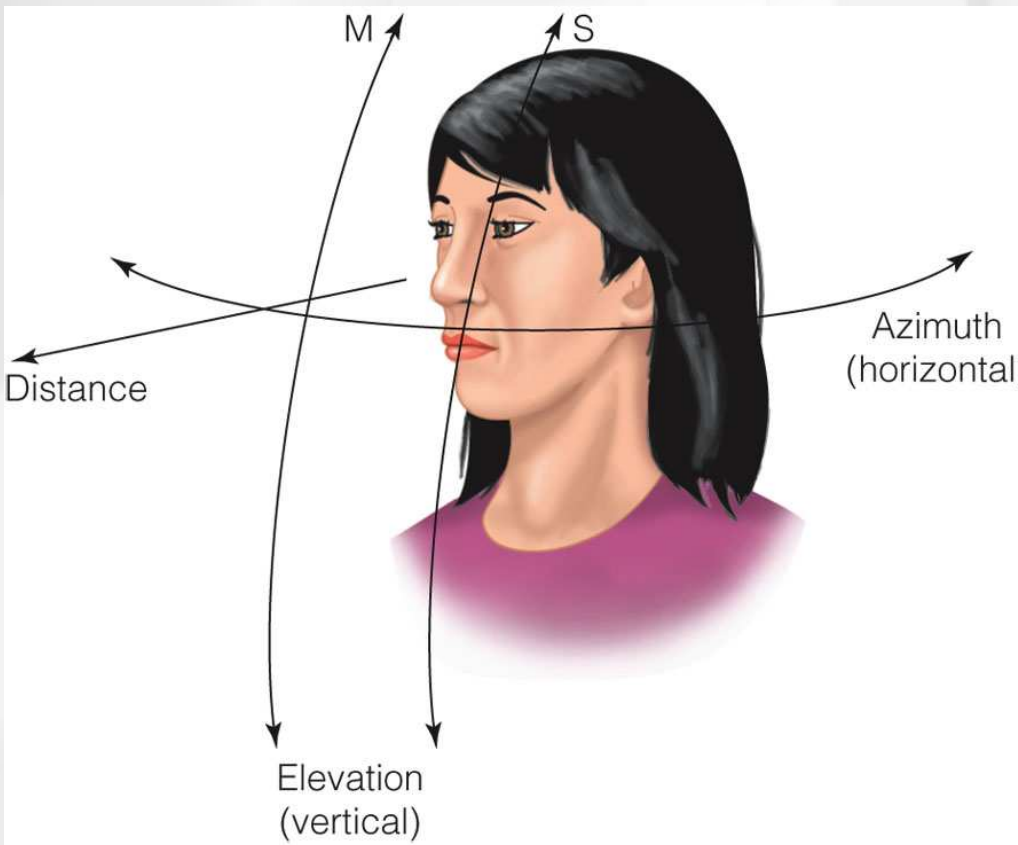
Chapter 12: Sound Localization and the Auditory Scene

- What makes it possible to tell where a sound is coming from in space?
- When we are listening to a number of musical instruments playing at the same time, how can we perceptually separate the sounds coming from the different instruments?
- Why does music sound better in some concert halls than in others?

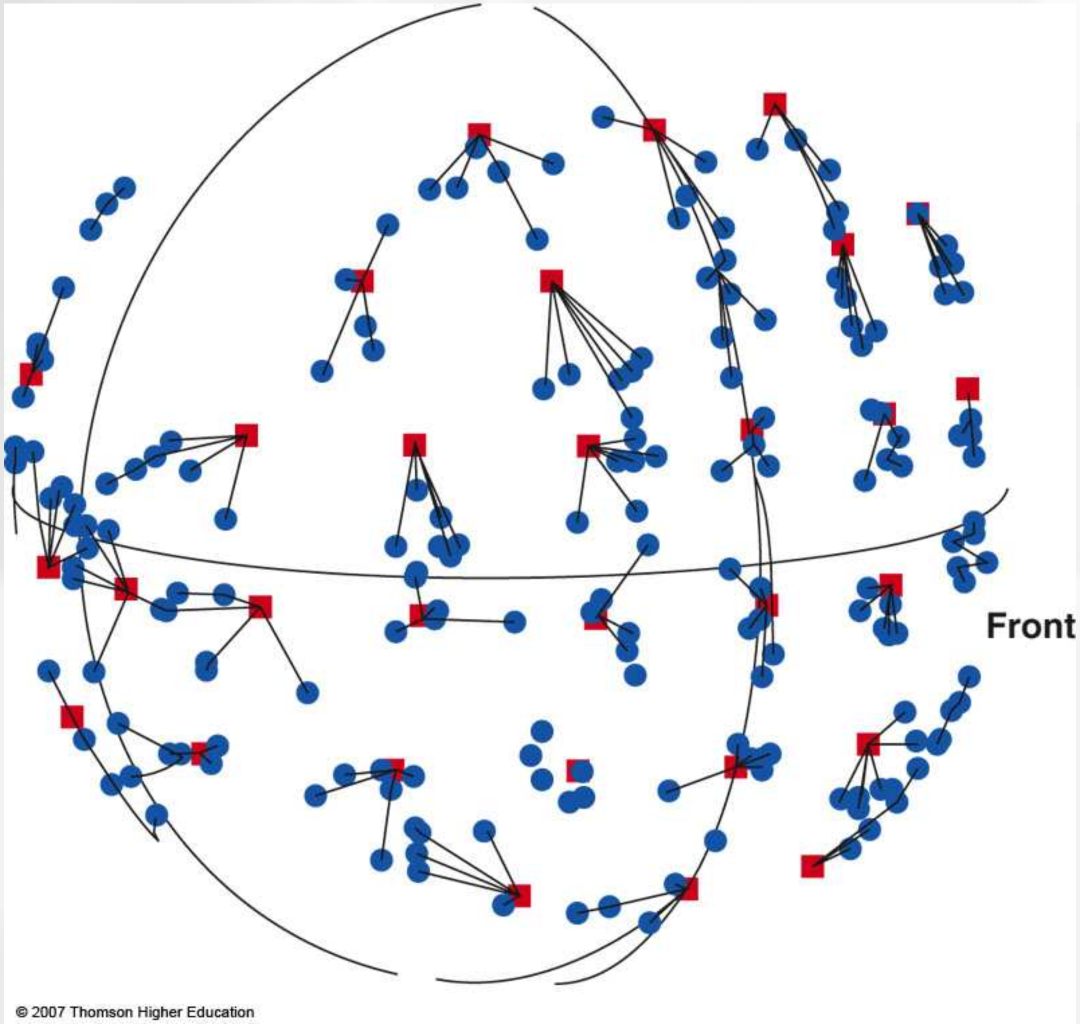
Auditory Localization; the ‘Where’ pathway for the auditory system

- Auditory space - surrounds an observer and exists wherever there is sound
- Researchers study how sounds are localized in space by using
  - Azimuth coordinates - position left to right
  - Elevation coordinates - position up and down
  - Distance coordinates - position from observer

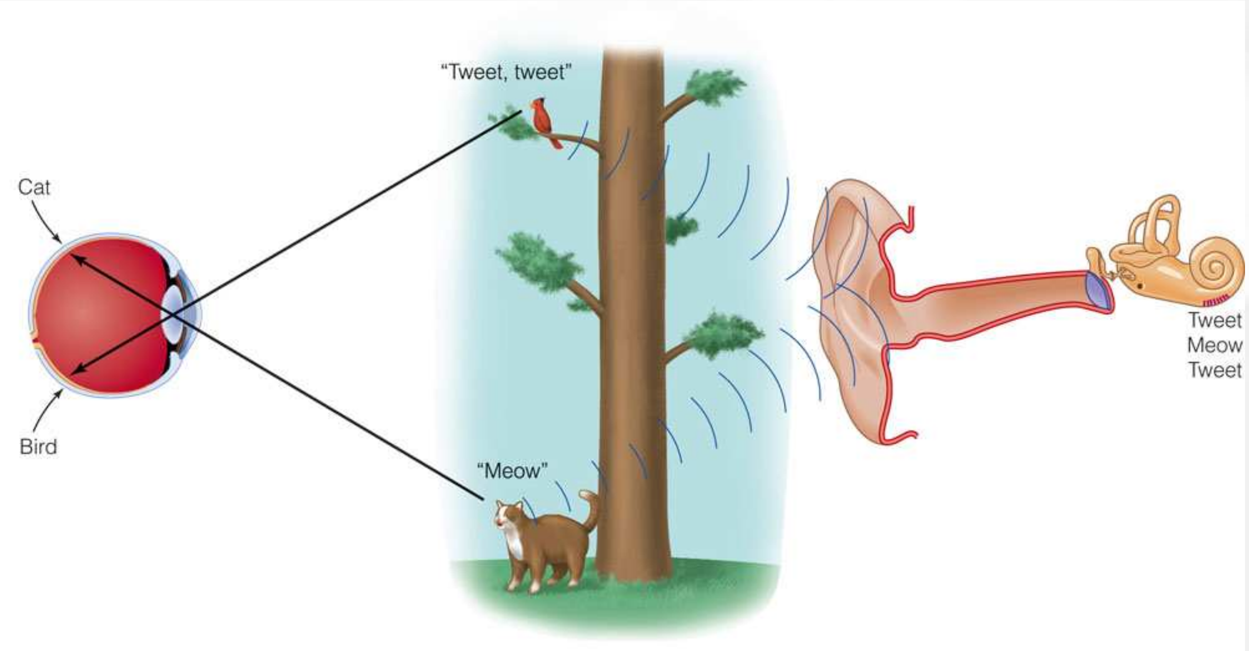


Auditory Localization

- On average, people can localize sounds
  - Directly in front of them most accurately
  - To the sides and behind their heads least accurately



Location cues are not contained in the receptor cells like on the retina in vision; location for sounds must be calculated through other cues.



3 primary cues for auditory localization:

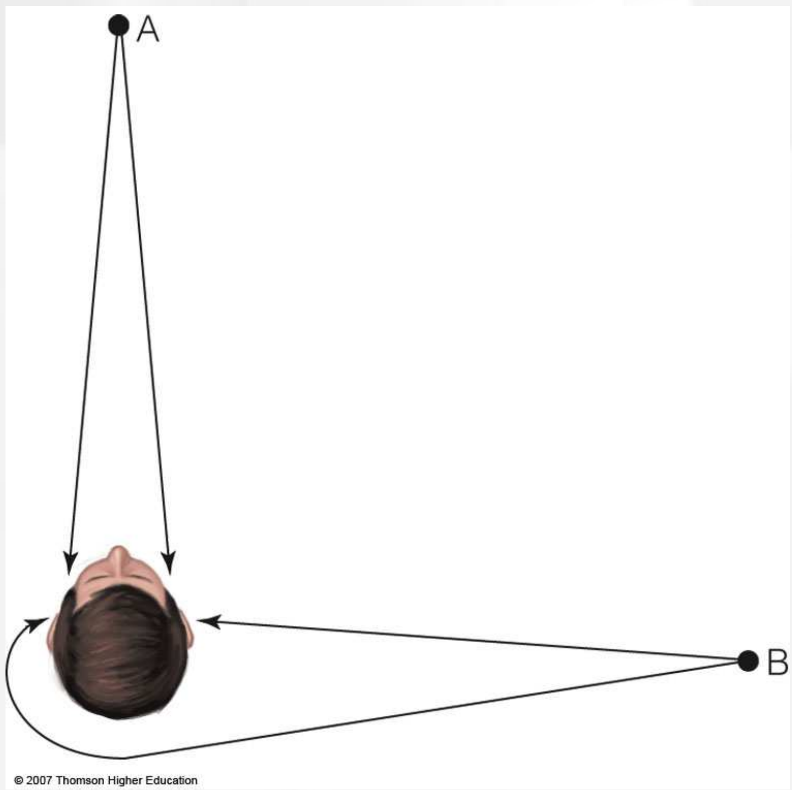
1. Interaural time difference (ITD)
2. Interaural level difference (ILD)
3. Head-related transfer function (HRTF)

Cues for Auditory Location

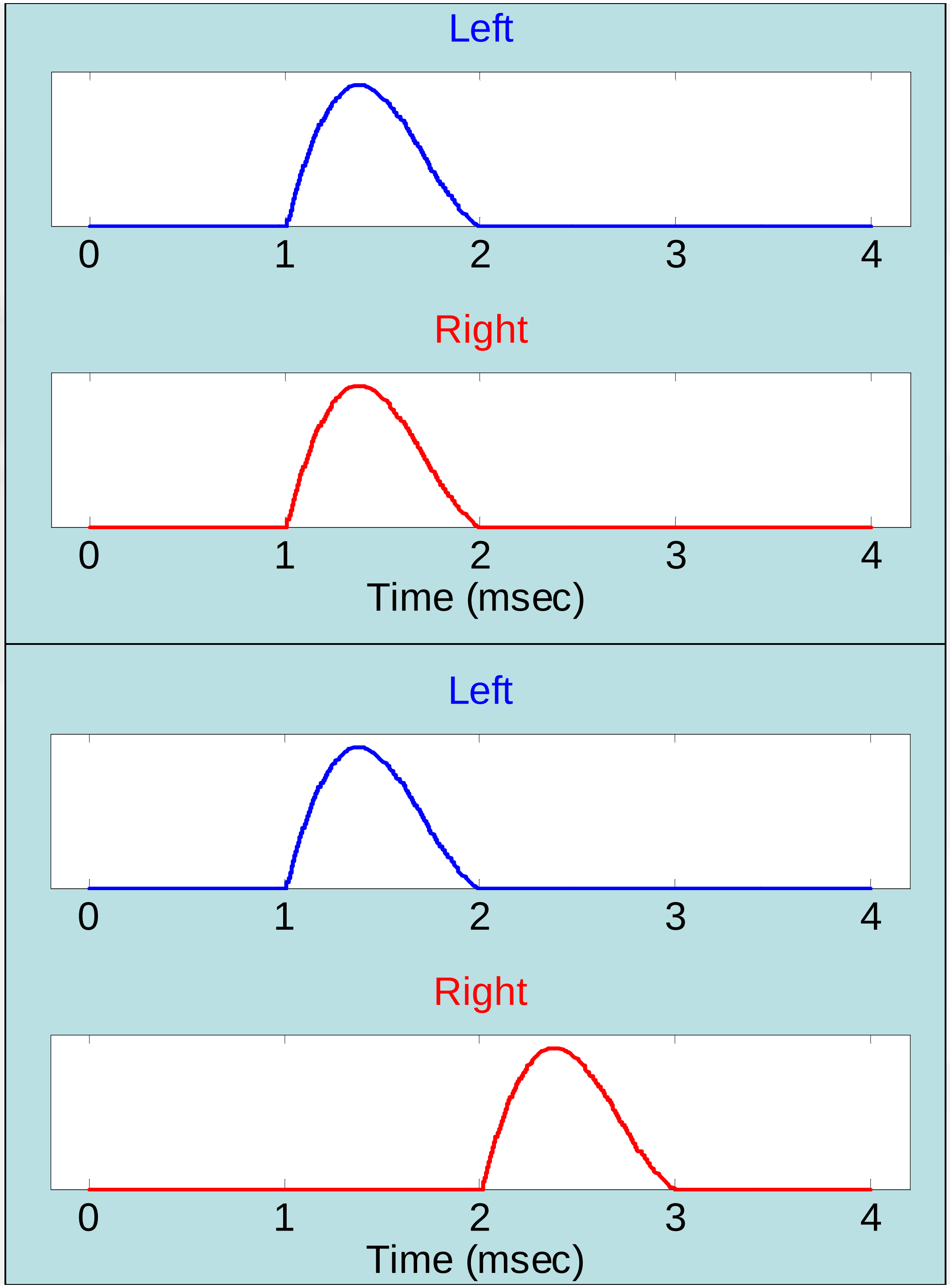
- Binaural cues - location cues based on the comparison of the signals received by the left and right ears

  Cue 1: Interaural *time* difference (ITD) - difference between the times sounds reach the two ears

    - When distance to each ear is the same, there are no differences in time
    - When the source is to the side of the observer, the times will differ



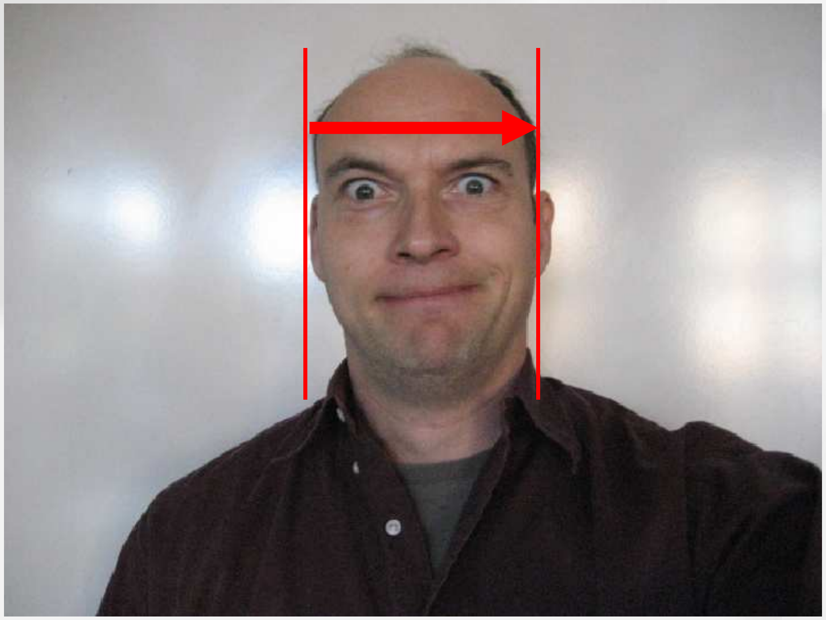
Interaural time difference (ITD)



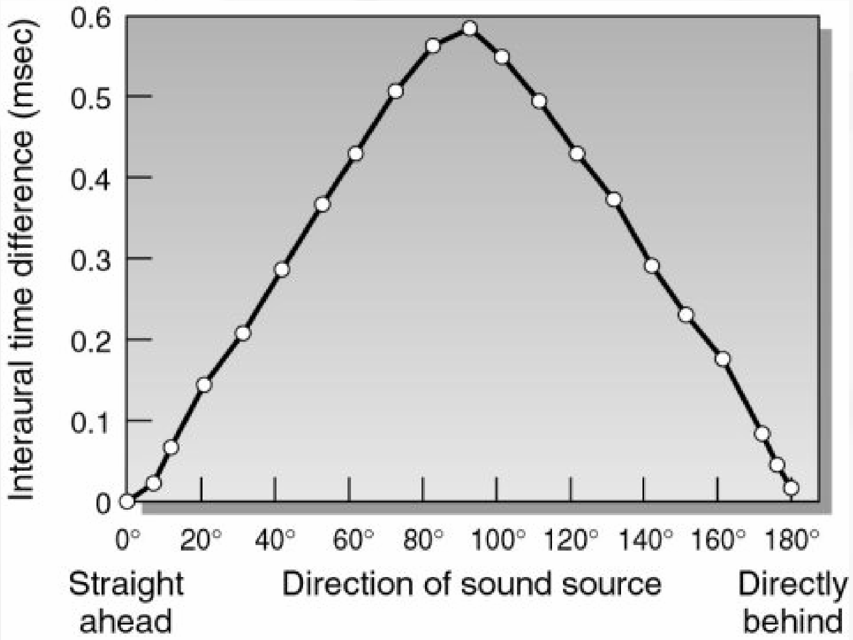
Interaural time difference (ITD)

Speed of sound at sea level: 761 mph = 6.22 inches/millisecond

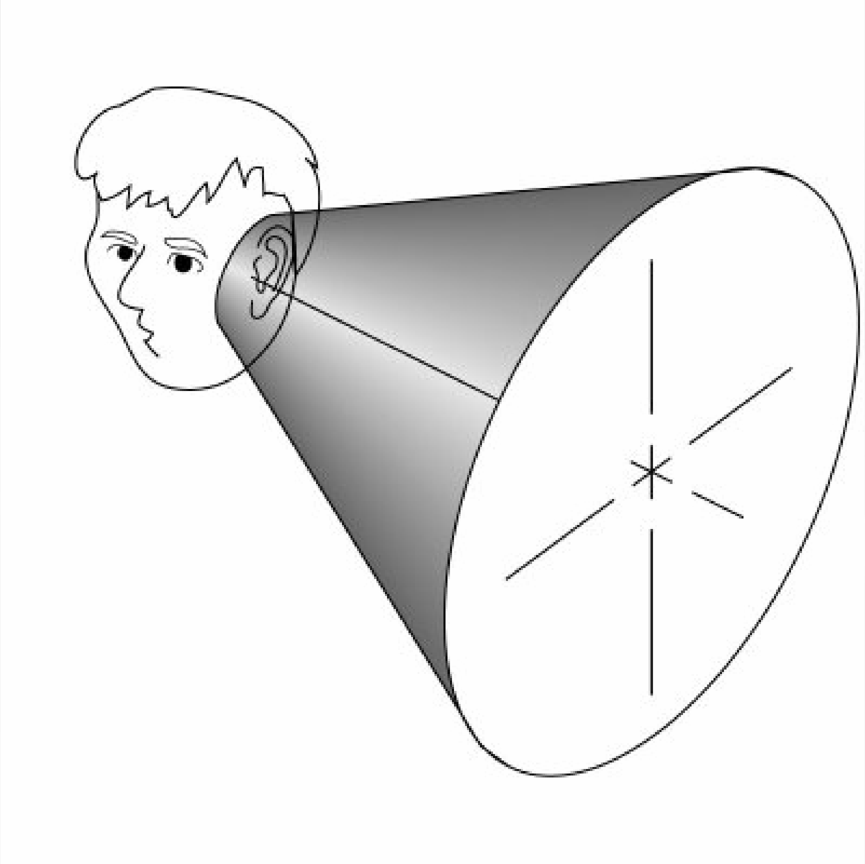
It should take about 0.6 msec for sound to travel the width of the average head.



ITD for different directions:



The ‘**Cone of Confusion**’: Set of locations that have the same interaural time differences (ITD)



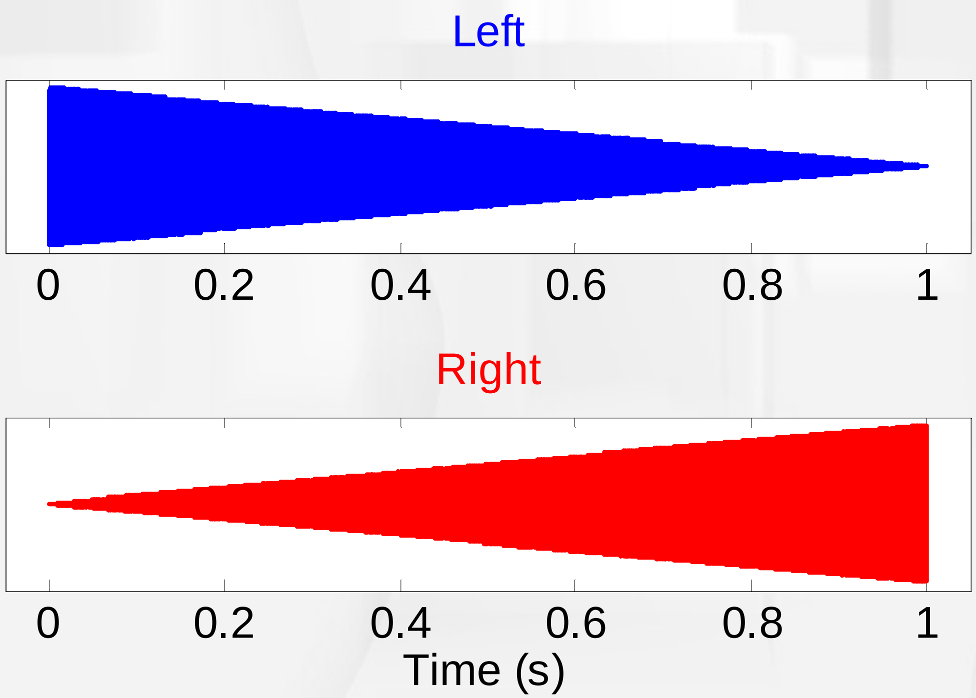
Cue 2: Interaural *level* difference ILD - difference in sound pressure level reaching the two ears

Reduction in sound level occurs for high frequency sounds for the far ear

The head casts an *acoustic shadow*

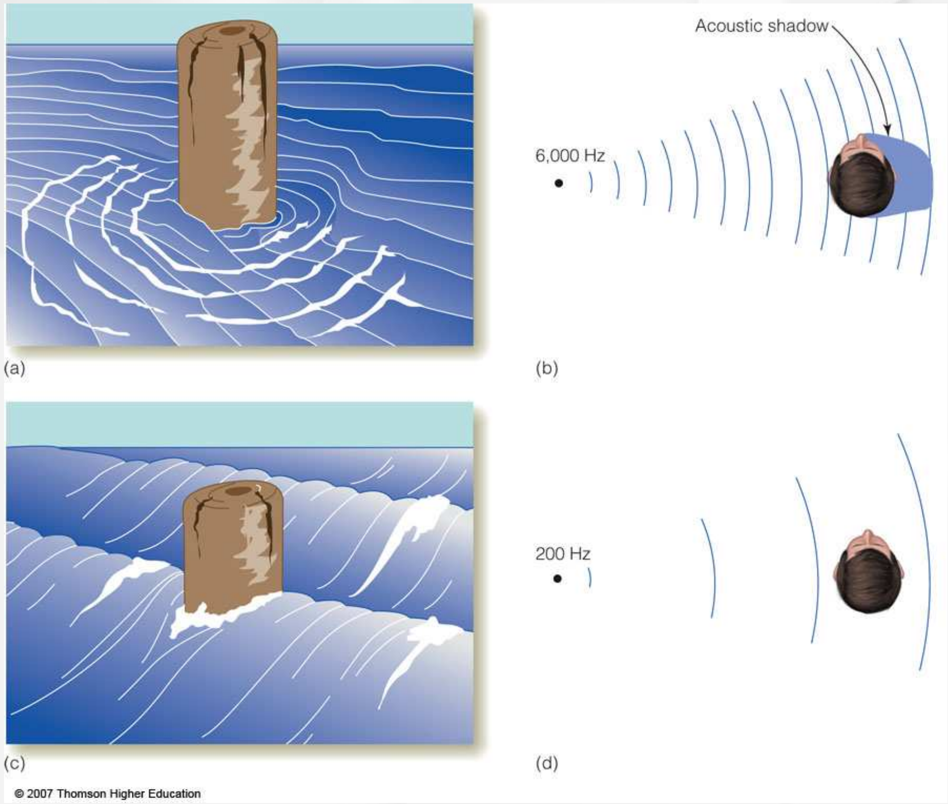
Demonstration of interaural level difference (ILD):

intensity sweep from left to right ear.

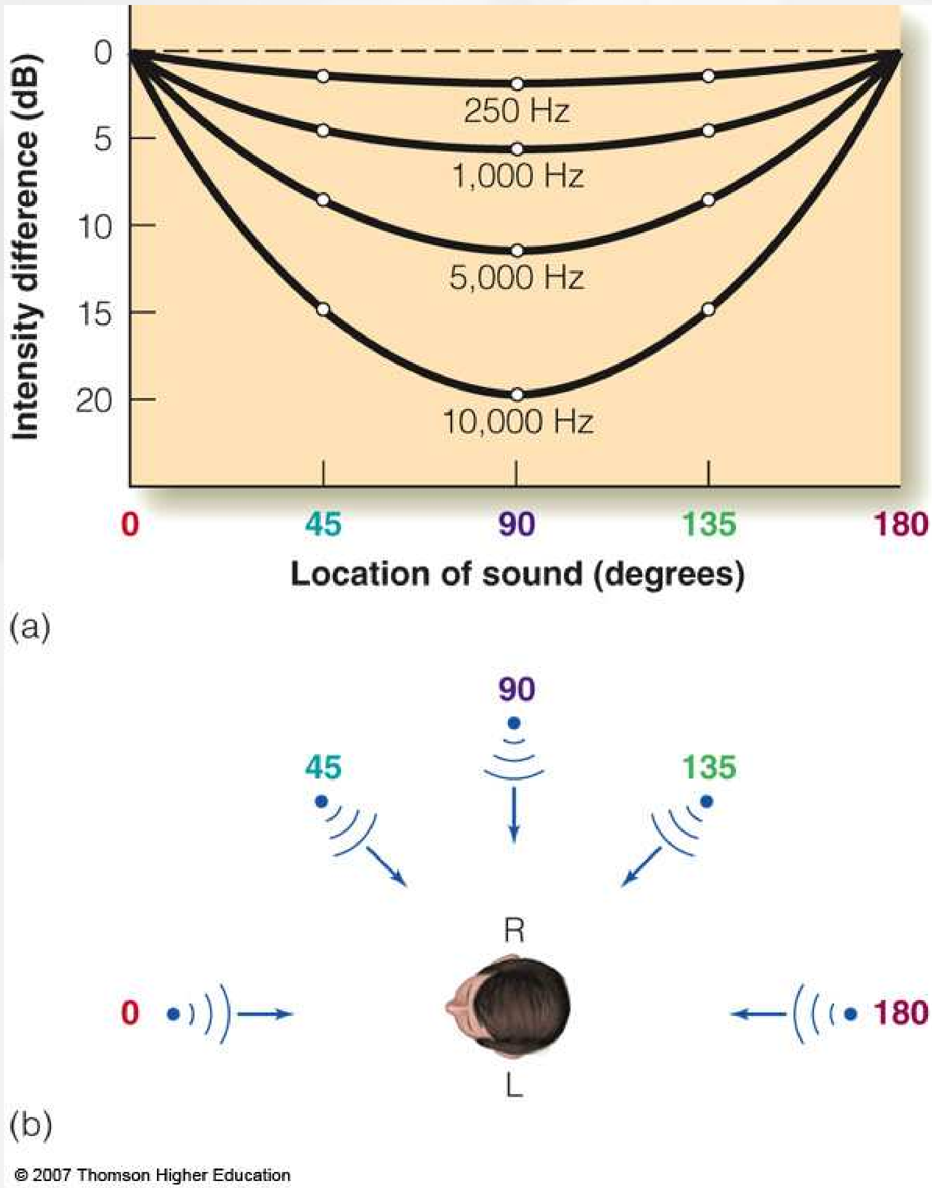


Interaural level difference (ILD) is best for high frequency sounds because low frequency sounds are not attenuated much by the head.

(think of how low frequency sounds pass through the wall from your neighbor next door)

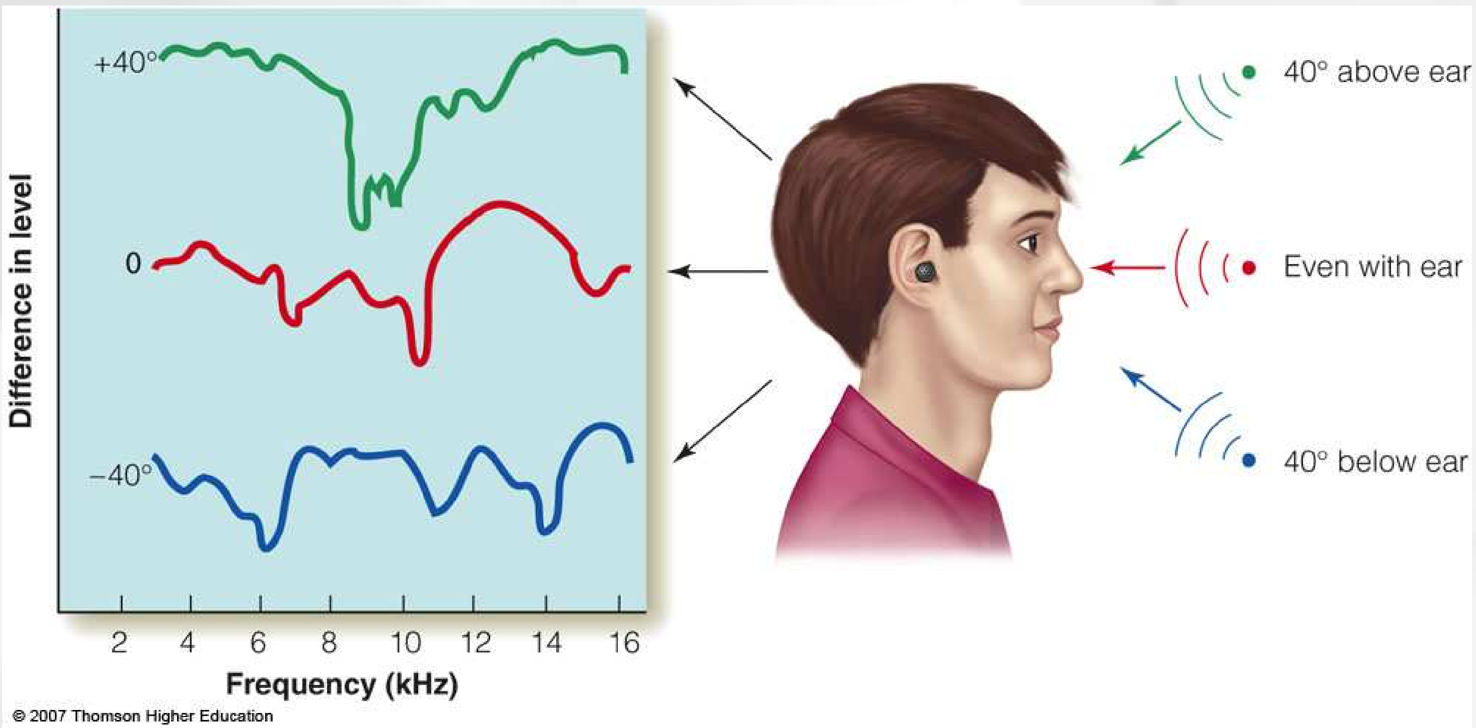


Interaural level difference (ILD) is best for high frequency sounds.



Cue 3: the head-related transfer function (HRTF)

- The pinna and head affect the intensities of frequencies
- Measurements have been performed by placing small microphones in ears and comparing the intensities of frequencies with those at the sound source
  - The difference is called the head-related transfer function (HRTF)
  - This is a spectral cue since the information for location comes from the spectrum of frequencies



Two ways to present sounds to subjects:

1) Free-field presentation - sounds are presented by speakers located around the listener's head in a dark room

Listener can indicate location by pointing or by giving azimuth and elevation coordinates

2) Headphone presentation of sounds

Advantage - experimenter has precise control over sounds

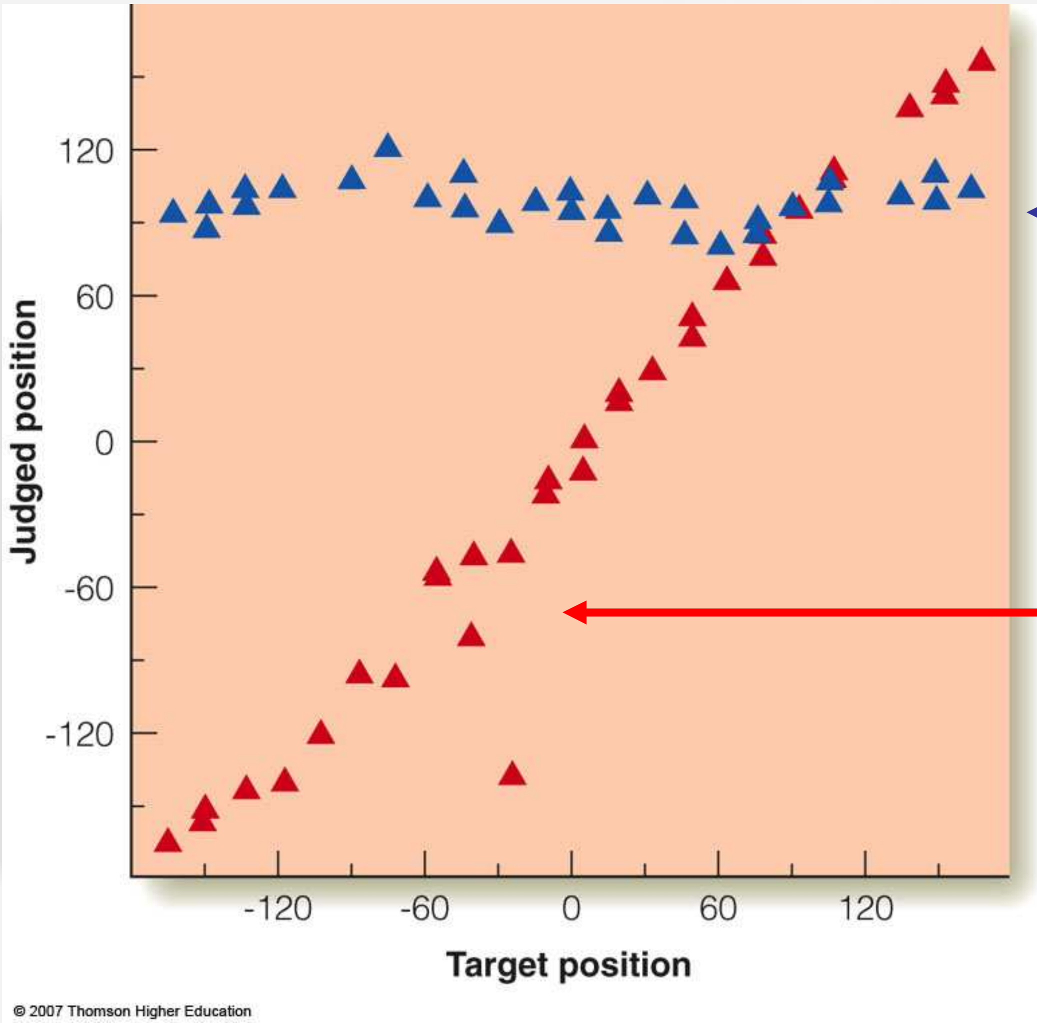
Disadvantage - cues from the pinna are eliminated, which results in the sound being *internalized*

- Sound can be *externalized* by using HTRFs to create a virtual auditory space

Which cues for sound localization do we actually use?

- Experiments by Wight and Kistler
  - Experiment 1 - used virtual auditory space
    - HRTFs, ITDs, & ILDs were used to indicate locations that varied from left to right
    - Listeners were fairly accurate

Low frequency tone:



ITD kept constant at 90 degrees:
Subjects don't use ILD cue and
ITD dominates judgment

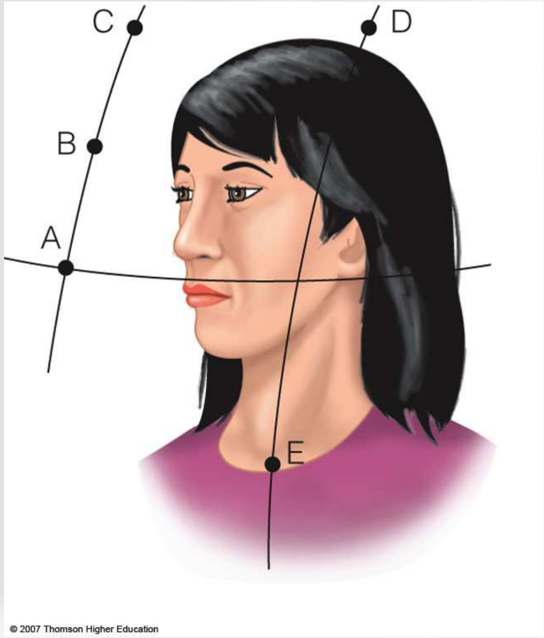
Both ITD and ILD cues:
Accurate judgment of azimuth

For low frequencies, ITD dominates judgment

For high frequencies, ILD dominates judgment

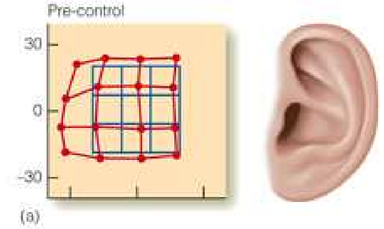
Judging Elevation

- ILD and ITD are not effective for judgments on elevation since in many locations they may be zero



You can turn elevation to azimuth by tilting your head.

Experiment investigating spectral cues:



Listeners were measured for performance locating sounds differing in elevation

They were then fitted with a mold that changed the shape of their pinnae: Right after the molds were inserted, performance was poor

After 19 days, performance was close to original performance

Once the molds were removed, performance stayed high.

This suggests that there might be two different sets of neurons—one for each set of spectral cues

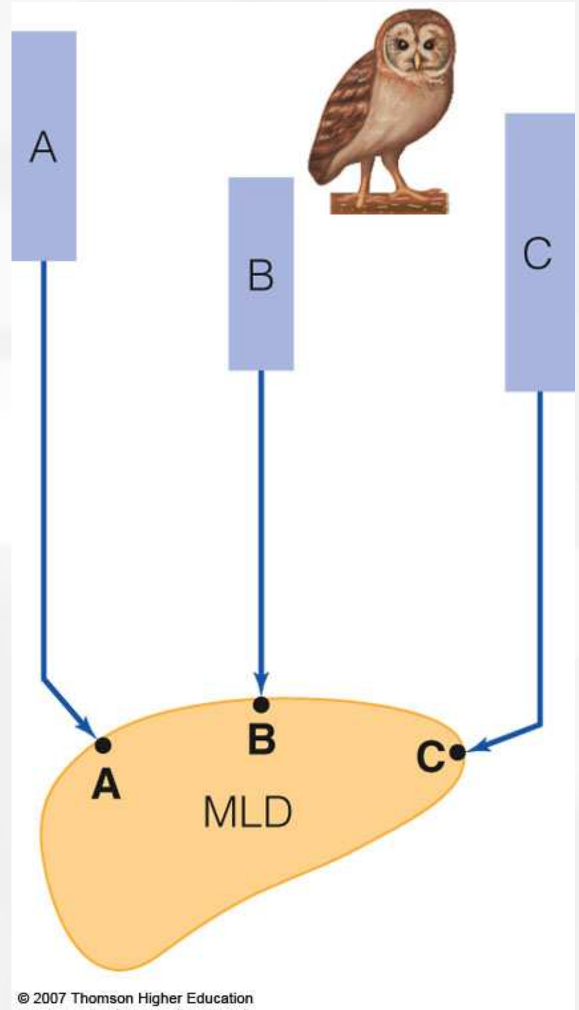
The Physiological Representation of Auditory Space

- Interaural time-difference (ITD) detectors - neurons that respond to specific interaural time differences
  - They are found in the auditory cortex and at the first nucleus (superior olivary) in the system that receives input from both ears

- Topographic maps - neural structure that responds to locations in space

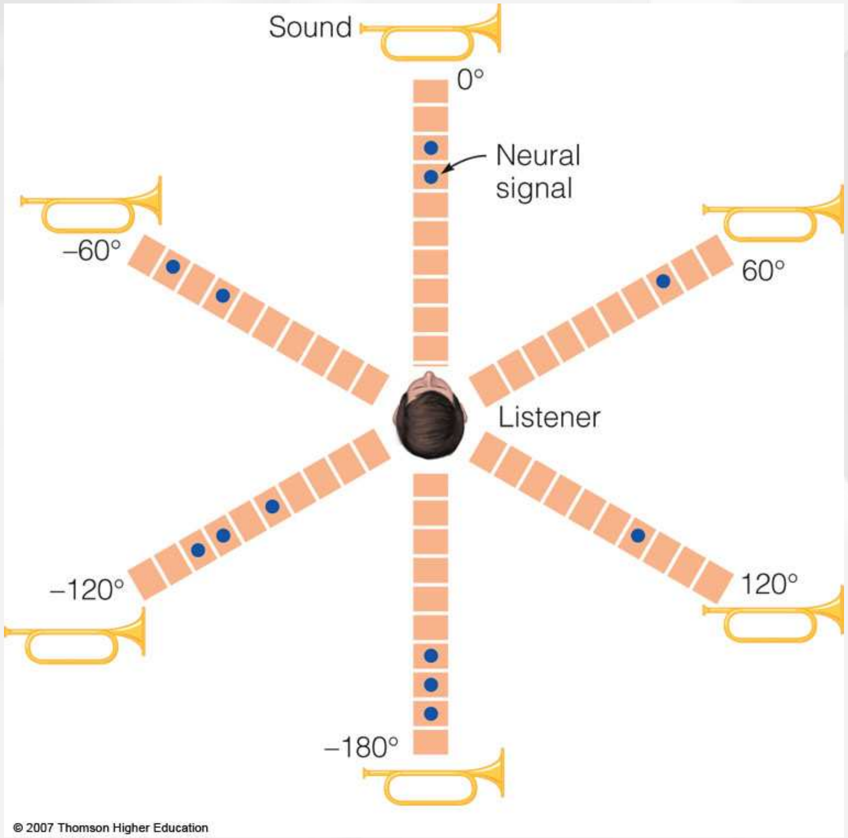
Topographic Maps

- Barn owls have neurons in the mesencephalicus lateralus dorsalis (MLD) that respond to locations in space
- Mammals have similar maps in the subcortical structures, such as the inferior colliculus
- These neurons have receptive fields for *sound location*

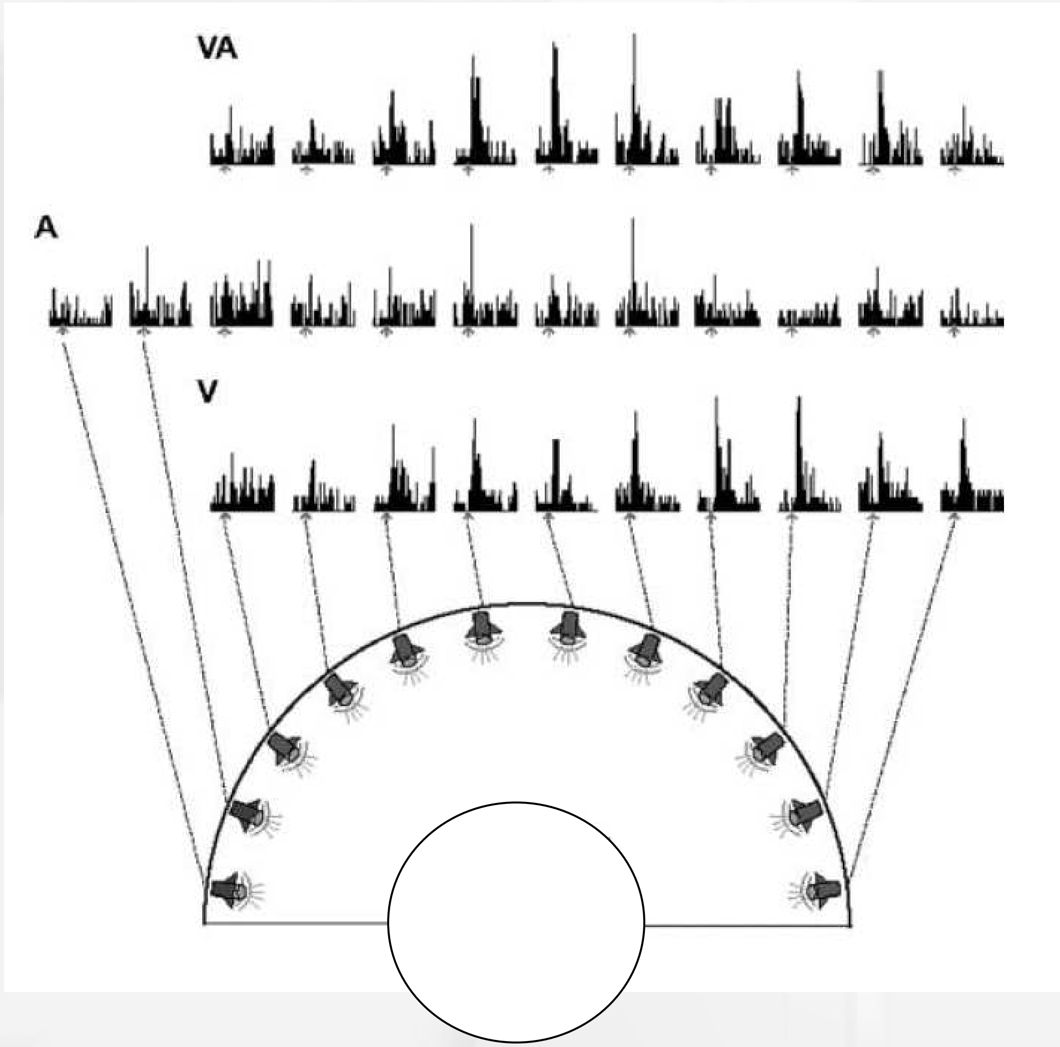


The Auditory Cortex

- Even though there are topographic maps in subcortical areas of mammals, there is no evidence of such maps in the cortex (to date).
- Instead, ***panoramic neurons*** have been found that signal location by their *pattern* of firing

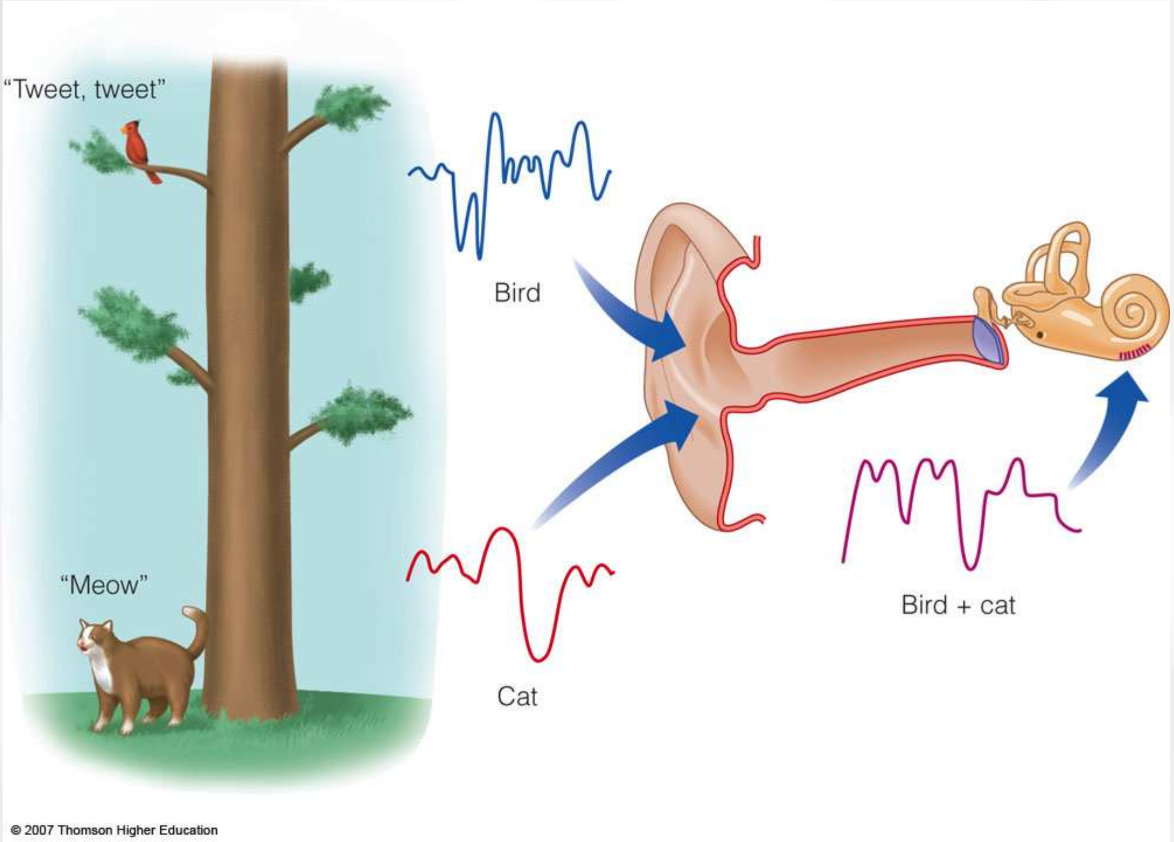


Evidence for ‘multimodal’ neurons coding spatial position in the association cortex of the cat.



The Auditory Scene; the ‘what’ pathway

- Auditory Scene - the array of all sound sources in the environment
- Auditory Scene Analysis - process by which sound sources in the auditory scene are separated into individual perceptions
- This does not happen at the cochlea since simultaneous sounds will be together in the pattern of vibration of the basilar membrane



Principles of Auditory Grouping

Auditory stimuli tend to group together by **similarity**. This includes:

1. Location - a single sound source tends to come from one location and to move continuously

2. Proximity in time - sounds that occur in rapid succession usually come from the same source

- This principle was illustrated in auditory streaming

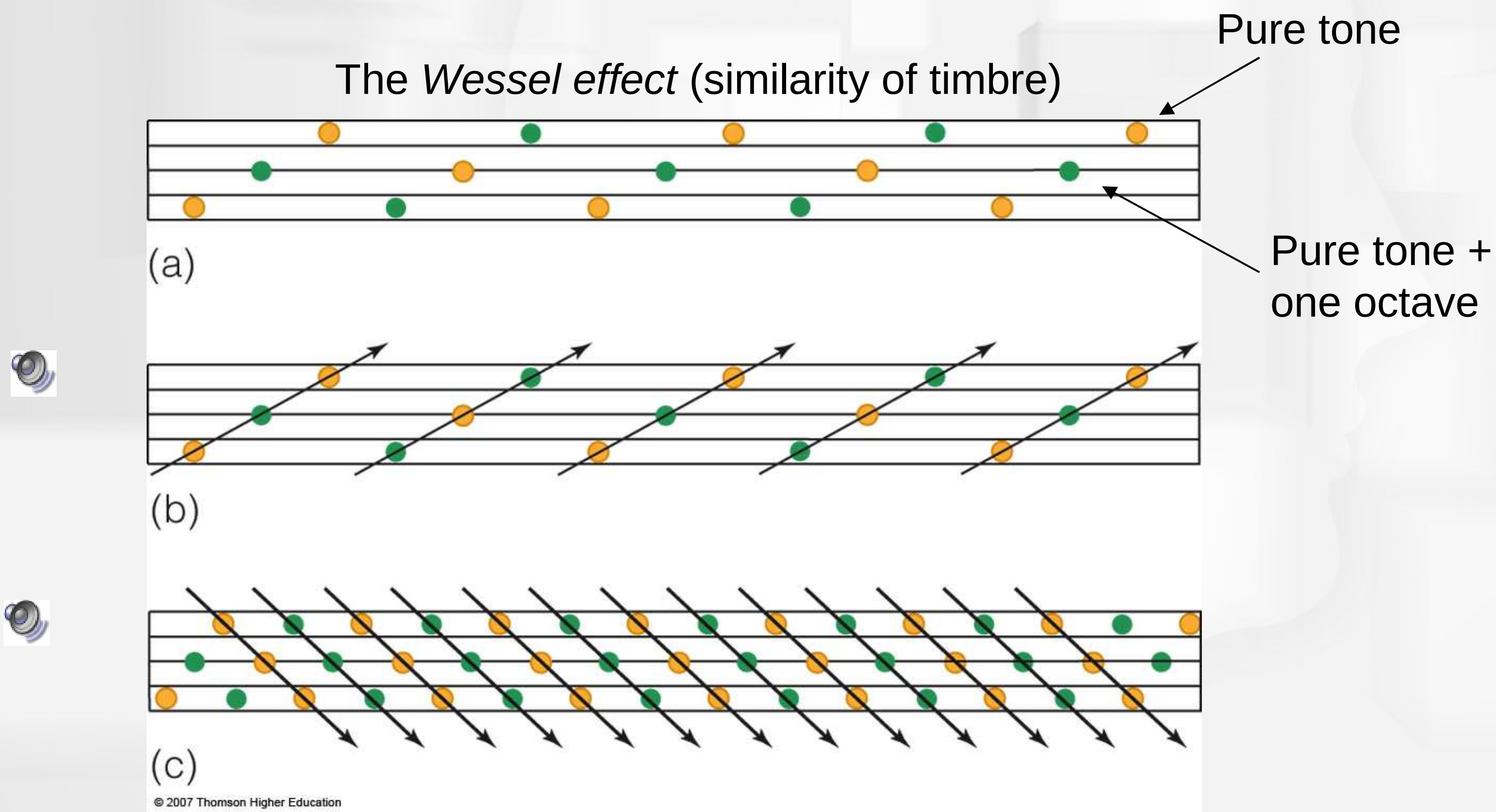
3. Good continuation - sounds that stay constant or change smoothly are usually from the same source

Principles of Auditory Grouping

4. Similarity of timbre and pitch - similar sounds are grouped together

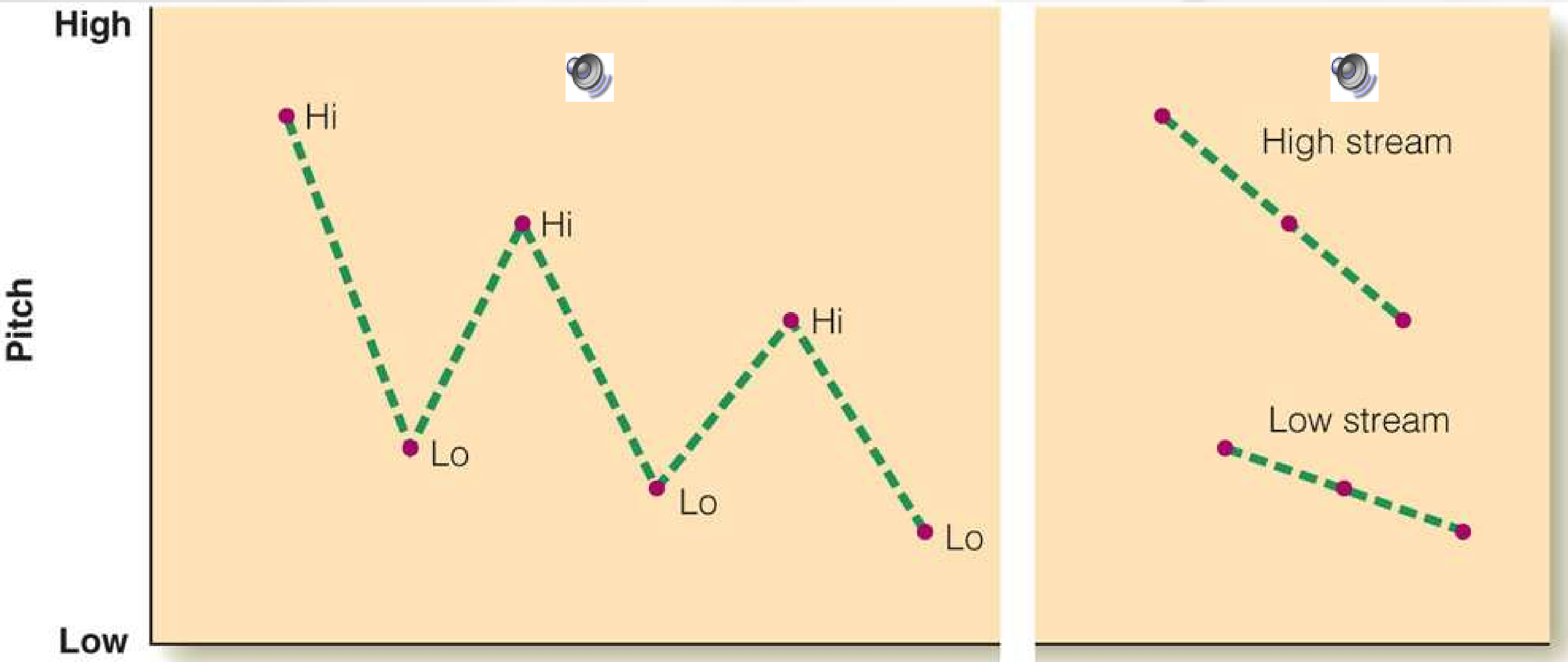
Sounds with similar frequencies sound like they come from the same source, which is usually true in the environment

The *Wessel effect* (similarity of timbre)



Similarity of timbre and pitch

- Experiment by Bregman and Campbell (similarity of pitch vs. proximity in time)
  - Stimuli were alternating high and low tones
  - When stimuli played slowly, the perception is hearing high and low tones alternating
  - When the stimuli are played quickly, the listener hears two streams; one high and one low

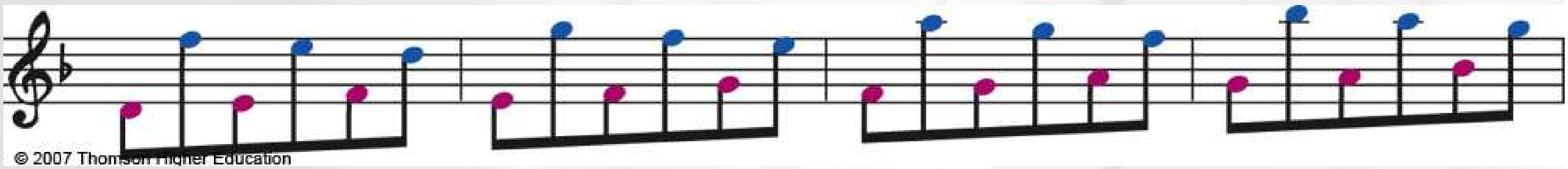


(a) Tones alternated slowly
Perception: Hi-Lo-Hi-Lo-Hi-Lo

(b) Tones alternated rapidly
Perception: Two separate streams

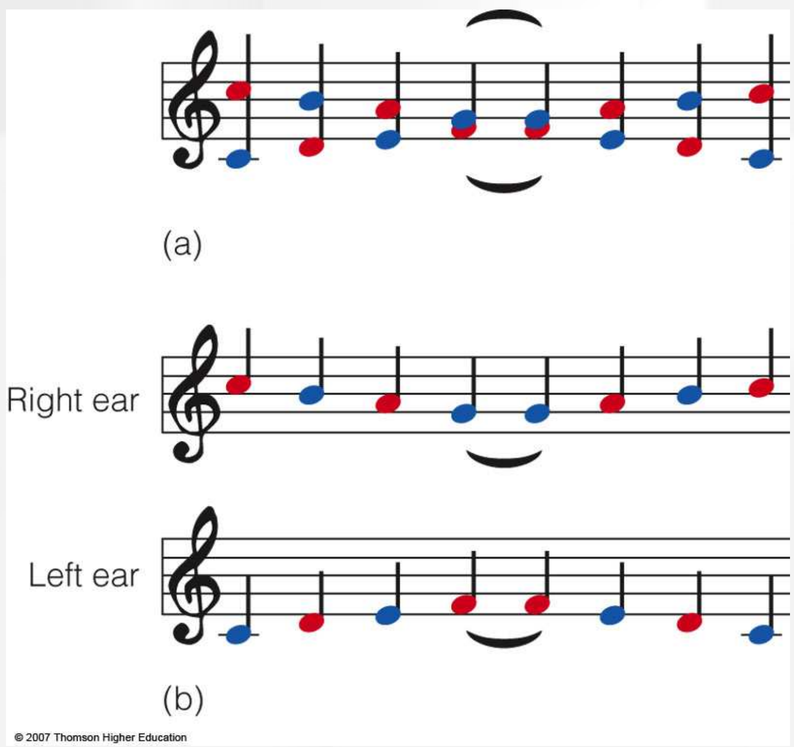
Similarity of timbre and pitch

Four measures of a composition by J. S. Bach (Chorale Prelude on *Jesus Christus unser Heiland,* 1739).



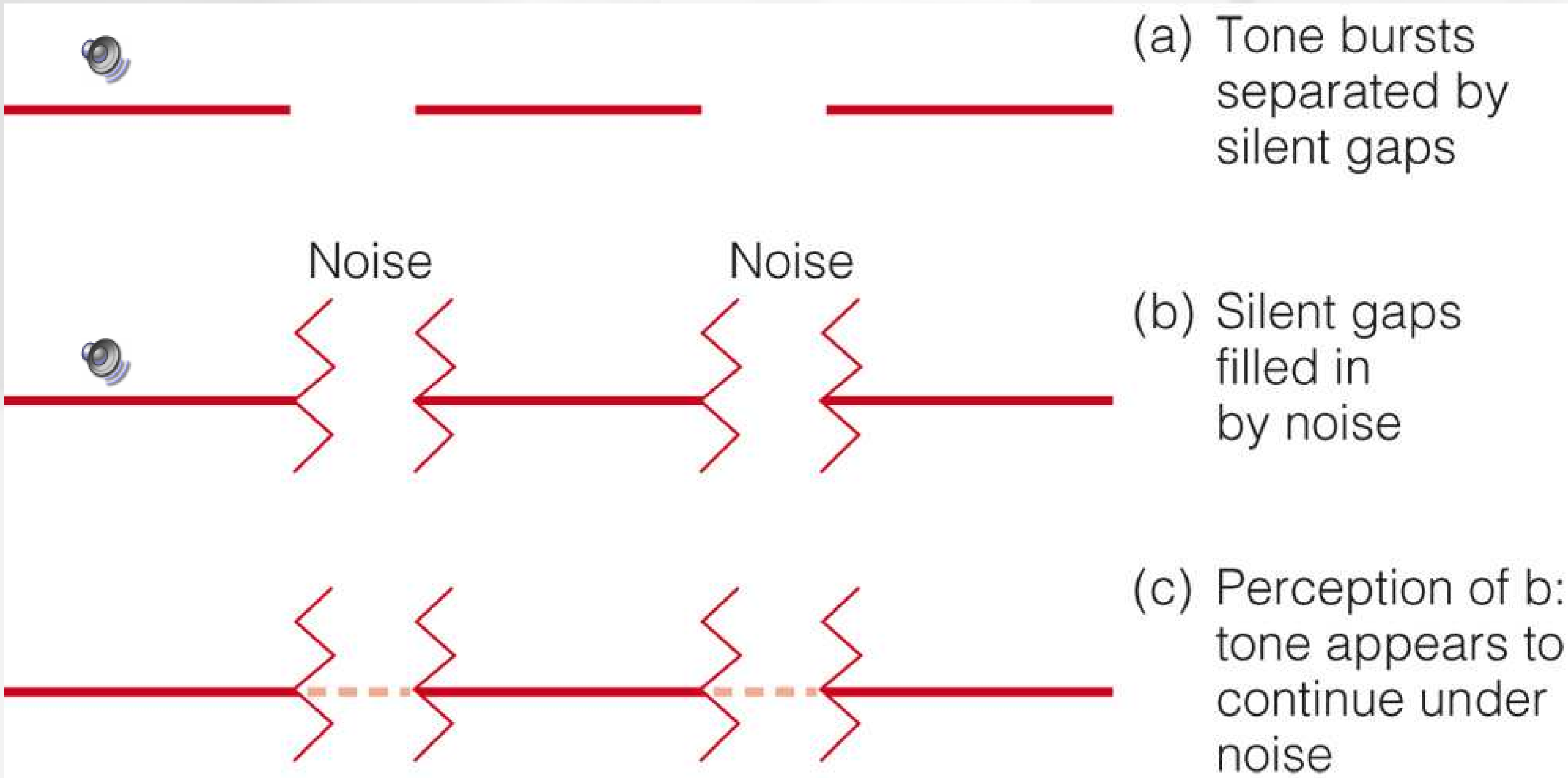
Auditory Stream Segregation - continued

- Experiment by Deutsch - the scale illusion or melodic channeling
  - Stimuli were two sequences alternating between the right and left ears
  - Listeners perceive two smooth sequences by grouping the sounds by **similarity in pitch**



Good Continuation

- Experiment by Warren et al.
  - Tones were presented interrupted by gaps of silence or by noise
  - In the silence condition, listeners perceived that the sound stopped during the gaps
  - In the noise condition, the perception was that the sound continued behind the noise

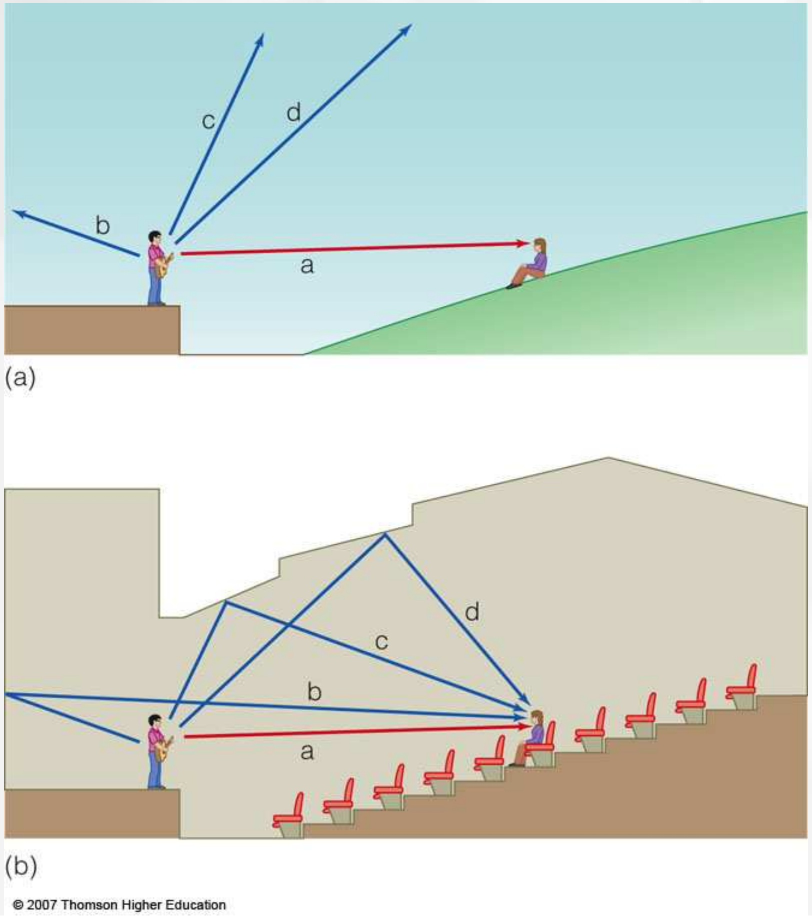


Principles of Auditory Grouping - continued

- Effect of past experience
  - Experiment by Dowling
    - Used two interleaved melodies (“Three Blind Mice” and “Mary Had a Little Lamb”)
    - Listeners reported hearing a meaningless jumble of notes
    - But listeners who were told to listen for the melodies were able to hear them by using *melody schema*

Hearing Inside Rooms

- Direct sound - sound that reaches the listeners' ears straight from the source
- Indirect sound - sound that is reflected off of environmental surfaces and then to the listener
- When a listener is outside, most sound is direct; however inside a building, there is direct and indirect sound



Experiment by Litovsky et al.

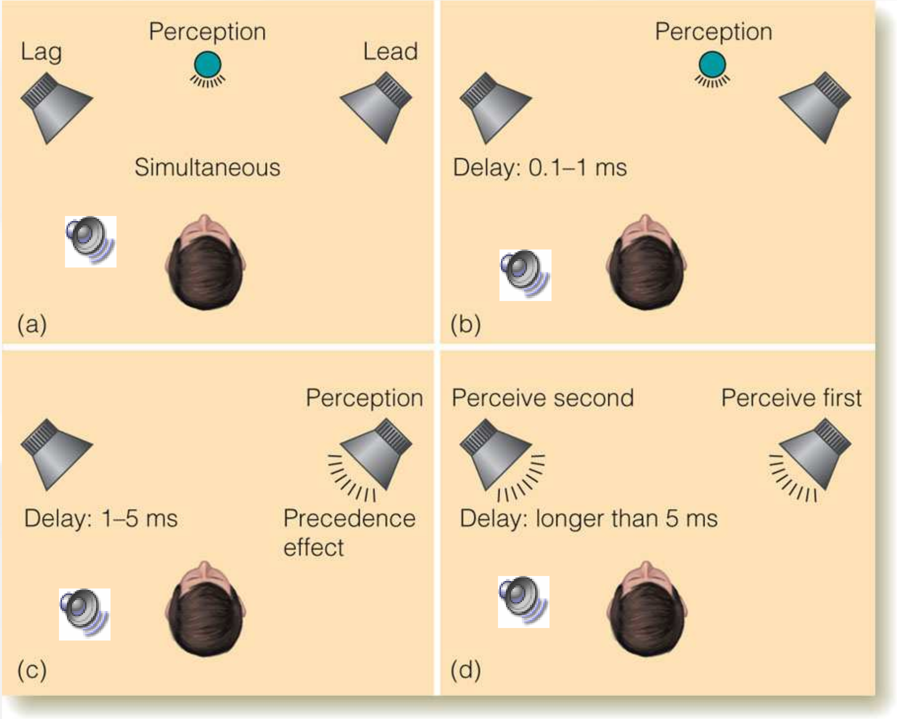
- Listeners sat between two speakers
  - Right speaker was the *lead speaker*
  - Left speaker was the *lag speaker*

(a) When two sounds were presented simultaneously, listeners heard a centered sound between speakers, the two sounds became *fused*

(b) Less than 1 ms before the lag speaker, a single sound nearer the lead speaker was heard

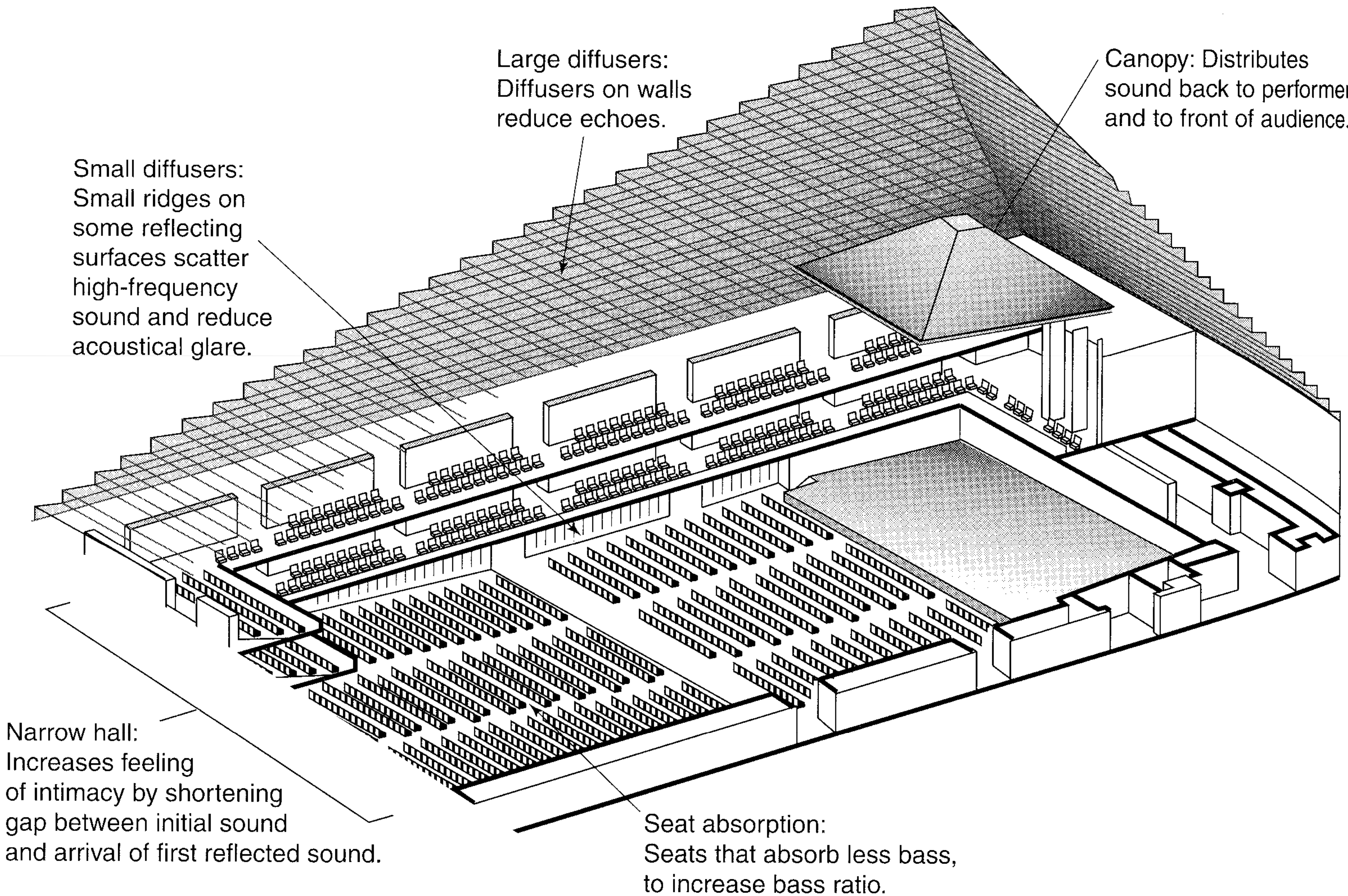
(c) From 1 to 5 ms before the lag speaker, sound appeared to come from lead speaker alone - called the *precedence effect*

(d) At intervals greater than 5 ms, two separate sounds were heard, one following the other - called the *echo threshold*



Architectural Acoustics

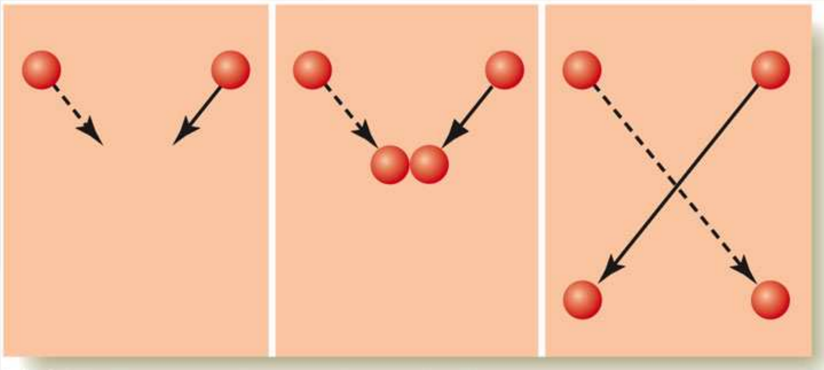
- The study of how sounds are reflected in rooms
- Factors that affect perception in concert halls
  - ***Reverberation time*** - the time is takes sound to decrease by 1/1000th of its original pressure
    - Best time is around 2 sec (1.5 for opera)
  - ***Intimacy time*** - time between when sound leaves its source and when the first reflection arrives
    - Best time is around 20 ms
  - ***Bass ratio -*** ratio of low to middle frequencies reflected from surfaces
    - High bass ratios are best
  - ***Spaciousness factor*** - fraction of all the sound received by listener that is indirect
    - High spaciousness factors are best


Large diffusers:
Diffusers on walls
reduce echoes.
Canopy: Distributes
sound back to performers
and to front of audience.
Small diffusers:
Small ridges on
some reflecting
surfaces scatter
high-frequency
sound and reduce
acoustical glare.
Narrow hall:
Increases feeling
of intimacy by shortening
gap between initial sound
and arrival of first reflected sound.
Seat absorption:
Seats that absorb less bass,
to increase bass ratio.

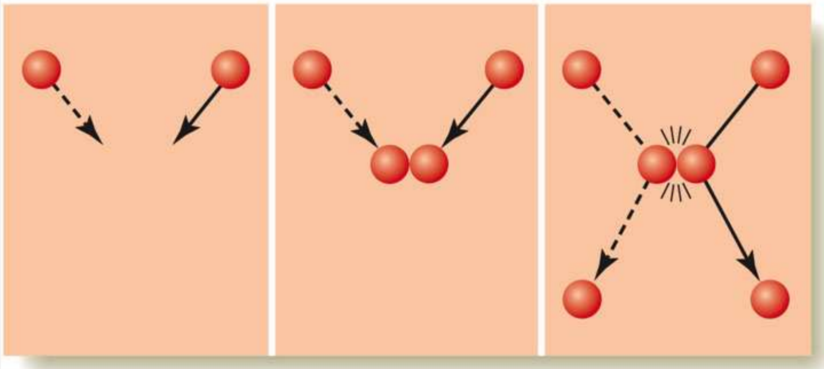
Interactions between sight and sound

Experiment by Sekuler et al.

- Balls moving without sound appeared to move past each other
- Balls with an added “click” appeared to collide



(a) Objects appear to pass by each other



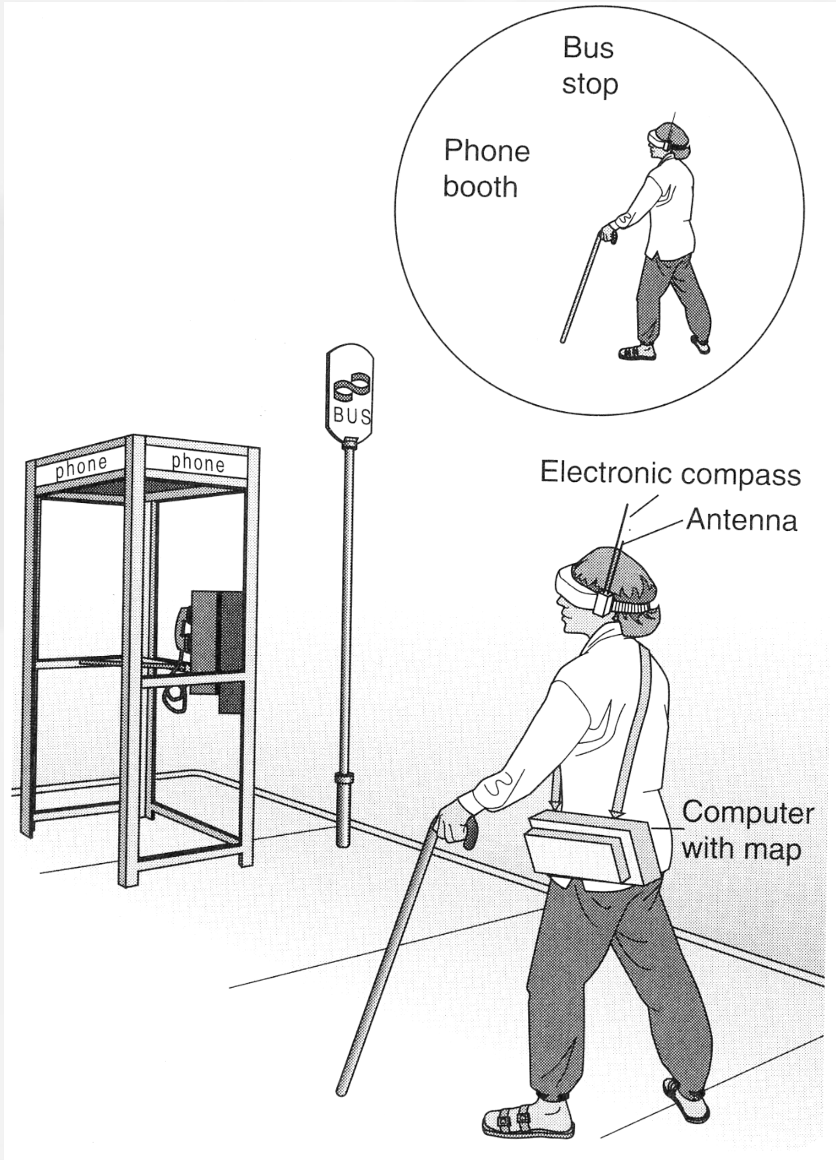
(b) Objects appear to collide

Sound-induced Illusory Flashing

Auditory clicks can influence perceived number of visual flashes.

http://shamslab.psych.ucla.edu/demos/

Using auditory stimuli to replace sight



http://www.senderogroup.com/