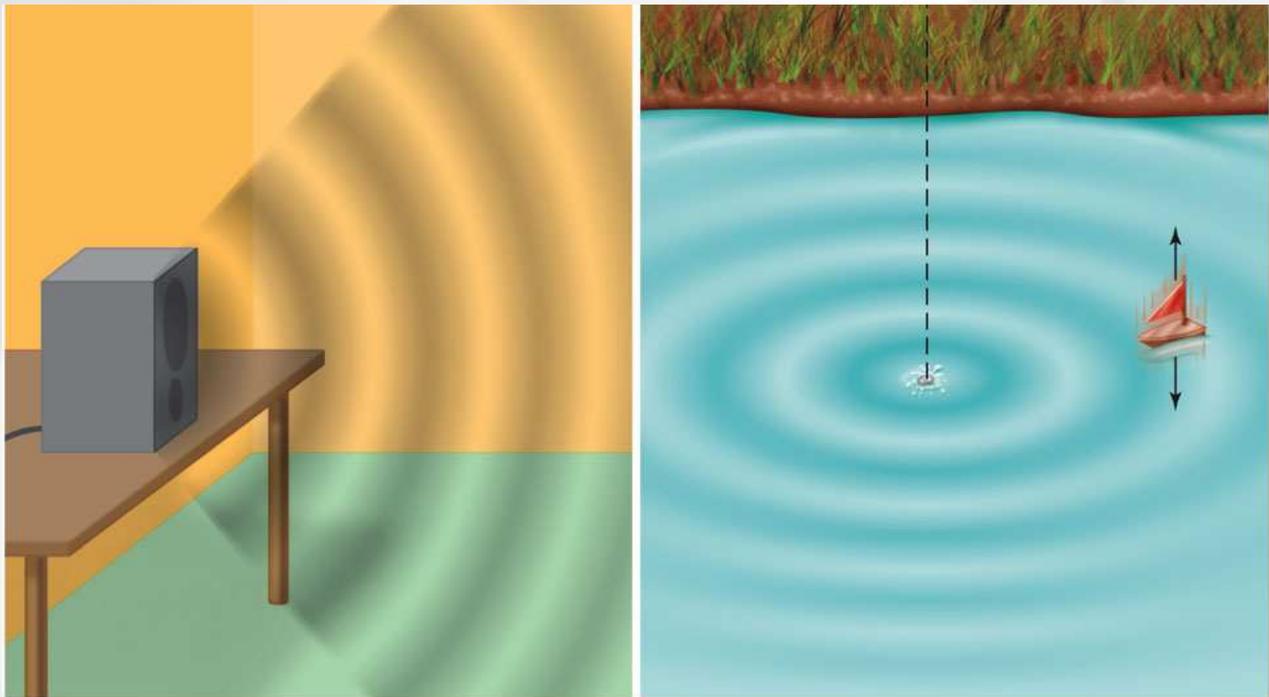


## Overview of Questions

- What is it that makes sounds high pitched or low pitched?
- How do sound vibrations inside the ear lead to the perception of different pitches?
- How are sounds represented in the auditory cortex?

# Pressure Waves and Perceptual Experience

- Two definitions of “sound”
  - Physical definition - sound is *pressure changes* in the air or other medium
  - Perceptual definition - sound is the *experience* we have when we hear

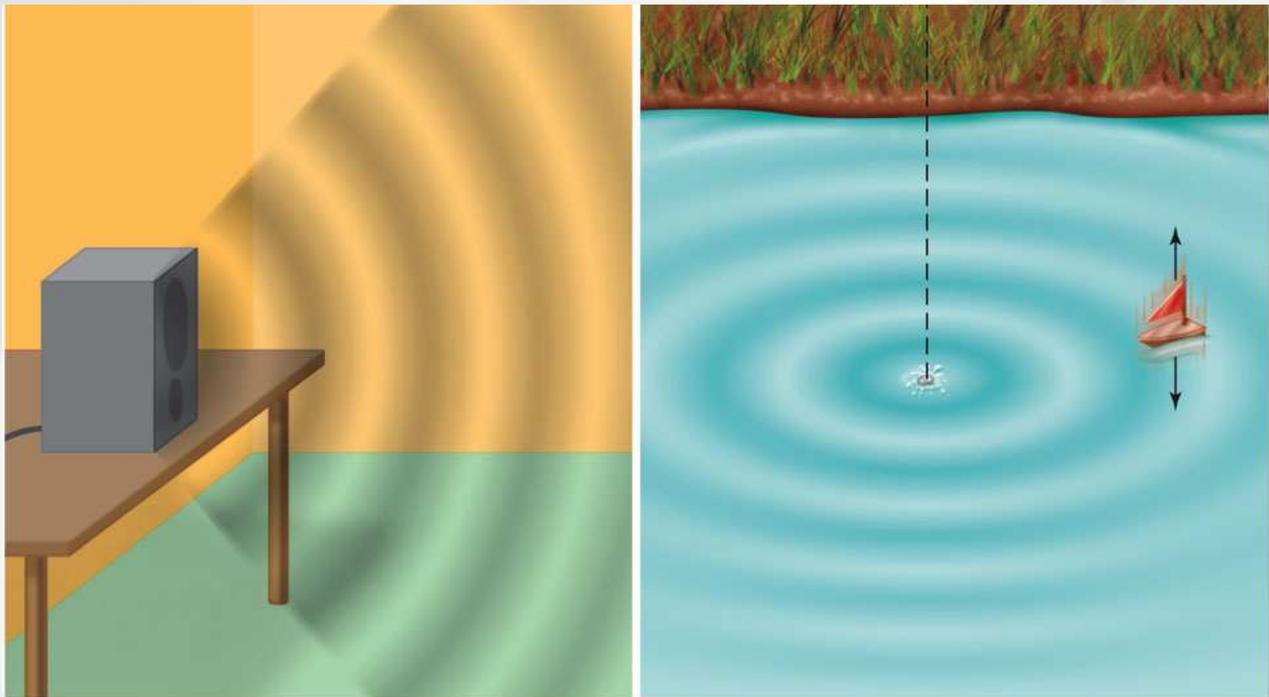


(a)

(b)

# Sound Waves

- Loud speakers produce sound by
  - The diaphragm of the speaker moves out, pushing air molecules together
  - The diaphragm also moves in, pulling the air molecules apart
  - The cycle of this process creates alternating high- and low-pressure regions that travel through the air

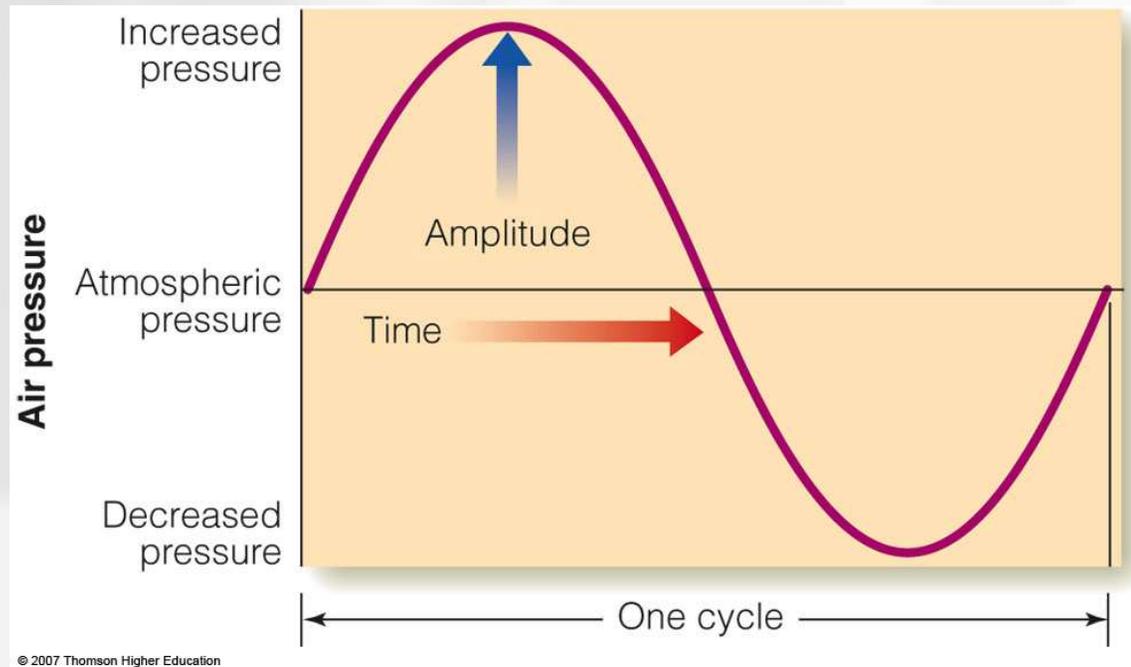


(a)

(b)

# Sound Waves

- Pure tone - created by a sine wave
  - Amplitude ( $p$ ) - difference in pressure between high and low peaks of wave

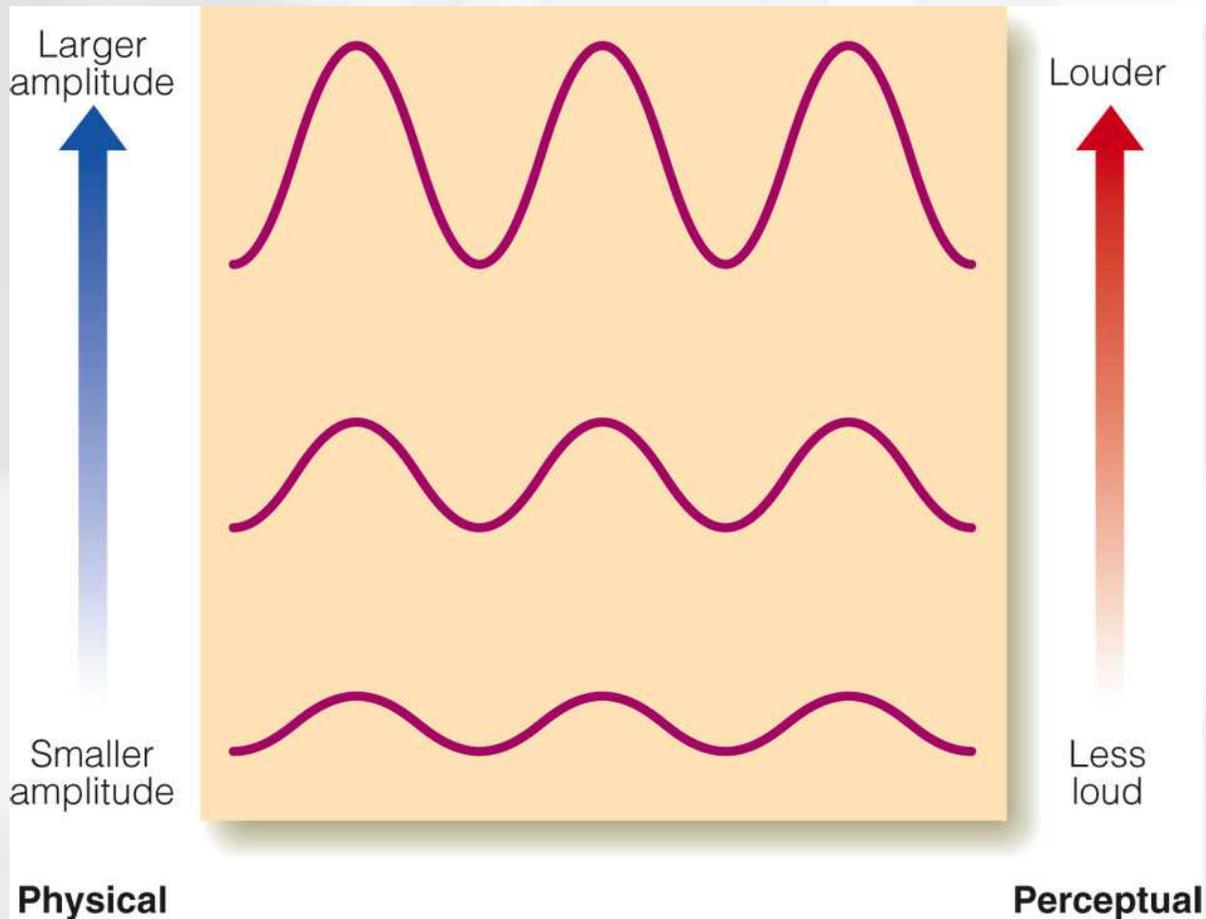


Perception of amplitude is loudness

Decibel (dB) is used as the measure of loudness

Number of dB =  $20 \log_{10}(p/p_0)$

The decibel scale relates the amplitude of the stimulus with the psychological experience of loudness



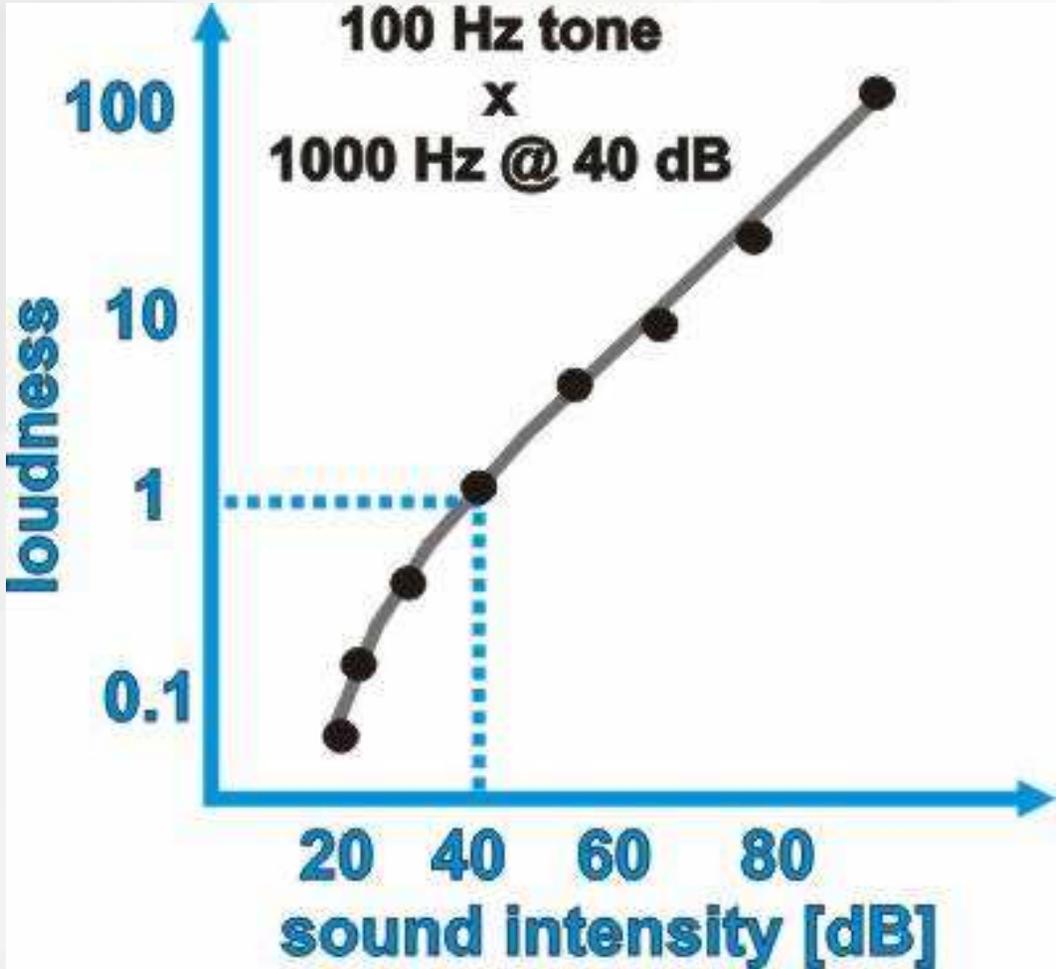
**Table 11.1** ■ *Relative amplitudes and decibels for environmental sounds*

Sound	Relative amplitude	Decibels (dB)
Barely audible (threshold)	1	0
Leaves rustling	10	20
Quiet residential community	100	40
Average speaking voice	1,000	60
Express subway train	100,000	100
Propeller plane at takeoff	1,000,000	120
Jet engine at takeoff (pain threshold)	10,000,000	140
Spacecraft launch at close range	100,000,000	160

© 2007 Thomson Higher Education

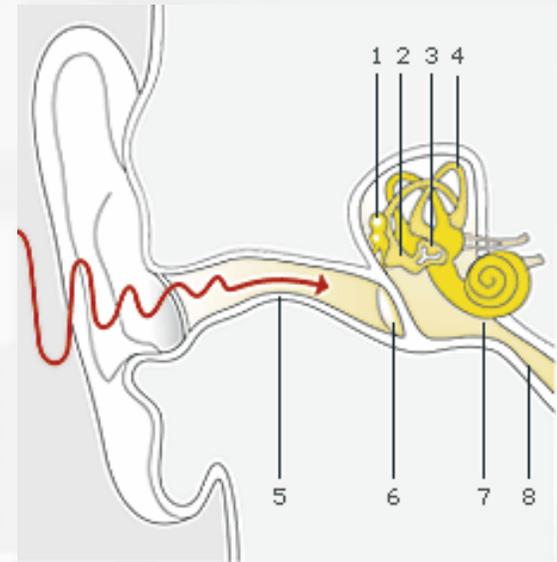
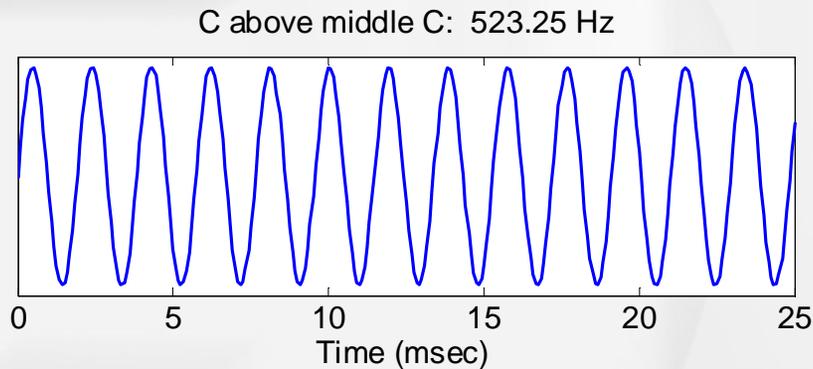
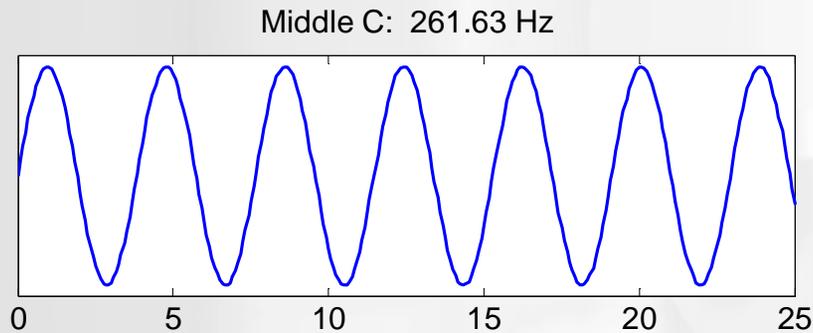
Table 11.1 Relative amplitudes and decibels for environmental sounds

Using magnitude estimation, loudness increases nearly linearly with decibels



# Frequency and Pitch

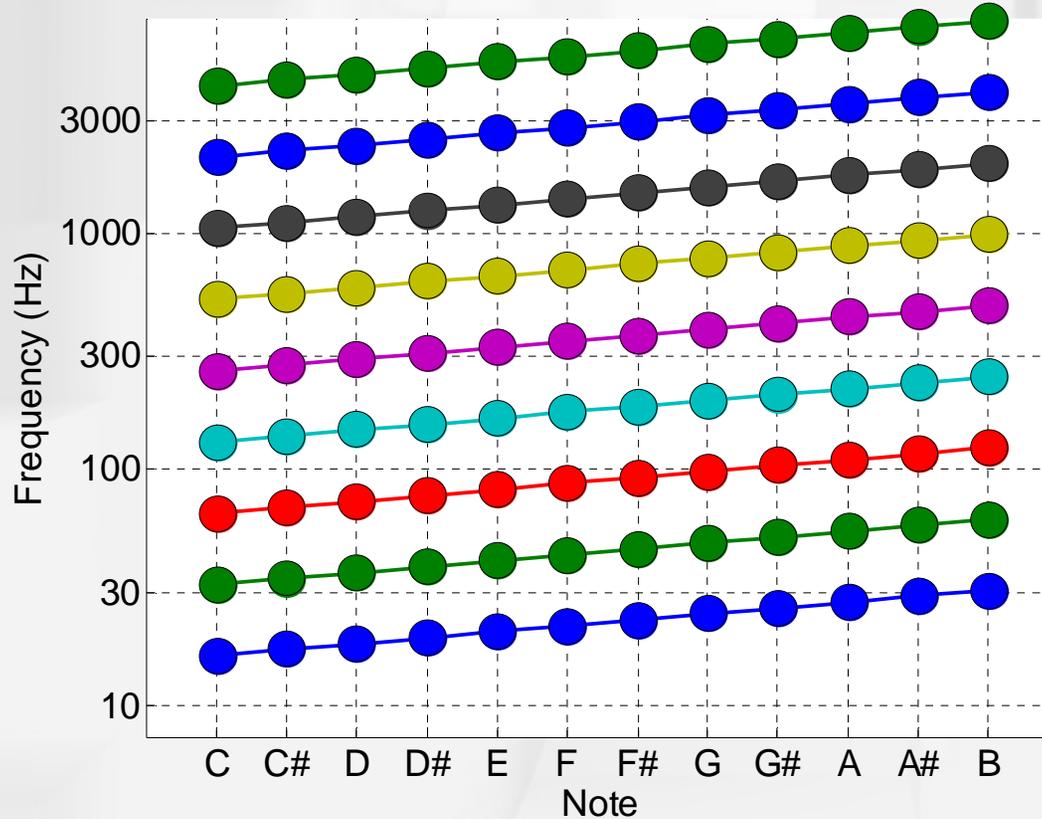
The psychological experience of *pitch* is related to the temporal *frequency* of vibrations of the air hitting the eardrum.



Doubling the frequency increases the pitch by one octave. The tone *chroma* is the same.

# Musical Scales and Frequency

- Letters in the musical scale repeat
- Notes with the same letter name (separated by octaves) sound similar - called tone chroma
- Notes separated by octaves have frequencies that are multiples of each other



## Frequencies of notes on the 12 tone equal tempered scale

Octave	8	4186.0	4434.9	4698.6	4978.0	5274.0	5587.6	5919.9	6271.9	6644.9	7040.0	7458.6	7902.1
	7	2093.0	2217.5	2349.3	2489.0	2637.0	2793.8	2960.0	3136.0	3322.4	3520.0	3729.3	3951.1
	6	1046.5	1108.7	1174.7	1244.5	1318.5	1396.9	1480.0	1568.0	1661.2	1760.0	1864.7	1975.5
	5	523.3	554.4	587.3	622.3	659.3	698.5	740.0	784.0	830.6	880.0	932.3	987.8
	4	261.6	277.2	293.7	311.1	329.6	349.2	370.0	392.0	415.3	440.0	466.2	493.9
	3	130.8	138.6	146.8	155.6	164.8	174.6	185.0	196.0	207.7	220.0	233.1	246.9
	2	65.4	69.3	73.4	77.8	82.4	87.3	92.5	98.0	103.8	110.0	116.5	123.5
	1	32.7	34.6	36.7	38.9	41.2	43.6	46.3	49.0	51.9	55.0	58.3	61.7
	0	16.4	17.3	18.4	19.4	20.6	21.8	23.1	24.5	26.0	27.5	29.1	30.9
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B	

12 Tone Equal Tempered Scale')

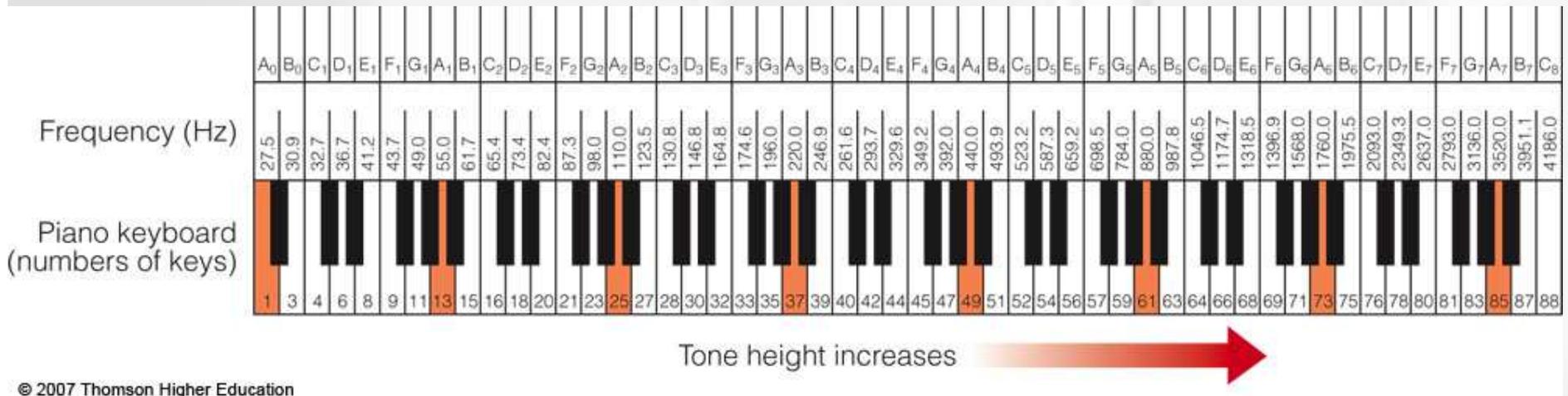
There are 12 equally spaced notes in an octave,

To get to the next note, we multiply the frequency by  $2^{1/12} = 1.0595$

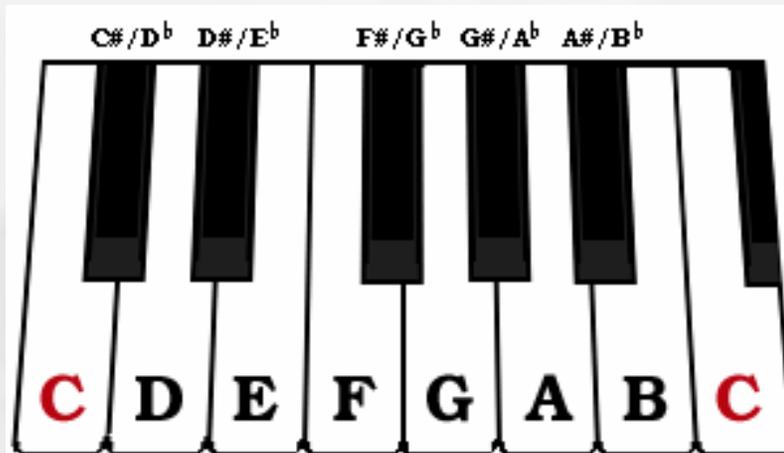
Example: C is 261.6 Hz,  $261.6 * 1.0595 = 277.2$ , or C sharp (D flat)

But a scale contains only 7 of these 12.

Why does it sound right to skip the sharps and flats on in the key of C?  
(I don't know)



Why are there 12 notes on the even tempered scale? Why not 10, or 20?



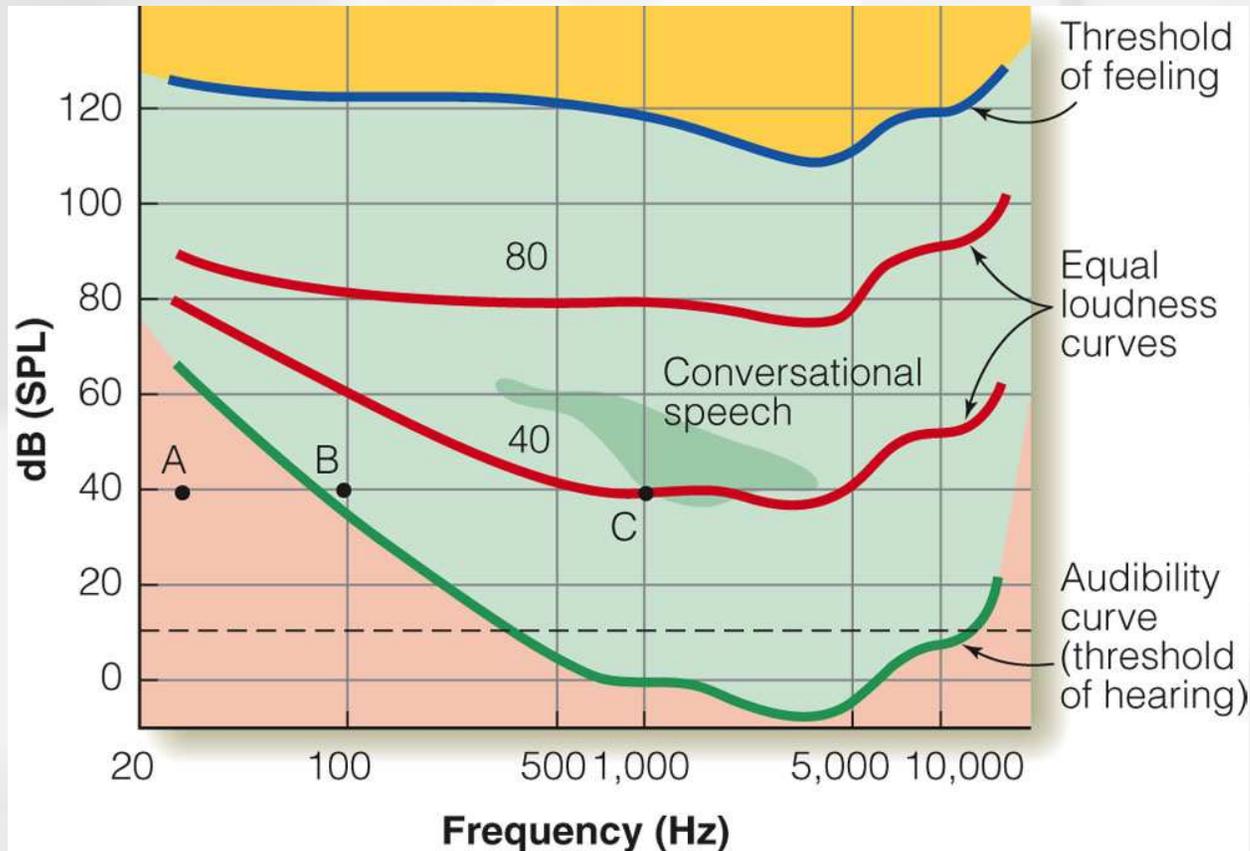
For some reason, notes with frequencies of certain ratios sound right together or are **consonant**

- 1/1 - C
- 2/1 - the octave
- 3/2 - the perfect fifth (C & G)
- 4/3 - the perfect fourth (C & F)
- 5/4 - the major third (C & E)
- 6/5 - the minor third (C & E flat)
- 5/3 - the major sixth (C & A)
- 8/5 - the minor sixth (C & A flat)

It turns out, the 12 note scale contains notes that has pairs that are almost exactly consonant. But not exactly!

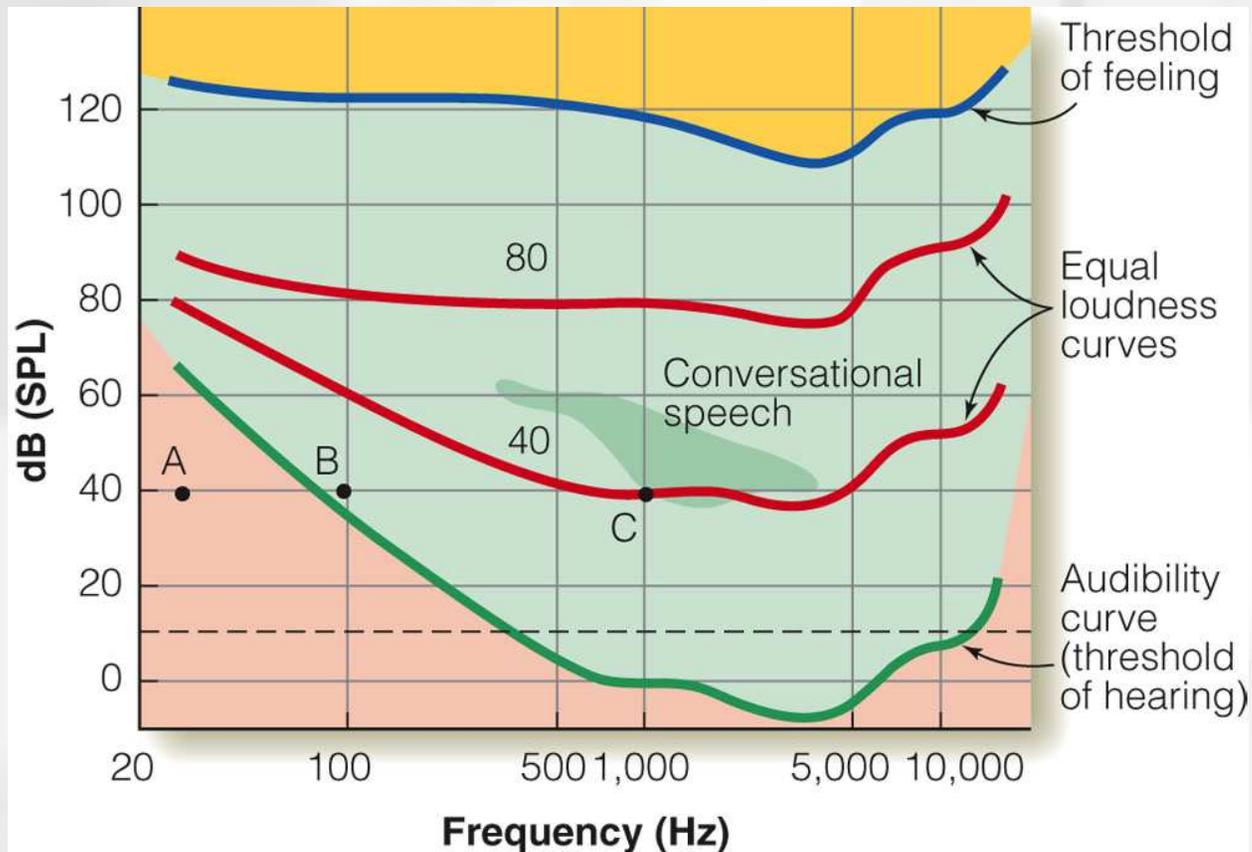
# Range of Hearing

- Human hearing range - 20 to 20,000 Hz (dogs 40KHz, cats 50KHz)
- Audibility curve - shows the threshold of hearing
  - Changes on this curve show that humans are most sensitive to 2,000 to 4,000 Hz
- Auditory response area - falls between the audibility curve and the threshold for feeling

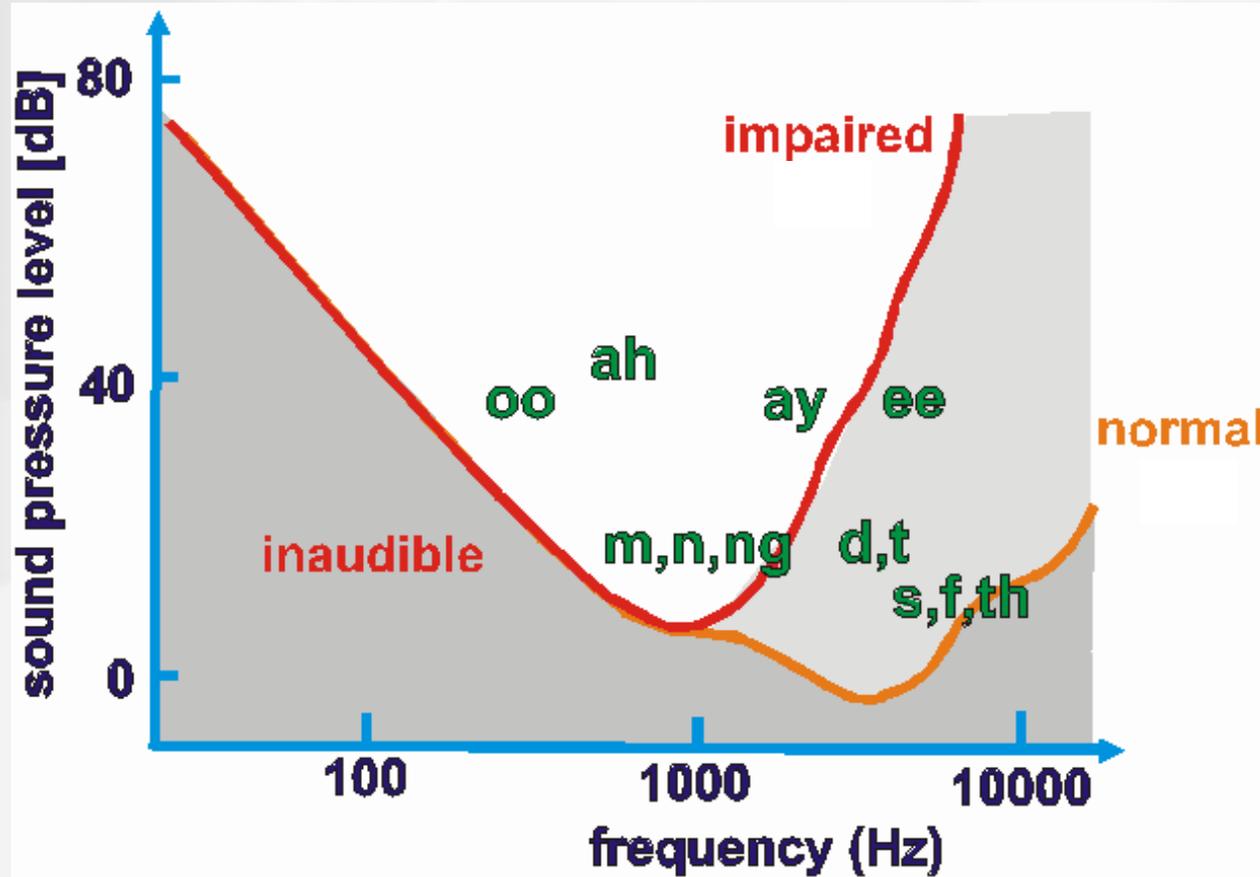


# Range of Hearing

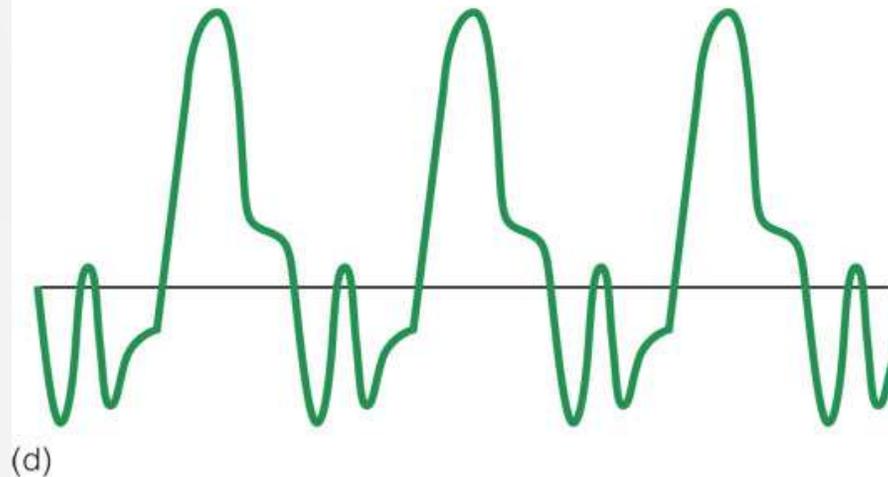
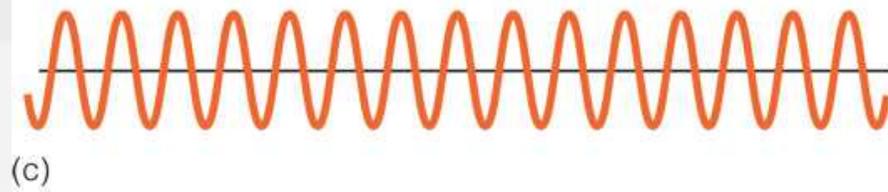
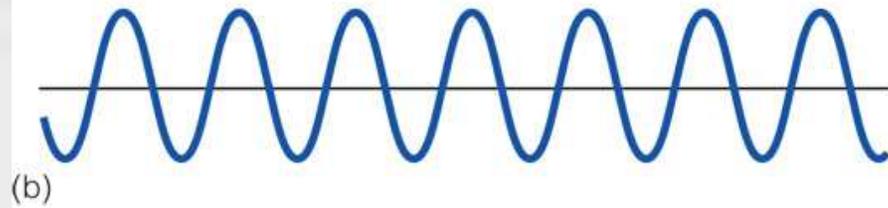
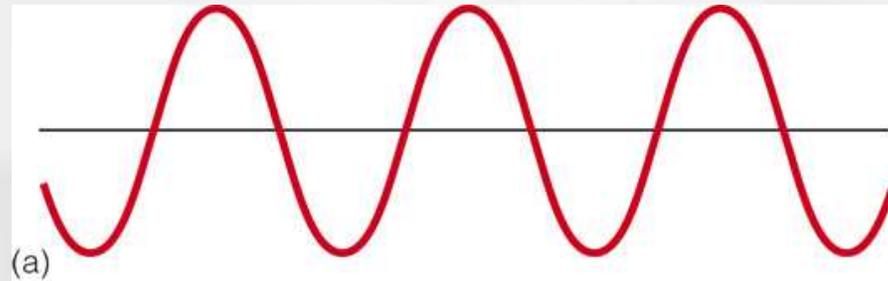
- Equal loudness curves - determined by using a standard 1,000 Hz tone
  - Almost equal loudness at 80 dB
  - High and low frequencies sound softer at 40 dB than the rest of the tones in the range



High frequency hearing loss with age and/or prolonged exposure to loud sounds.



# Additive synthesis - process of adding harmonics to create complex sounds

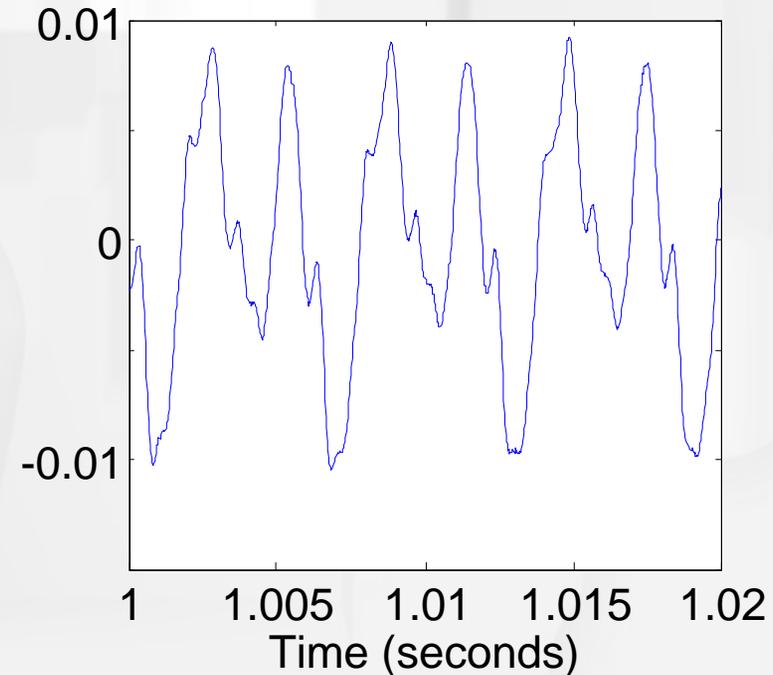
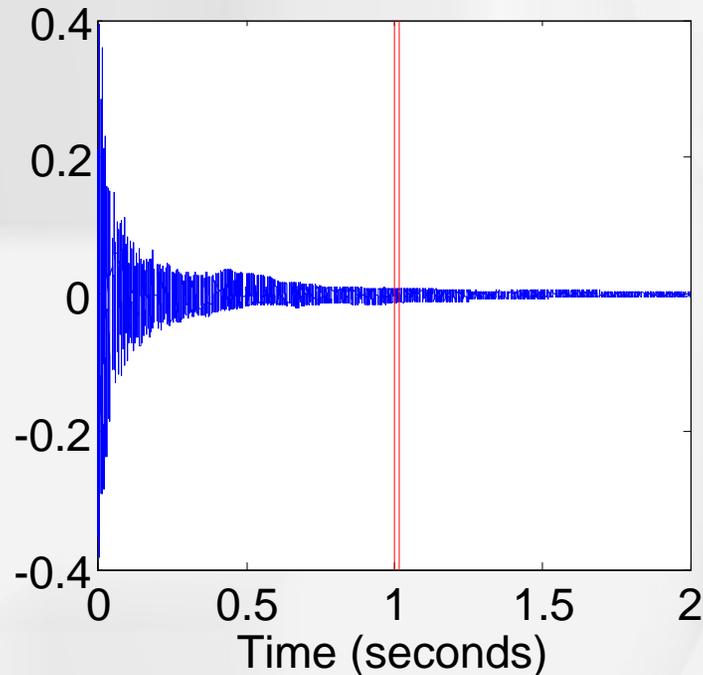


# Sound Quality: Timbre

- All other properties of sound except for loudness and pitch constitute timbre
- Timbre is created partially by the multiple frequencies that make up complex tones
  - Fundamental frequency is the first (lowest) harmonic
  - Musical tones have additional harmonics that are multiples of the fundamental frequency

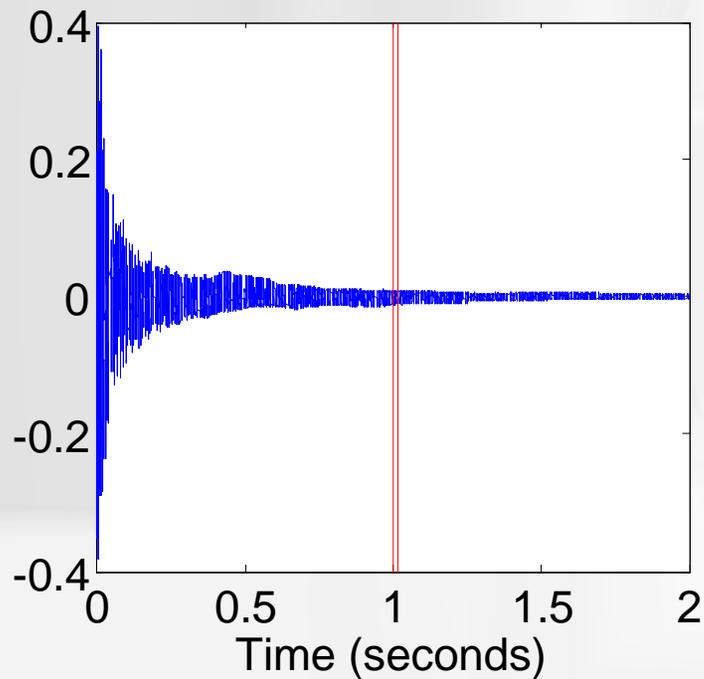


Guitar



# Sound Quality: Timbre

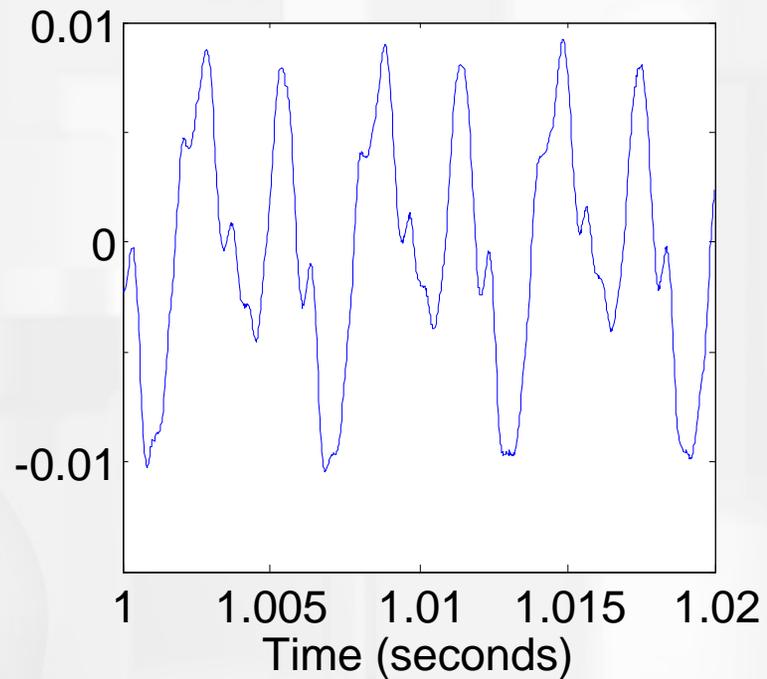
- Frequency spectrum - display of harmonics of a complex sound
- Attack of tones - buildup of sound at the beginning of a tone
- Decay of tones - decrease in sound at end of tone

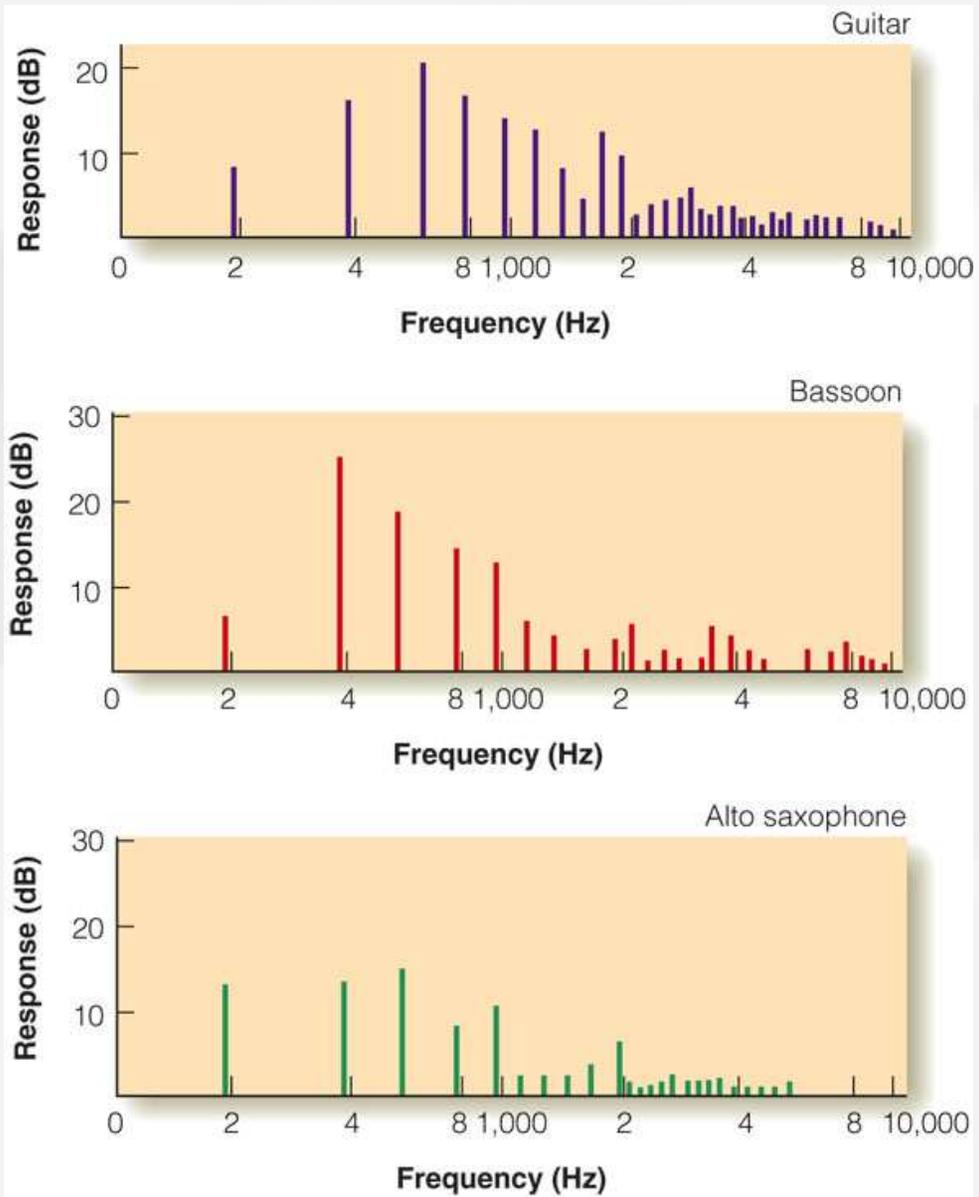


Guitar



Guitar backwards

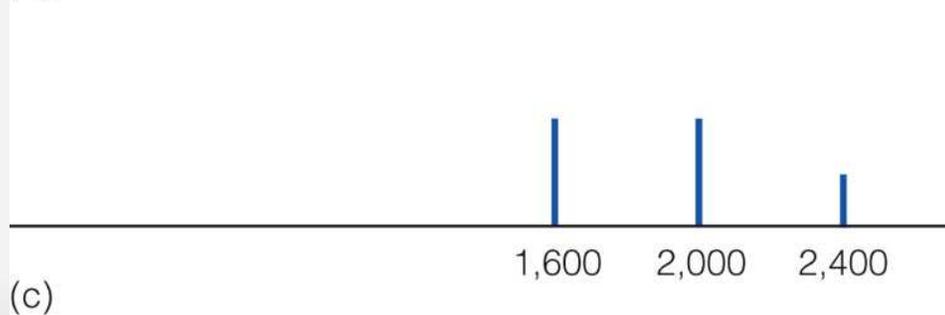
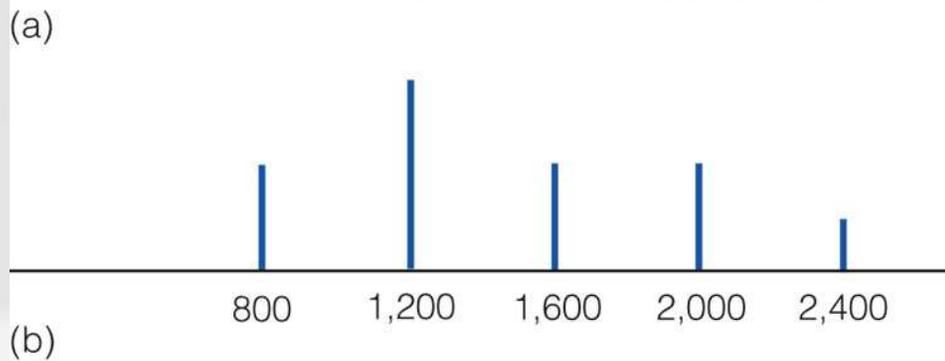
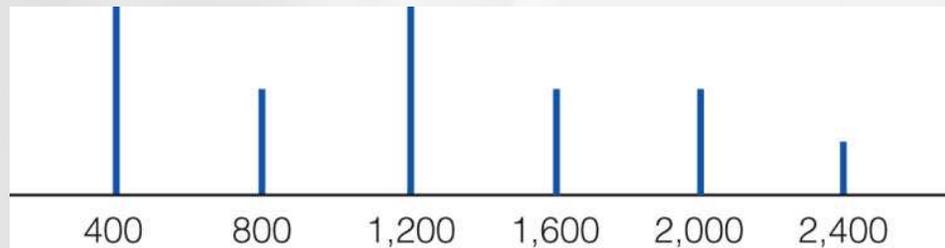




# Effect of the Missing Fundamental

- The fundamental frequency is the lowest frequency in a complex tone
- When the fundamental and other lower harmonics are removed, the perceived pitch is the same, but the timbre changes
- The pitch perceived in such tones is called periodicity pitch

400 Hz



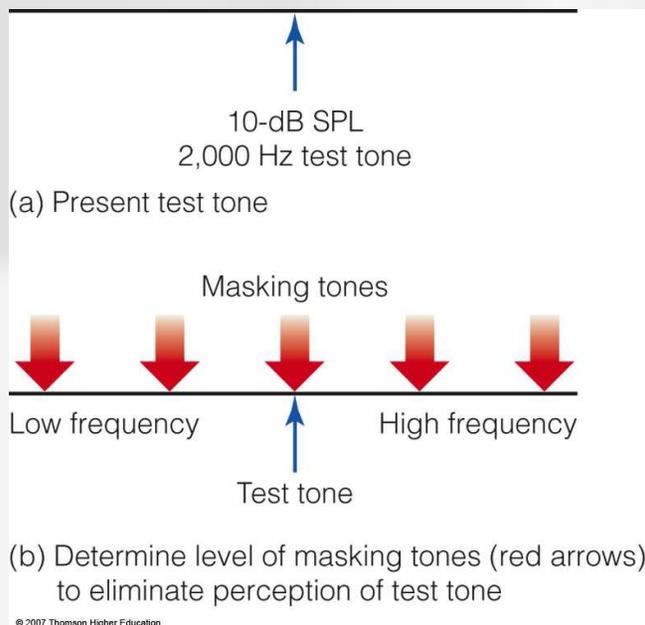
© 2007 Thomson Higher Education

Circular pitch illusion



# Psychophysical Tuning Curves

- Experimental procedure
  - First, a low level test tone is presented
  - Then, masking tones are presented with frequencies above and below the test tone
  - Measures are taken to determine the level of each masking tone needed to eliminate the *perception* of the test tone
  - Assumption is that the masking tones must be causing activity at same location as test tone



Can you hear the 400Hz tone?



400 Hz



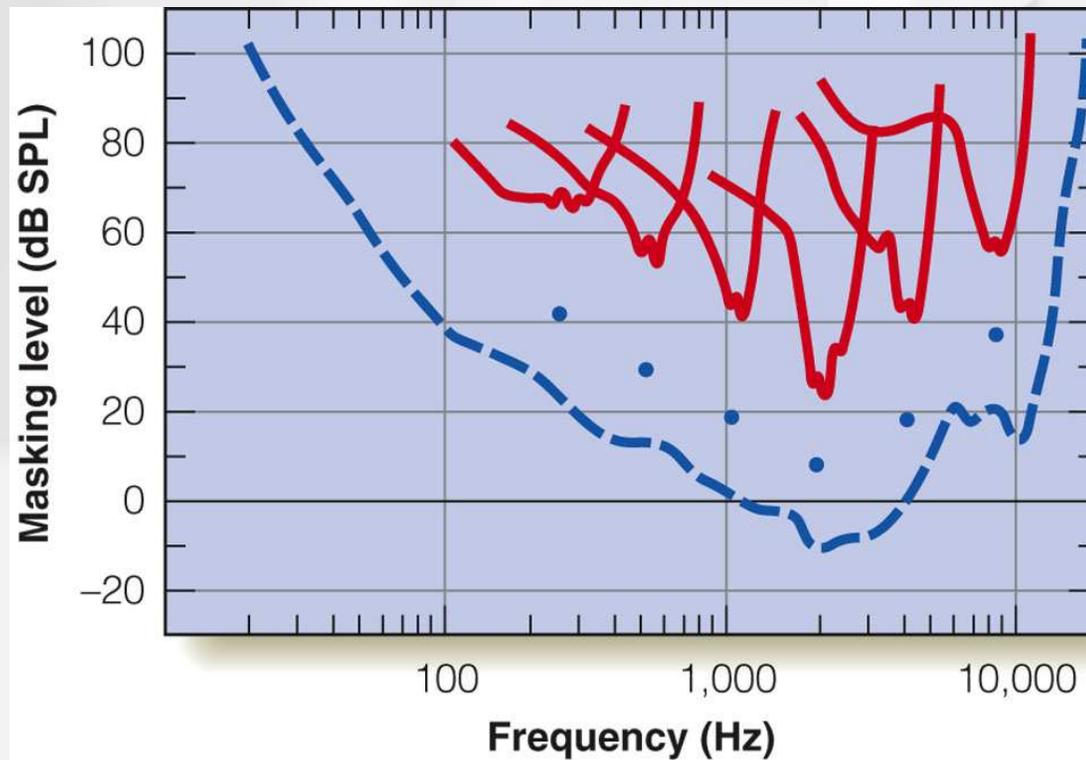
400 & 500 Hz



400 & 2000 Hz

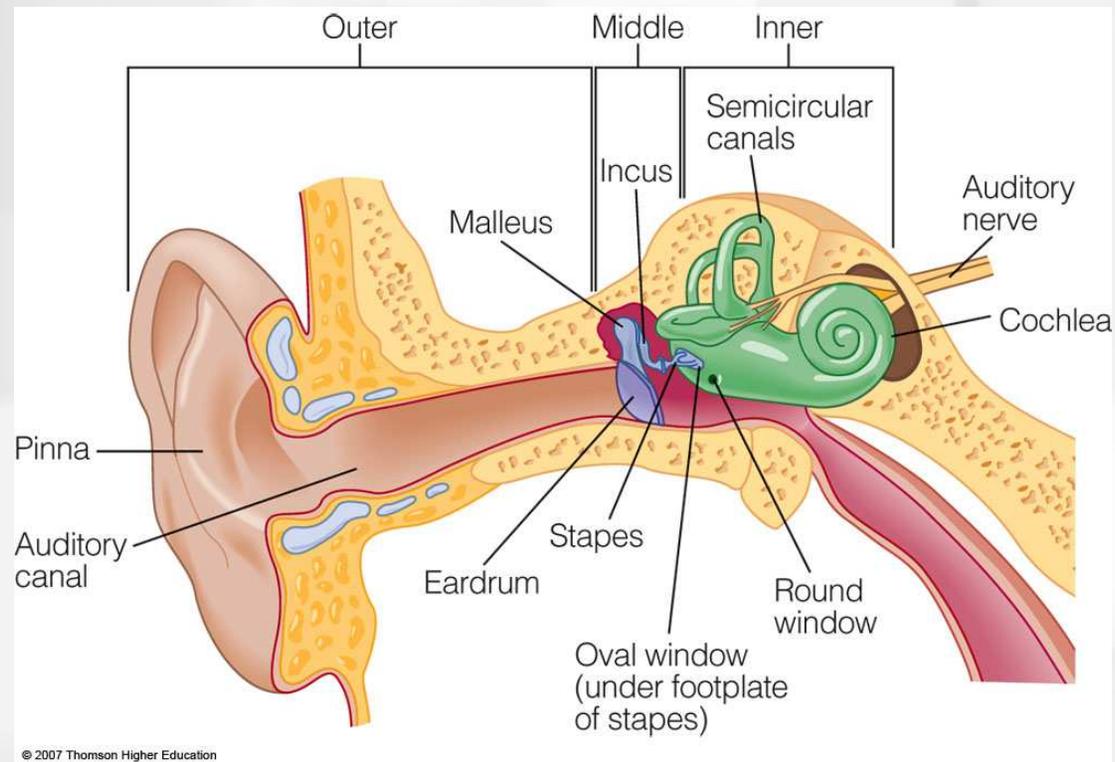
# Psychophysical Tuning Curves

- Resulting tuning curves show that the test tone is affected by a narrow range of masking tones
- Psychophysical tuning curves show the same pattern as neural tuning curves which reveals a close connection between perception and the firing of auditory fibers



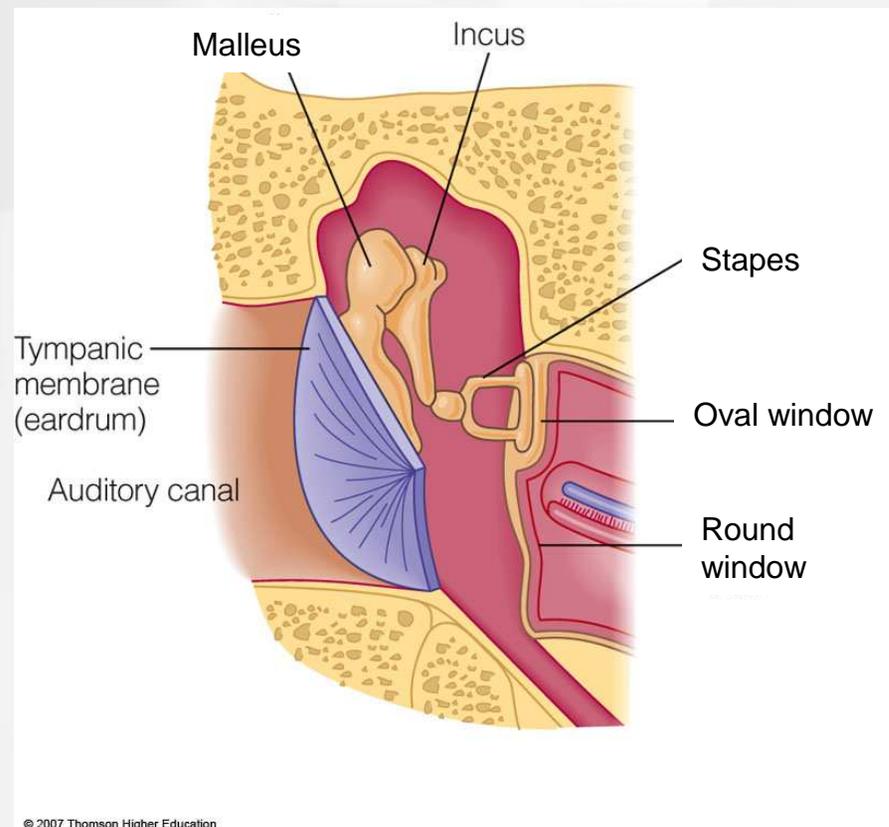
# The Ear

- **Outer ear** - pinna and auditory canal
  - Pinna helps with sound location
  - Auditory canal - tube-like 3 cm long structure
    - Protects the tympanic membrane at the end of the canal
    - Resonant frequency of the canal amplifies frequencies between 2,000 and 5,000 Hz



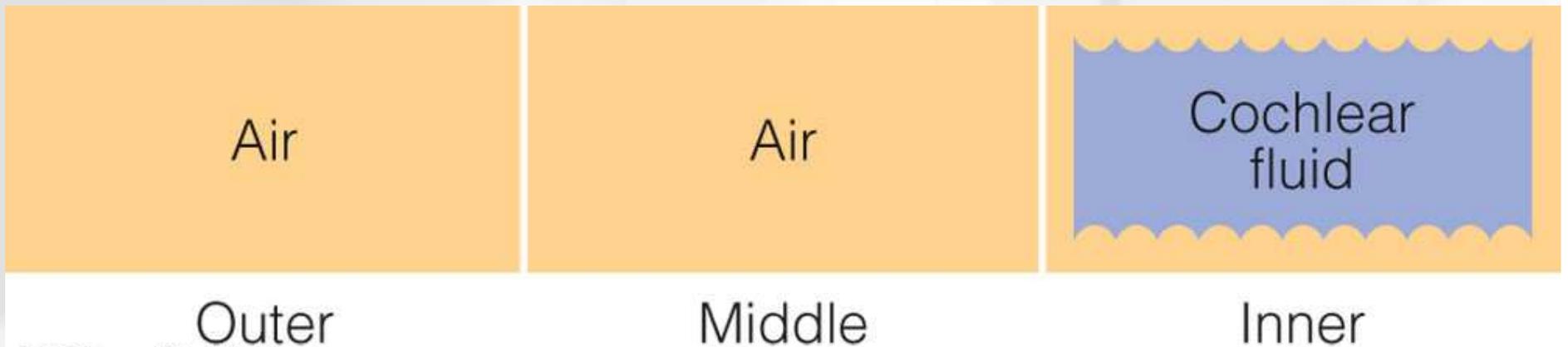
# The Middle Ear

- 2 cubic centimeter cavity separating inner from outer ear
- It contains the three ossicles (the smallest bones in the body!)
  - Malleus - moves due to the vibration of the tympanic membrane
  - Incus - transmits vibrations of malleus
  - Stapes - transmit vibrations of incus to the inner ear via the oval window of the cochlea

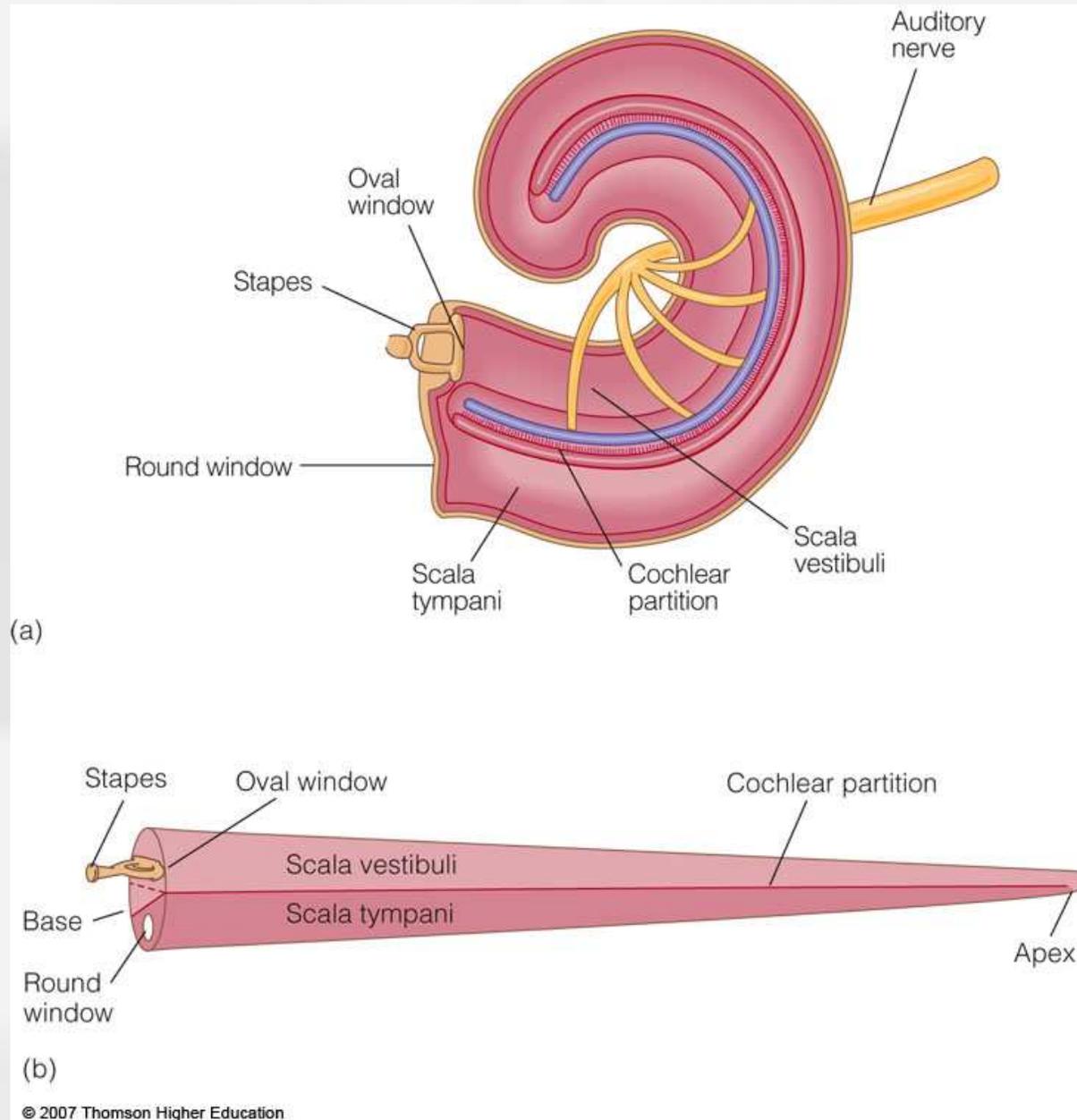


# Function of Ossicles

- Outer and inner ear are filled with air
- Inner ear filled with fluid that is much denser than air
- Pressure changes in air transmit poorly into the denser medium
- Ossicles act to amplify the vibration for better transmission to the fluid

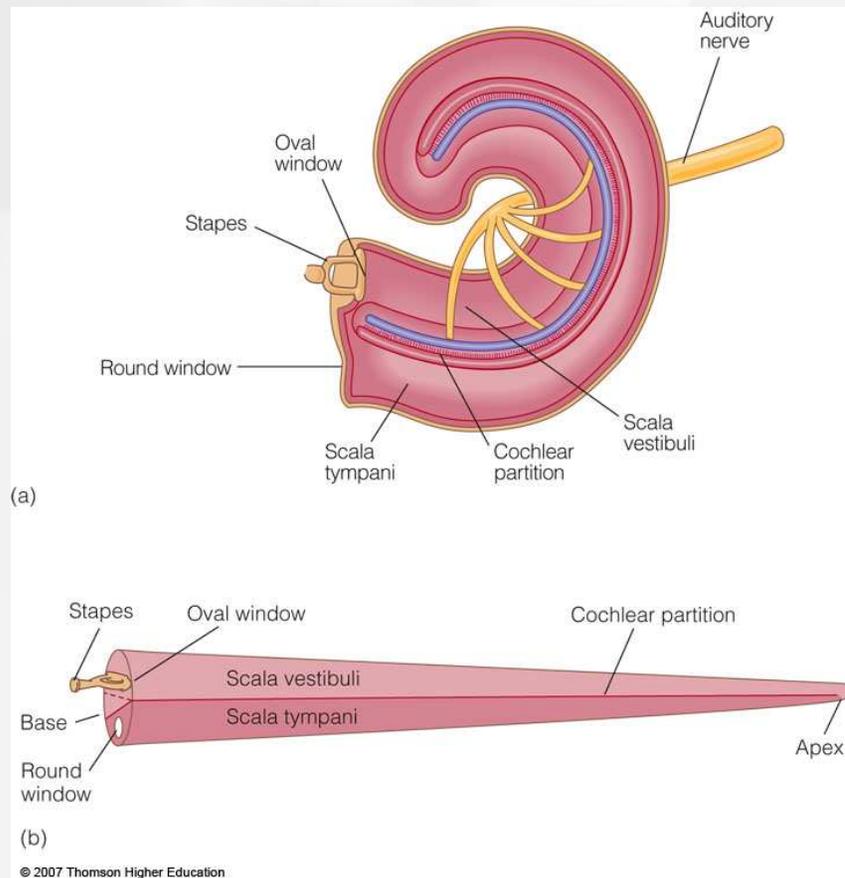


# The Inner Ear



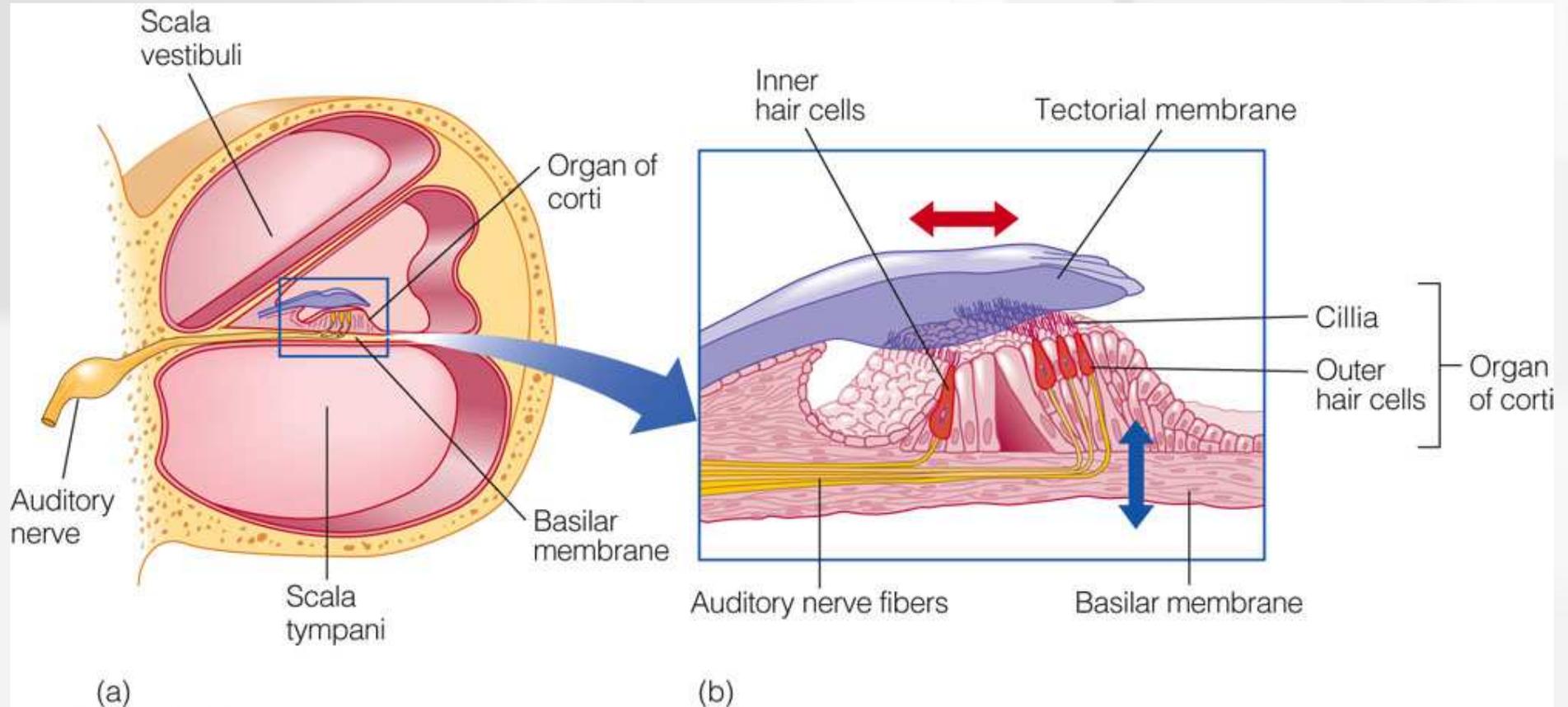
# The Cochlea

- Fluid-filled snail-like structure set into vibration by the stapes
- Divided into the scala vestibuli and scala tympani by the cochlear partition
- Cochlear partition extends from the base (stapes end) to the apex (far end)
- Organ of Corti contained by the cochlear partition



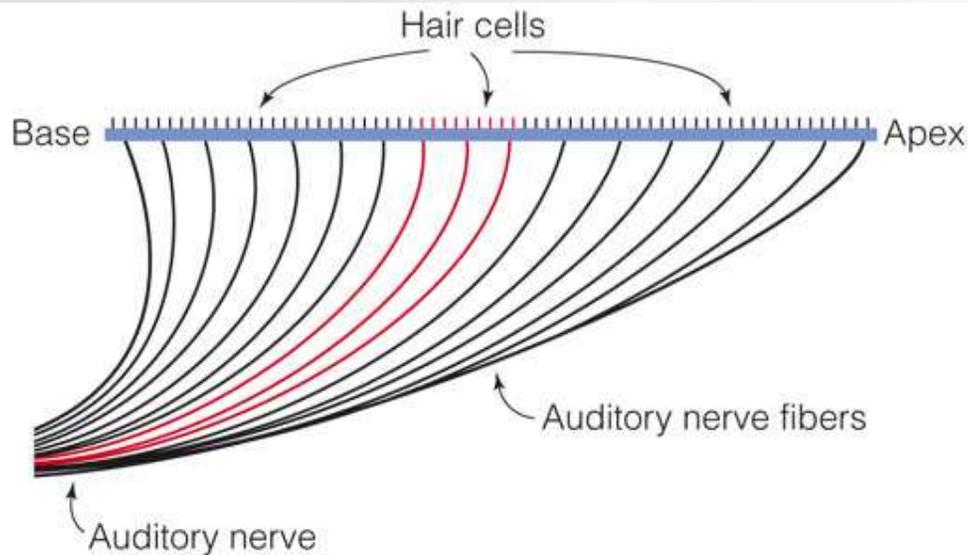
# The Organ of Corti

- Key structures
  - Basilar membrane vibrates in response to sound and supports the organ of Corti
  - Inner and outer hair cells are the receptors for hearing
  - Tectorial membrane extends over the hair cells
- Transduction at the hair cells takes place due to the interaction of these structures

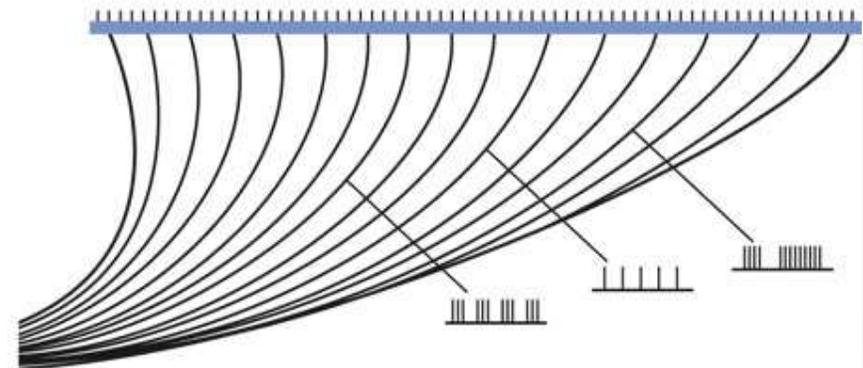


# Neural Signals for Frequency

- There are two ways nerve fibers signal frequency
  - **Which** fibers are responding
    - Specific groups of hair cells on basilar membrane activate a specific set of nerve fibers
  - **How** fibers are firing
    - Rate or pattern of firing of nerve impulses



Which fibers fire



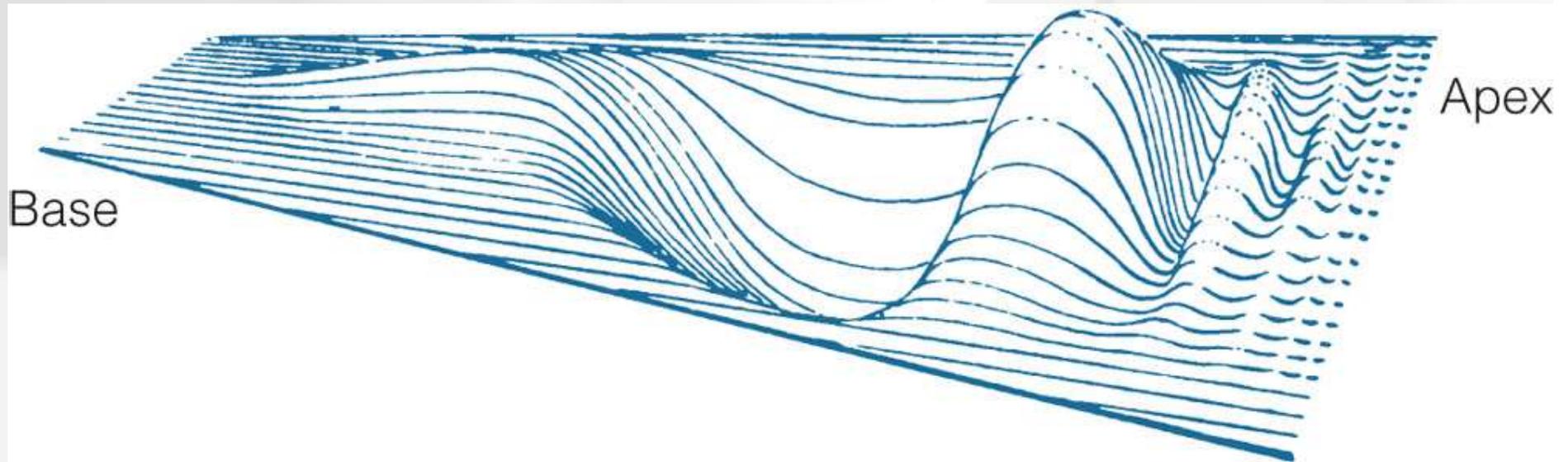
How fibers fire

(a)

(b)

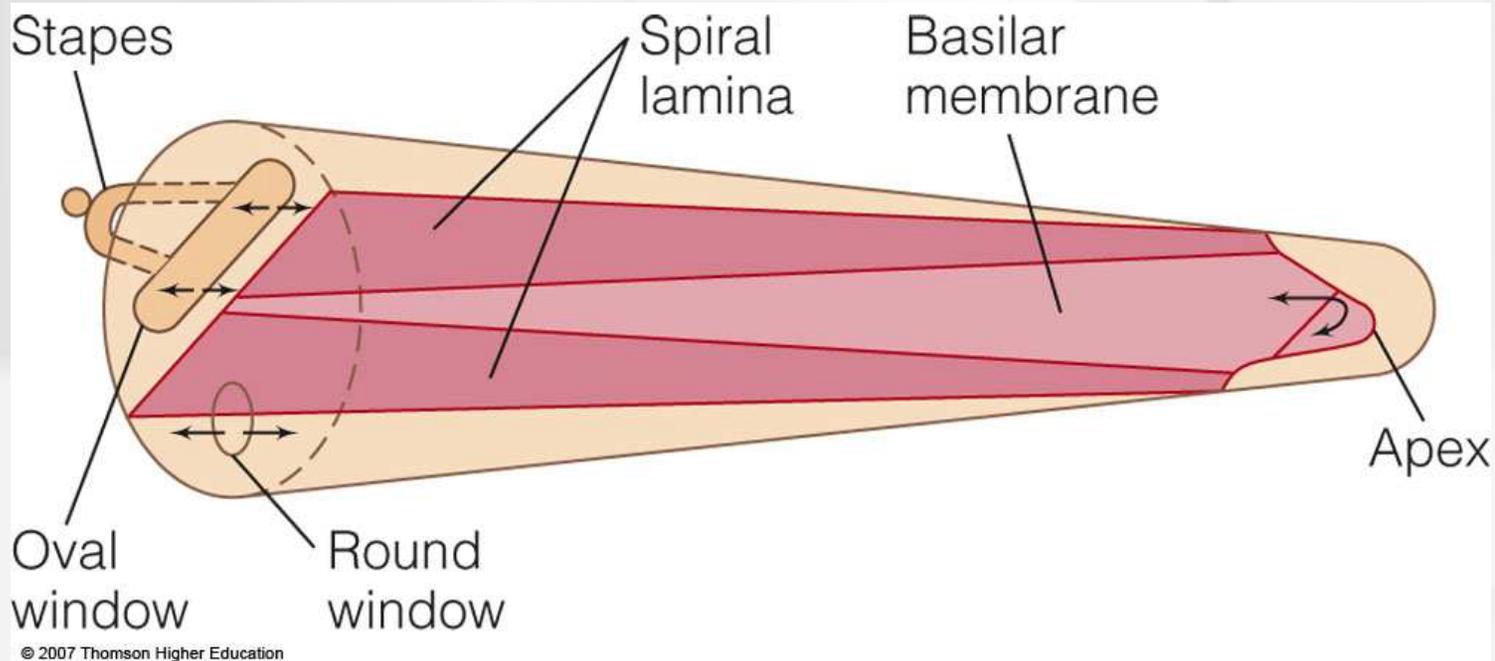
# Békésy's Place Theory of Hearing

- Frequency of sound is indicated by the place on the organ of Corti that has the highest firing rate
- Békésy determined this in two ways
  - Direct observation of the basilar membrane in a cadaver
  - Building a model of the cochlea using the physical properties of the basilar membrane



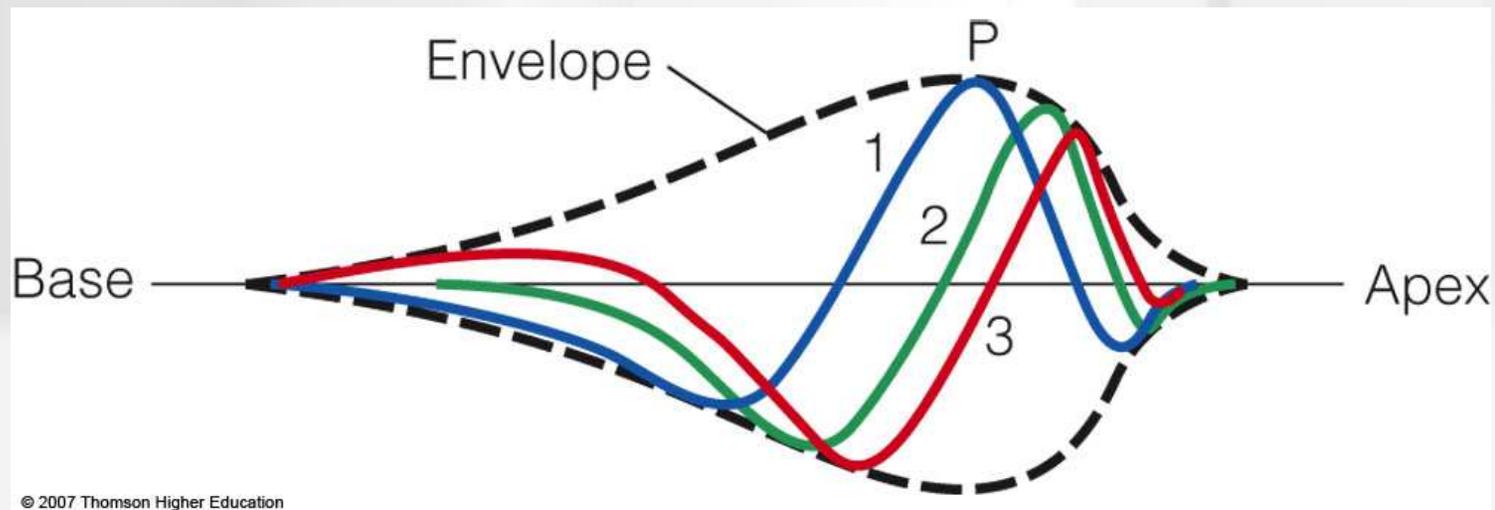
# Békésys' Place Theory of Hearing

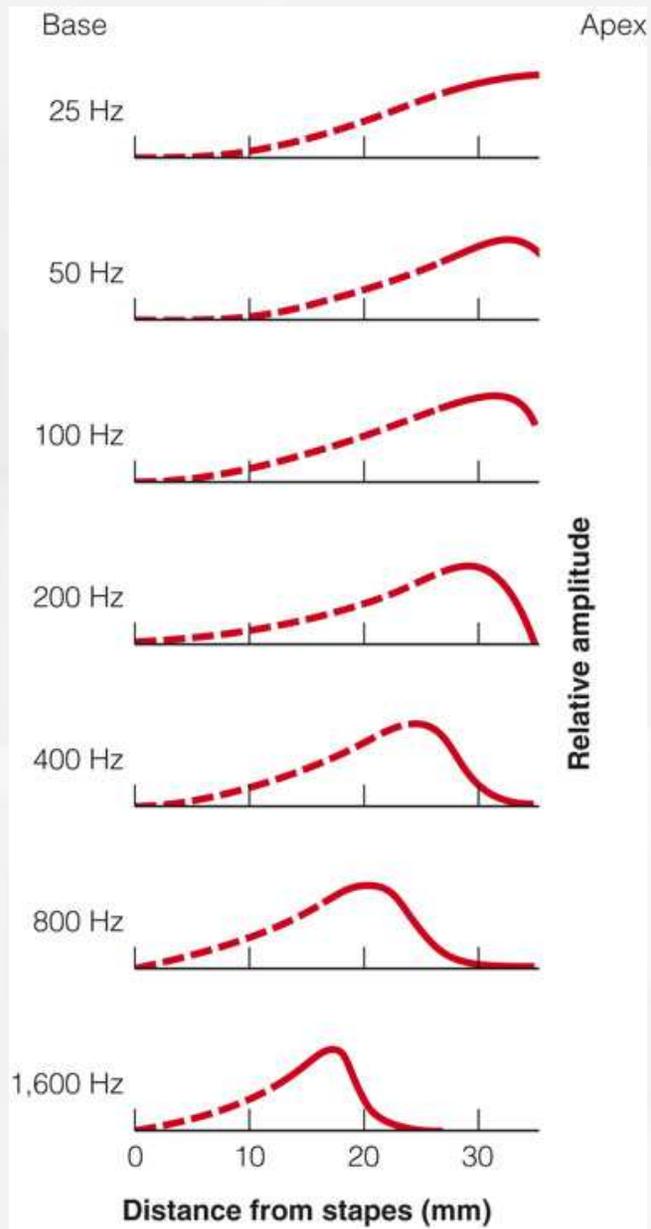
- Physical properties of the basilar membrane
  - Base of the membrane (by stapes) is
    - 3 to 4 times narrower than at the apex
    - 100 times stiffer than at the apex
- Both the model and the direct observation showed that the vibrating motion of the membrane is a traveling wave



# Békésys' Place Theory of Hearing

- Envelope of the traveling wave
  - Indicates the point of maximum displacement of the basilar membrane
  - Hair cells at this point are stimulated the most strongly leading to the nerve fibers firing the most strongly at this location
  - Position of the peak is a function of frequency





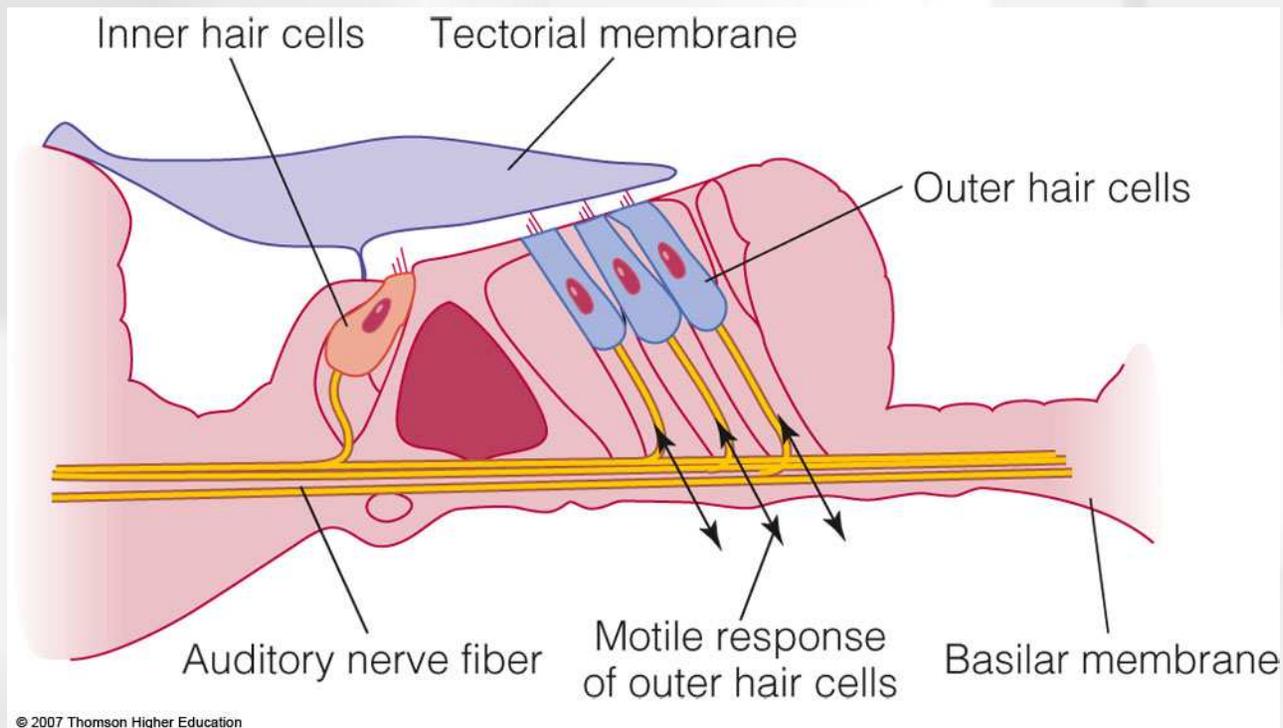


# Evidence for Place Theory

- Neural frequency tuning curves
  - Pure tones are used to determine the threshold for specific frequencies measured at single neurons
  - Plotting thresholds for frequencies results in tuning curves
  - Frequency to which the neuron is most sensitive is the characteristic frequency

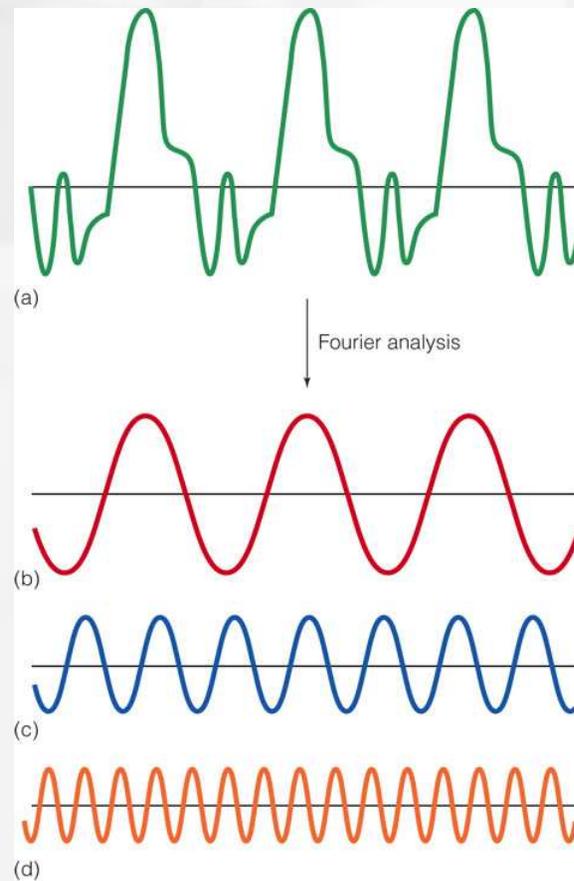
# Updating Békésy's Place Theory

- Békésy used unhealthy basilar membranes and his results showed no difference in response for close frequencies that people can distinguish.
- New research with healthy membranes show that the entire **outer hair cells** respond to sound by slight tilting and a change in length
  - This is called the motile response and helps to amplify action on the membrane

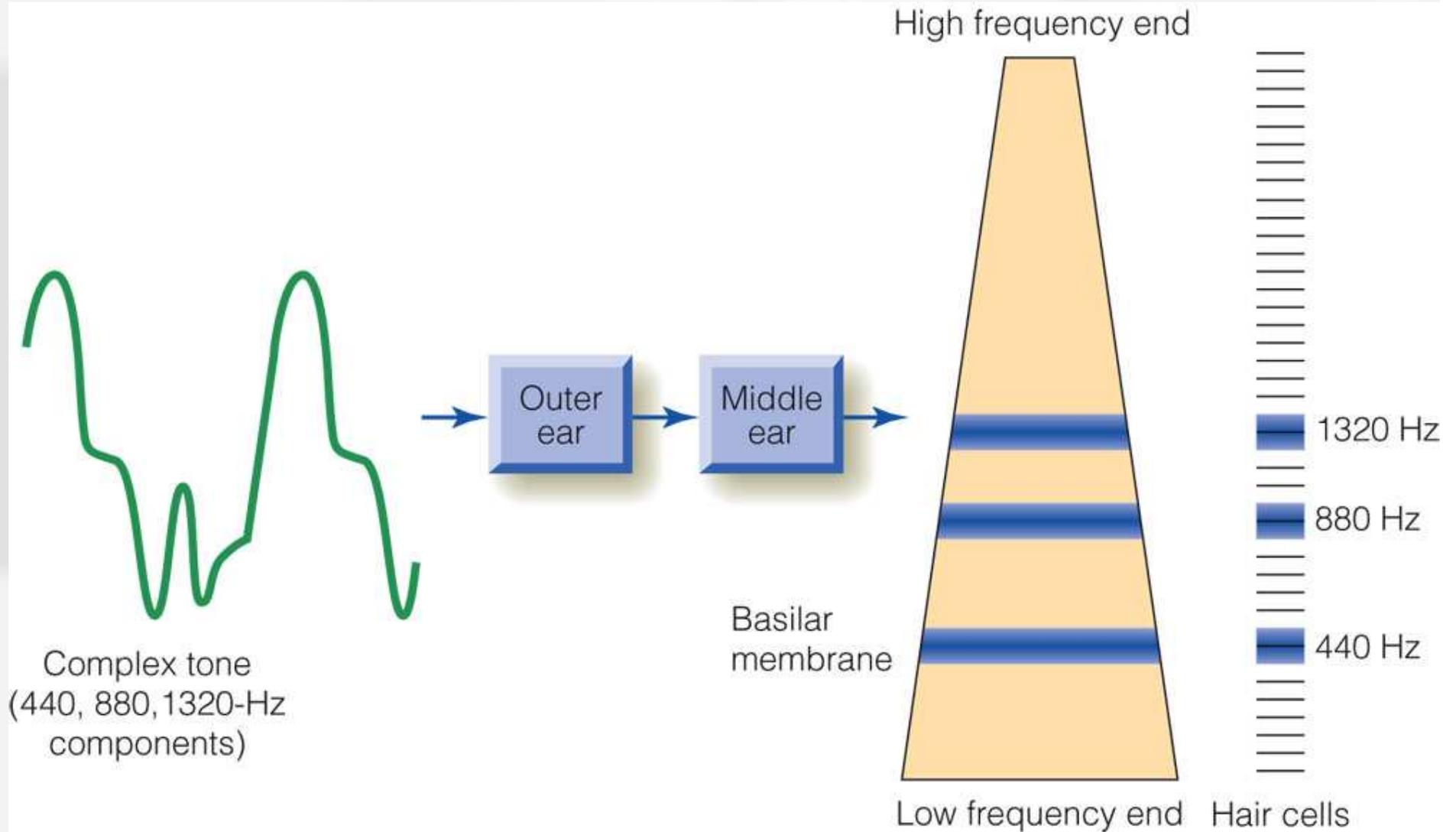


# Response of Basilar Membrane to Complex Tones

- Fourier analysis - mathematic process that separates complex waveforms into a number of sine waves
- Research on the response of the basilar membrane shows the highest response in auditory nerve fibers with characteristic frequencies that correspond to the sine-wave components of complex tones
- Thus the cochlea is called a frequency analyzer

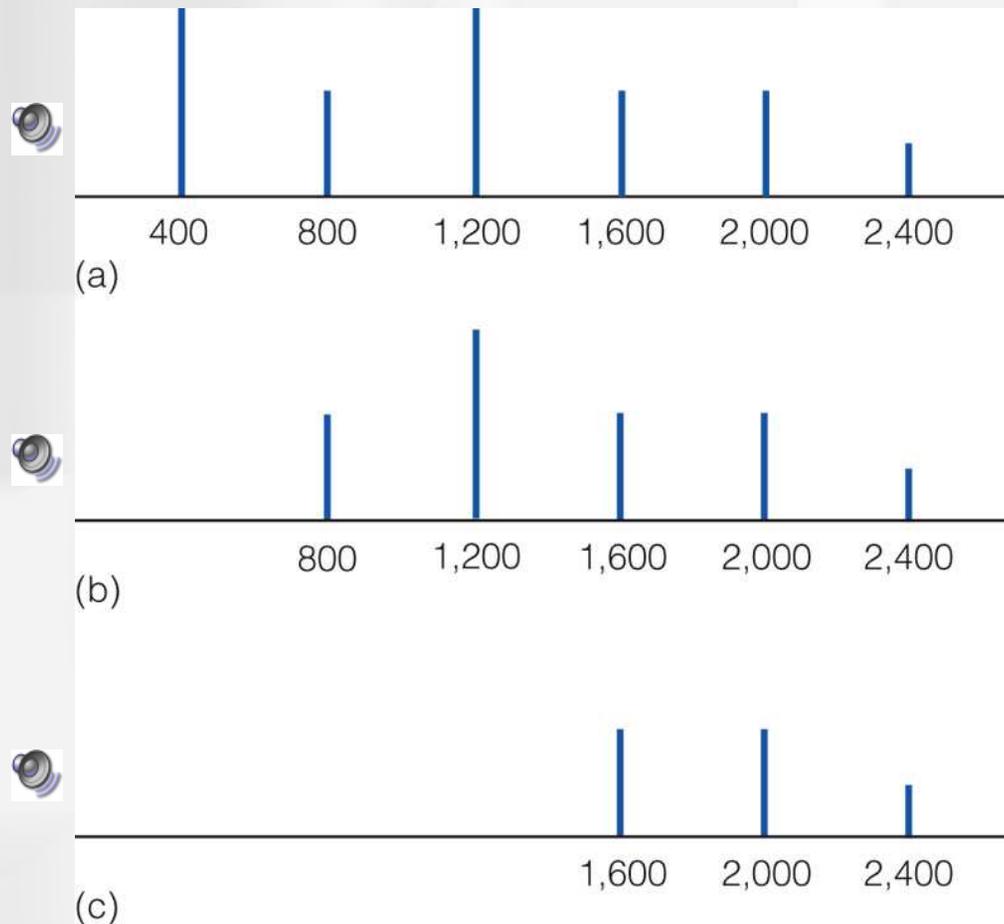


The Cochlea automatically breaks down complex tones into their component frequencies – it's performing Fourier analysis.



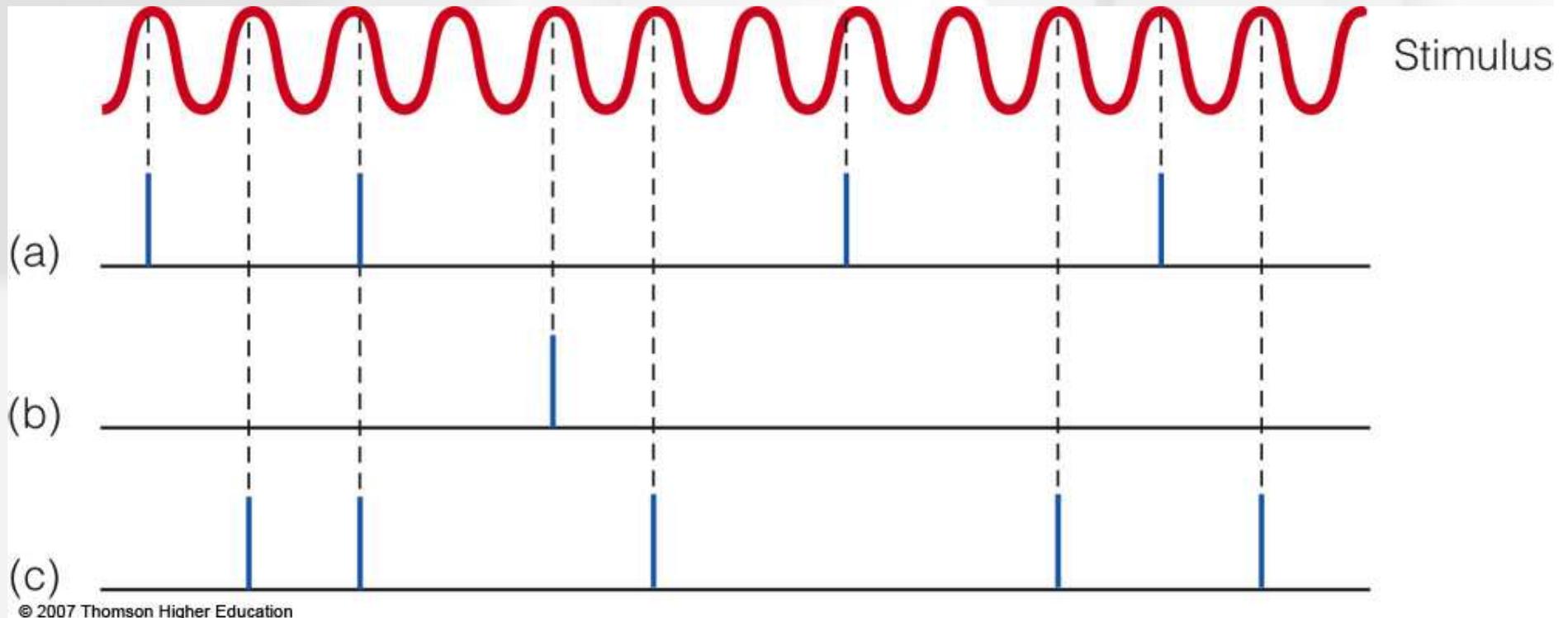
# Missing Fundamental: evidence against Place Theory

- Pattern of stimulation on the basilar membrane cannot explain this phenomenon since removing the fundamental and harmonics creates different patterns
- Periodicity pitch is perceived even when the tones are presented to two ears



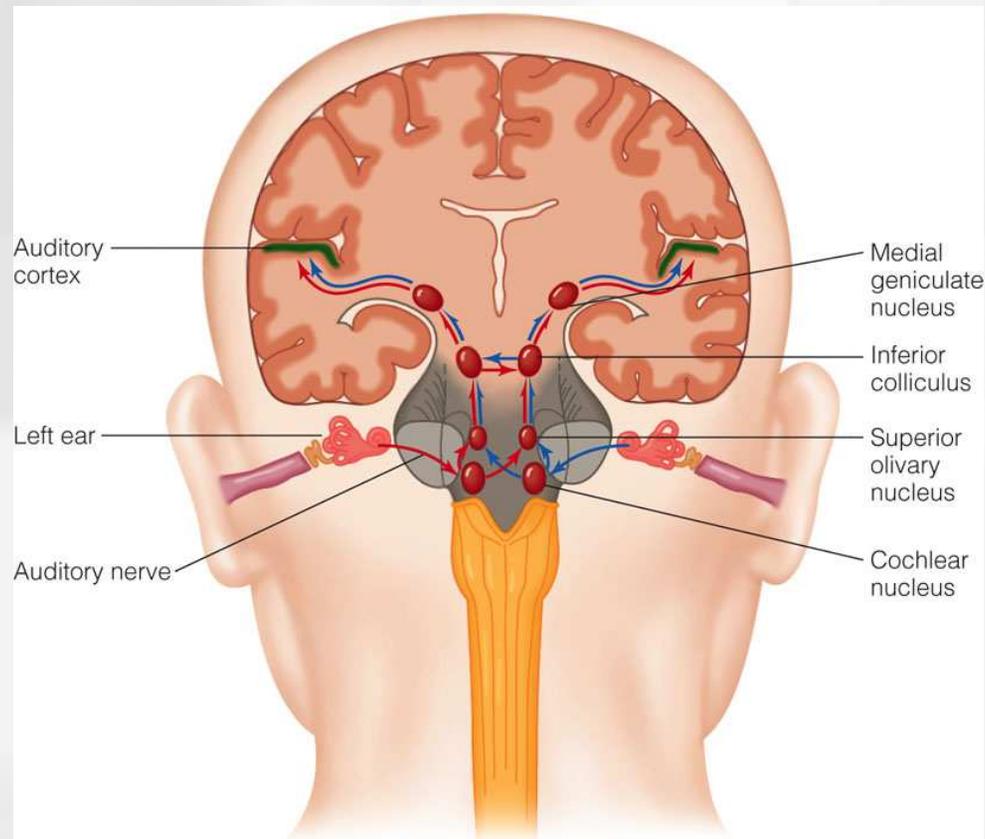
# Timing of Neural Firing and Frequency

- Phase locking
  - Nerve fibers fire in bursts
  - Firing bursts happen at or near the peak of the sine-wave stimulus
  - Thus, they are “locked in phase” with the wave
  - Groups of fibers fire with periods of silent intervals creating a pattern of firing



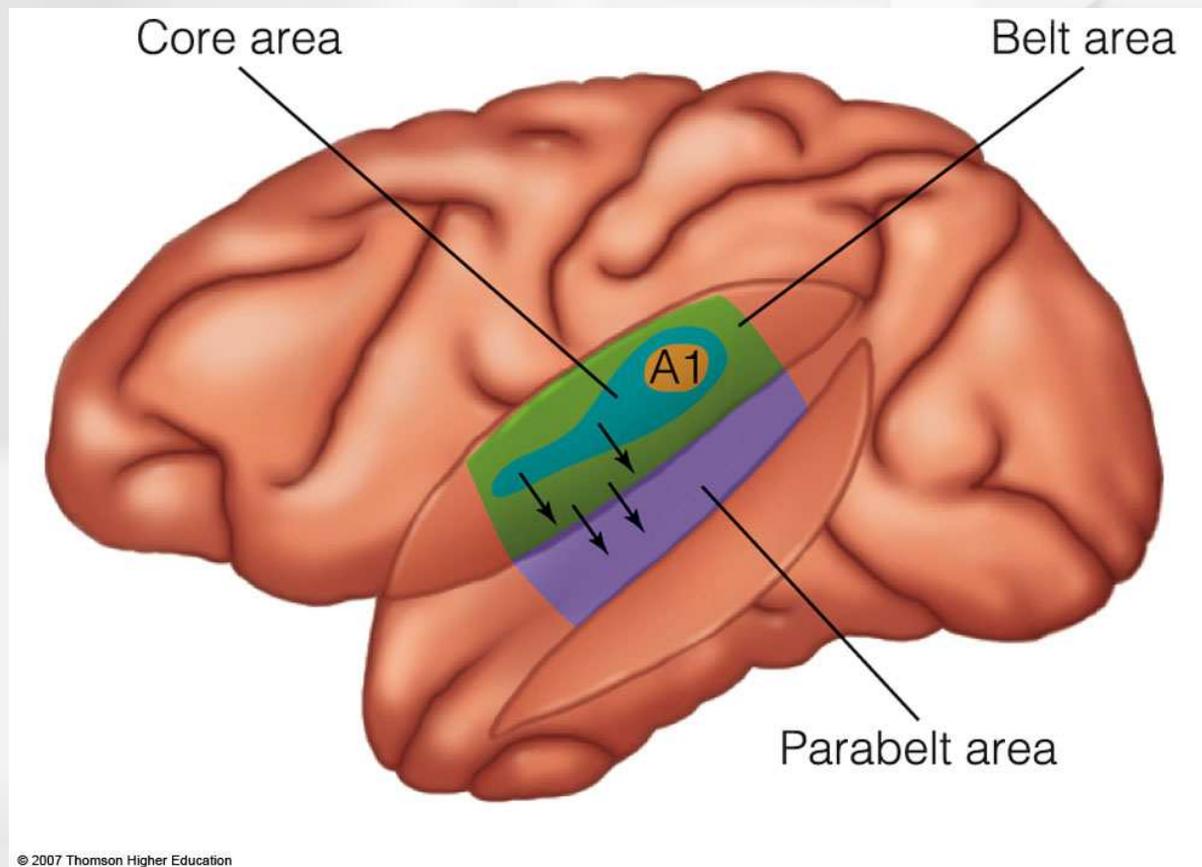
# Pathway from the Cochlea to the Cortex

- Auditory nerve fibers synapse in a series of subcortical structures
  - Cochlear nucleus
  - Superior olivary nucleus (in the brain stem)
  - Inferior colliculus (in the midbrain)
  - Medial geniculate nucleus (in the thalamus)
  - Auditory receiving area (A1 in the temporal lobe)



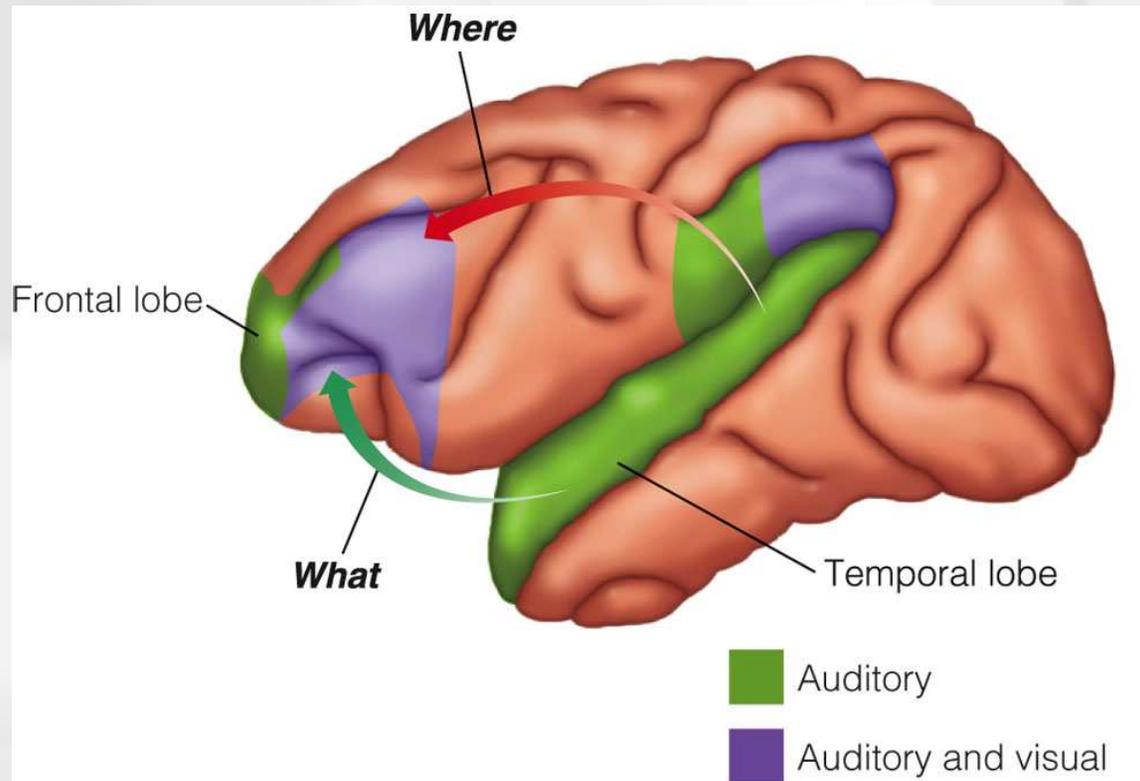
# Auditory Areas in the Cortex

- Hierarchical processing occurs in the cortex
  - Neural signals travel through the core, then belt, followed by the parabelt area
  - Simple sounds cause activation in the core area
  - Belt and parabelt areas are activated in response to more complex stimuli made up of many frequencies



# What and Where Streams for Hearing

- *What* or ventral stream starts in the anterior portion of the core and belt and extends to the prefrontal cortex
  - It is responsible for identifying sounds
- *Where* or dorsal stream starts in the posterior core and belt and extends to the parietal and prefrontal cortices
  - It is responsible for locating sounds
- Evidence from neural recordings, brain damage, and brain scanning support these findings



# Perceiving Pitch and Complex Sounds

- Tonotopic maps are found in A1
  - Neurons that respond better to low frequencies are on the left and those that respond best to high frequencies are on the right
  - However, early research did not show a direct relationship between pitch perception and the tonotopic map

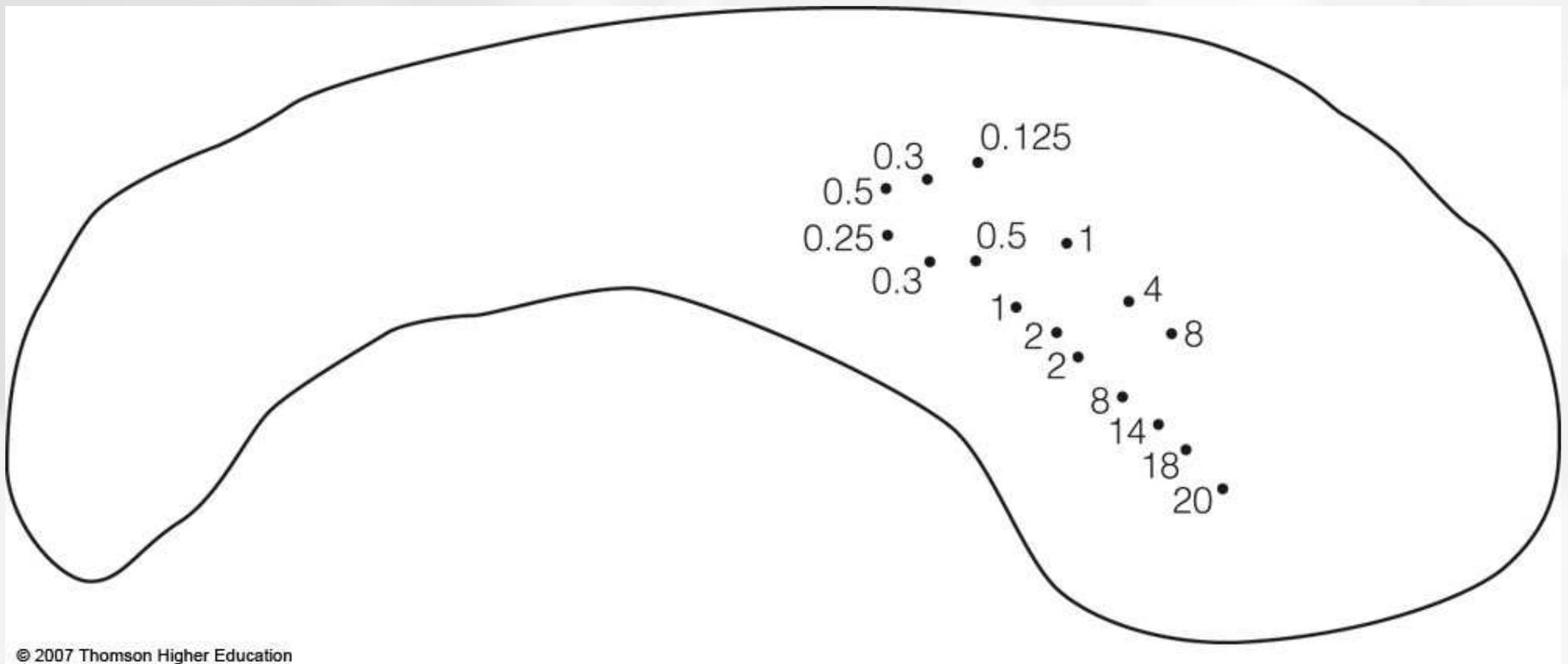
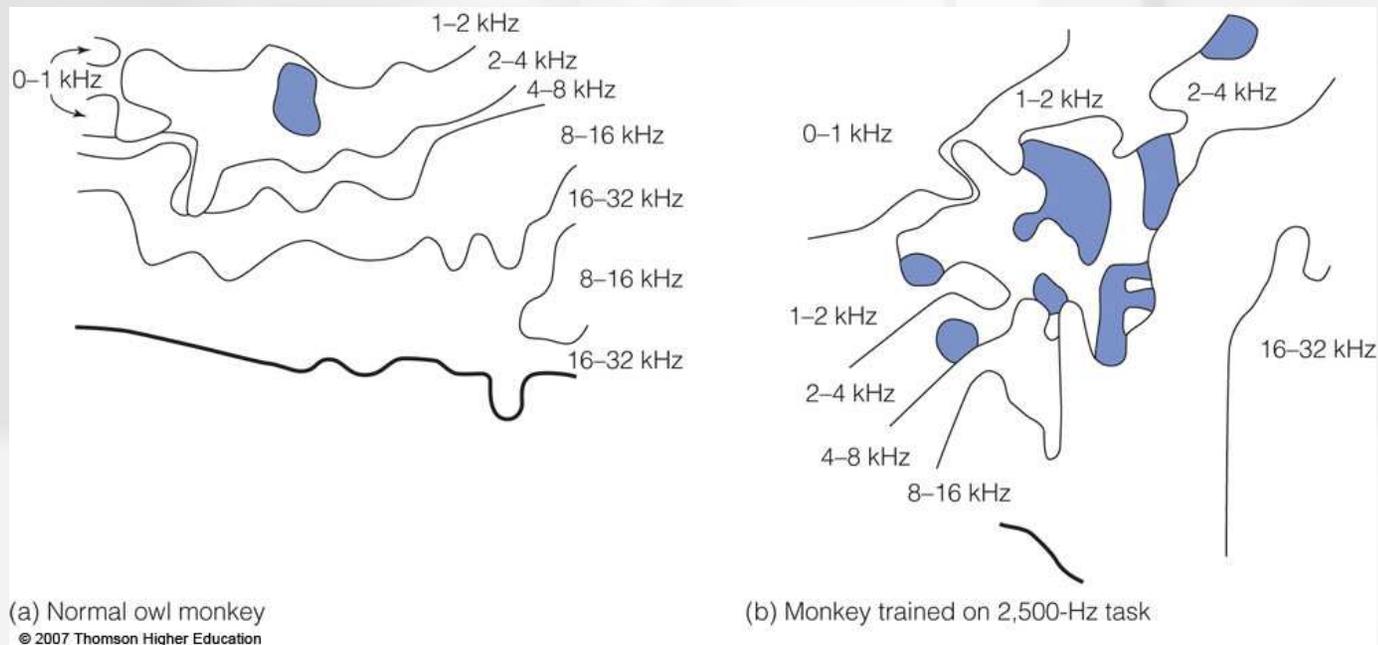


Figure 11.41 The outline of the core area of the monkey auditory cortex, showing the tonotopic map in the primary auditory receiving area, A1, which is located within the core. The numbers represent the characteristic frequencies (CF) of neurons in thousands of Hz. Low CF's are on the left, and high CF's are on the right. (Adapted from Kosaki et al., 1997).

# Recent Evidence of Pitch Perception in A1

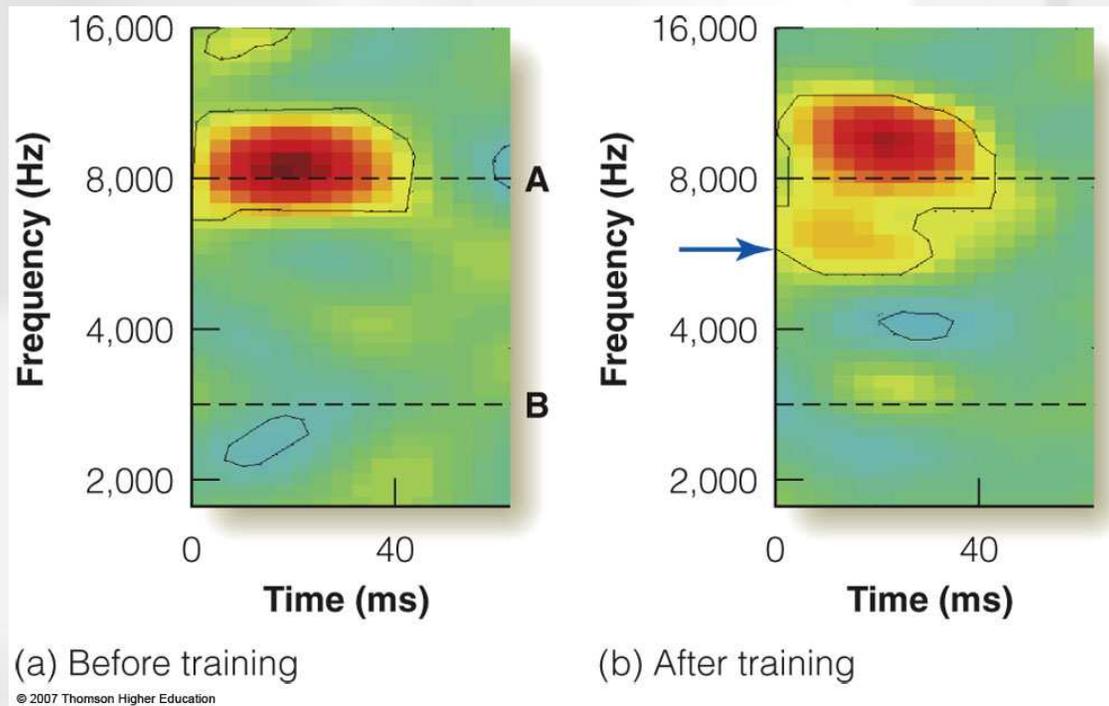
- Effect of training on tonotopic maps
  - Owl monkeys were trained to discriminate between two frequencies near 2,500 Hz
  - Trained monkeys showed tonotopic maps with enlarged areas with neurons that responded to 2,500 Hz compared to untrained monkeys



(a) Tonotopic map of the owl monkey's primary auditory receiving area (A1), showing areas that contain neurons with the characteristic frequencies indicated. The blue area contains neurons with CF = 2,500 Hz. (b) Tonotopic map of an owl monkey that was trained to discriminate between frequencies near 2,500 Hz. The blue areas indicate that after training more of the cortex responds best to 2,500 Hz. (From Recanzone *et al.*, 1993).

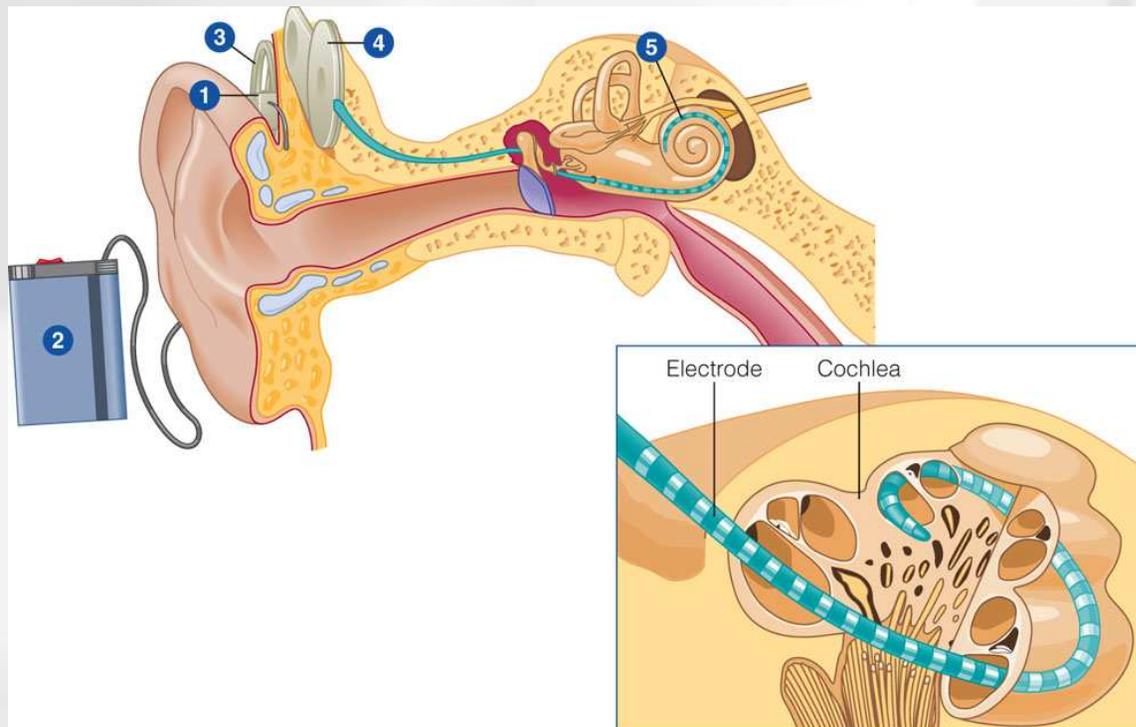
# Effect of Experience on the Auditory Cortex

- Musicians show enlarged auditory cortices that respond to piano tones and stronger neural responses than non-musicians
- Experiment by Fritz et al.
  - Marmosets were trained to lick a water spout in response to a pure tone embedded within a stream of complex tones
  - Neurons became quickly tuned to the target frequency and maintained the effect for hours after the testing session



# Cochlear Implants

- Electrodes are inserted into the cochlea to electrically stimulate auditory nerve fibers
- The device is made up of
  - A microphone worn behind the ear
  - A sound processor
  - A transmitter mounted on the mastoid bone
  - A receiver surgically mounted on the mastoid bone



# Cochlear Implants

- Implants stimulate the cochlea at different places on the tonotopic map according to specific frequencies in the stimulus
- These devices help deaf people to hear some sounds and to understand language
- They work best for people who receive them early in life or for those who have lost their hearing, although they have caused some controversy in the deaf community

