KTH ROYAL INSTITUTE
OF TECHNOLOGY

DT2350 Human Perception for Information Technology

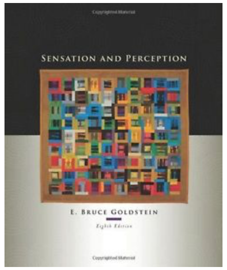
Sound, the Auditory System, and Pitch Perception

Roberto Bresin

Outlook

- The nature of sound
- How we experience sounds
- Physiology behind the perception of pitch
- Hearing loss
- How different parts of the brain respond to sound
- Sound localization
- Sound design: getting attention with sound

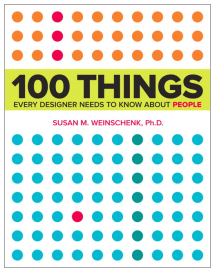
Literature



Goldstein, E. (2009/2014). Sensation and Perception.

Chapter 11 (2009 edition): *Sound, the Auditory System, and Pitch Perception*

Chapter 11 (2014 edition): *Hearing*



Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People.

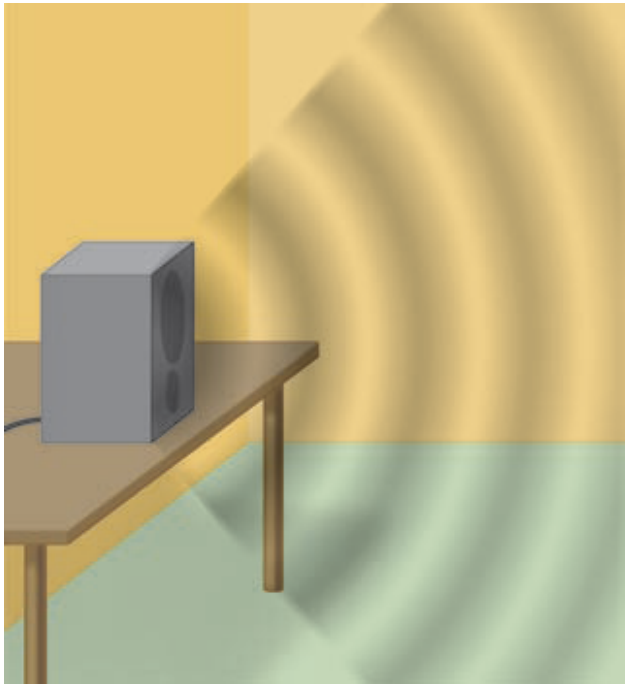
Chapters #48*: Loud noises startle and get attention*

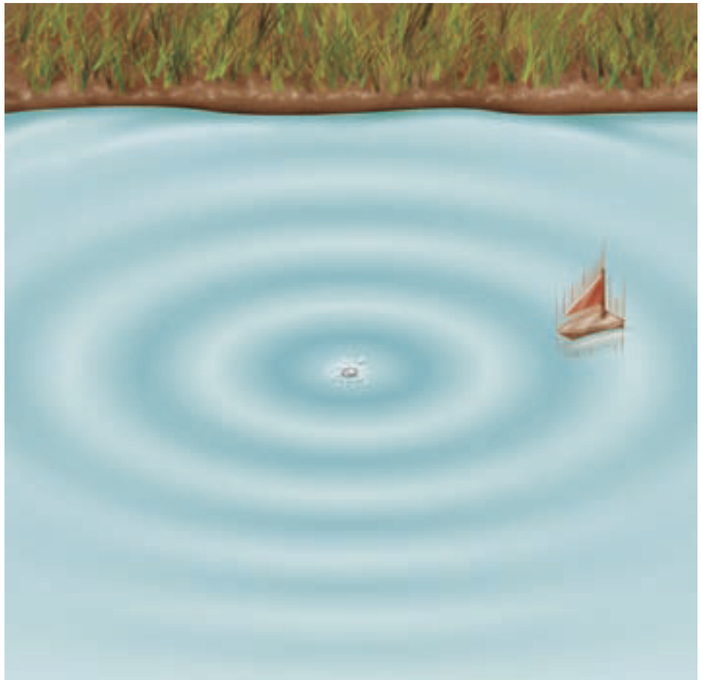
Test

Which sense would you choose to keep if you had to pick between hearing and vision?

Why? (discuss between peers in 3 minutes)

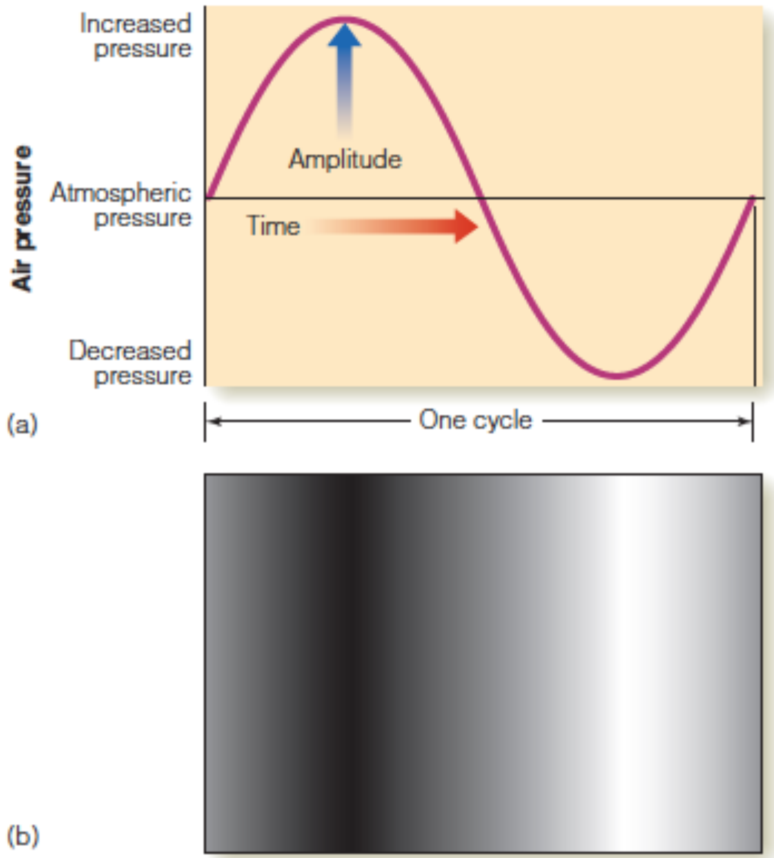
What is sound?
Physical vs perceptual definition





Sound wave speed: 340 m/s (air), 1500 m/s (water), 5120 m/s (iron)

Air pressure changes



Frequency [Hz]
Number of times/second
that the pressure changes repeat.

Humans can perceive frequencies in the range from 20 Hz to 20,000 Hz.

Decibel [dB]
dB = 20 x log (p/p_o)

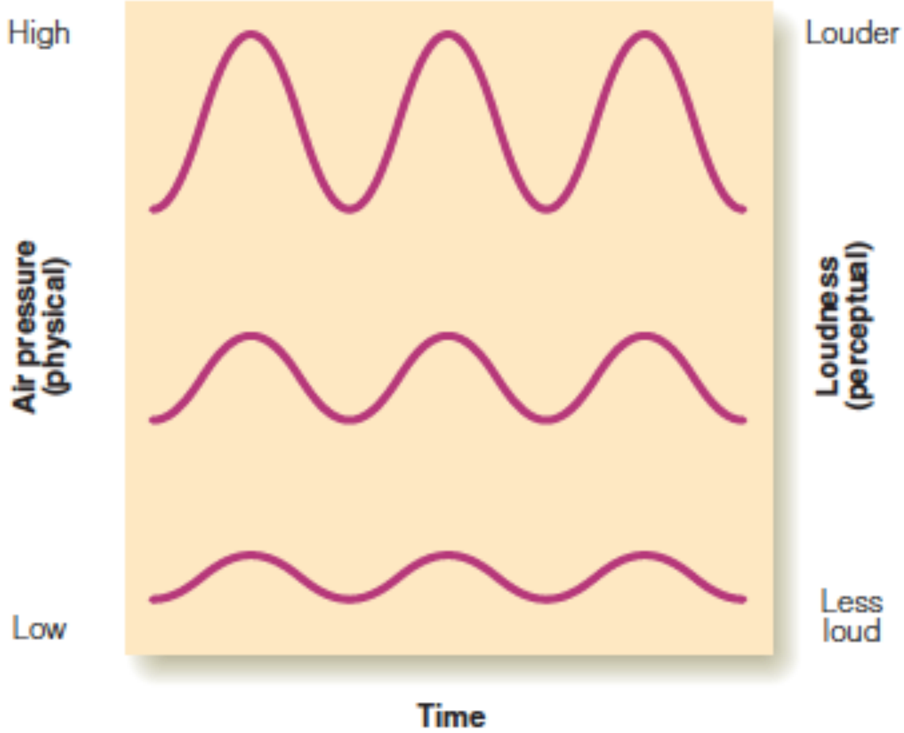
p = stimulus sound pressure
p_o = standard sound pressure
(20 micropascals = hearing threshold)

Sound level [dB]

Higher amplitude (physical) ➔ Higher sound level (perceptual)

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



Frequency [Hz]

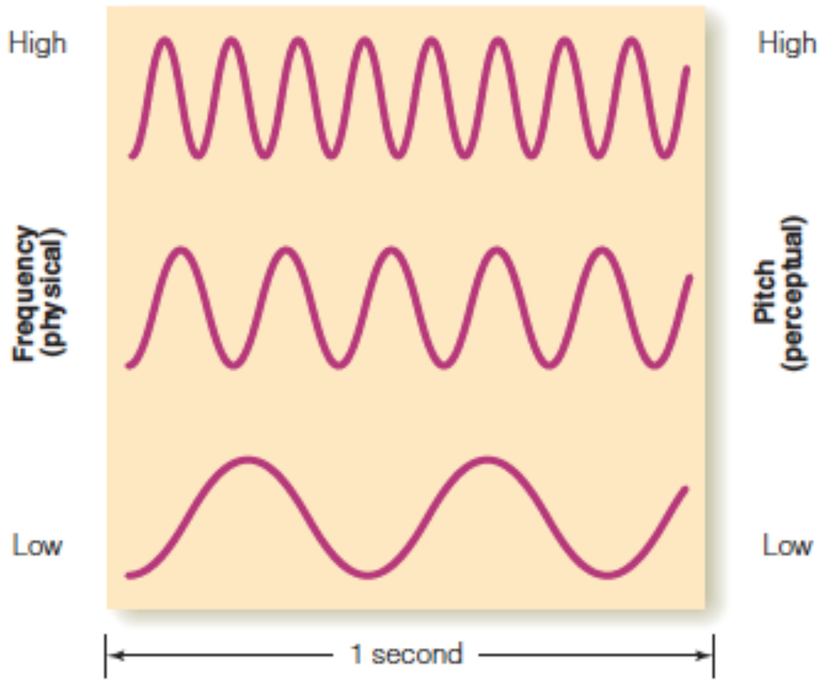
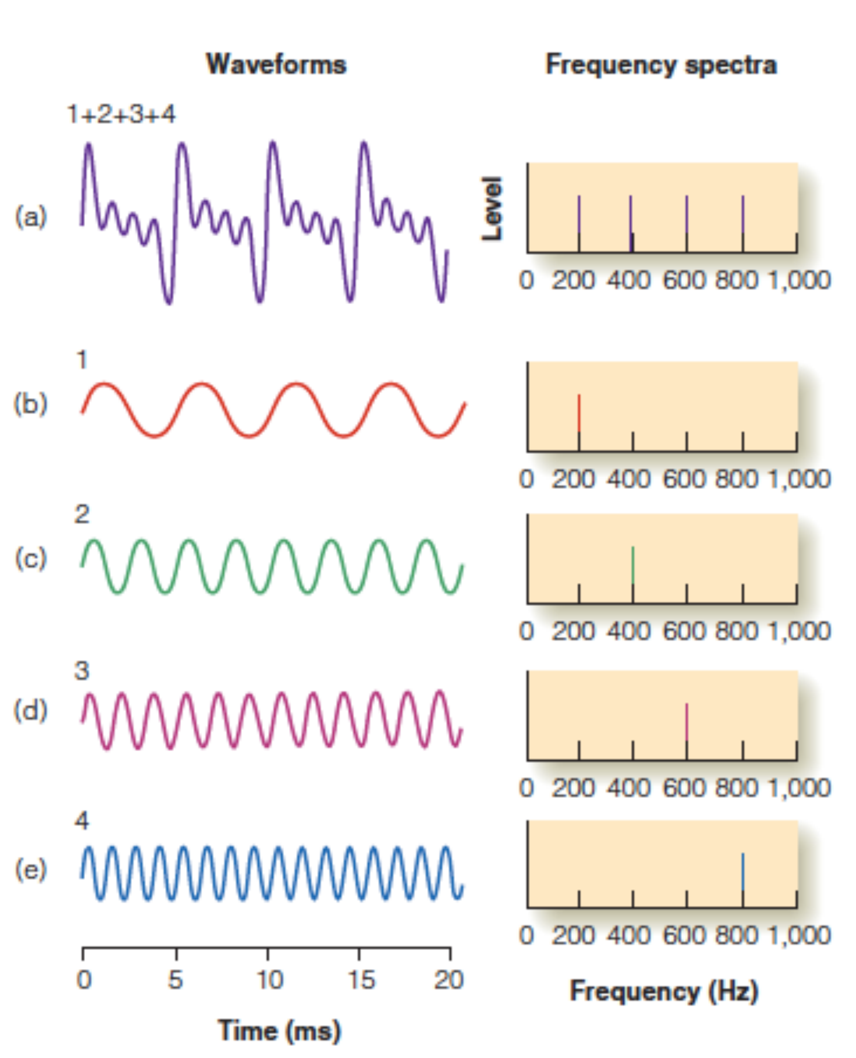
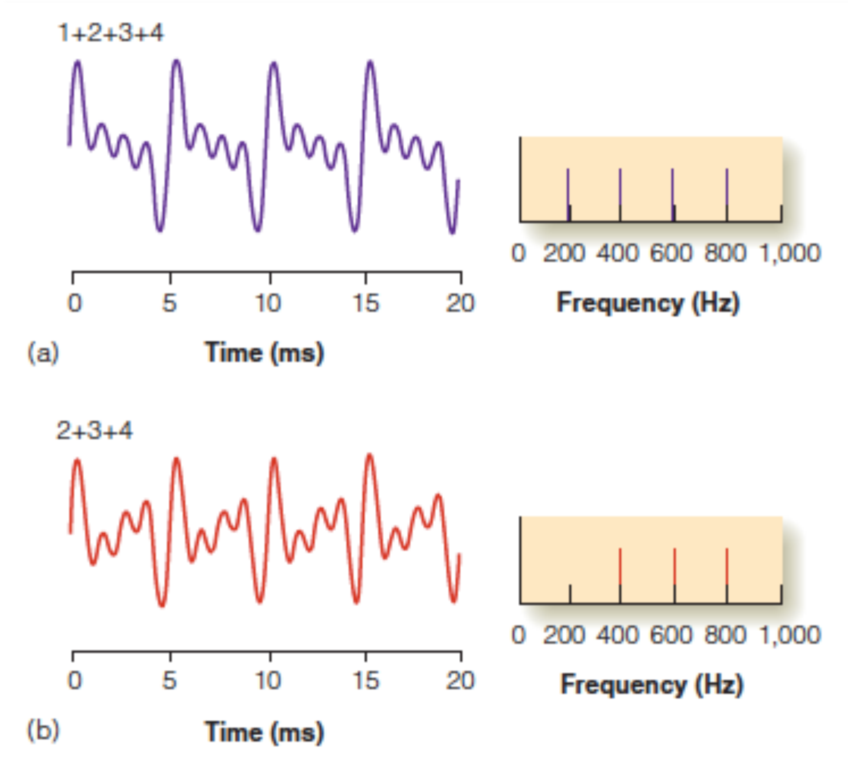


Figure 11.4 Three different frequencies of a pure tone. Higher frequencies are associated with the perception of higher pitches.

Complex wave forms / Frequency Spectra





Loudness

Loudness was judged relative to a standard of a 1000 Hz tone at 40 dB, which was assigned a value of 1.

Thus, a tone that sounds 10 times louder than this standard would be judged to have a loudness of 10.

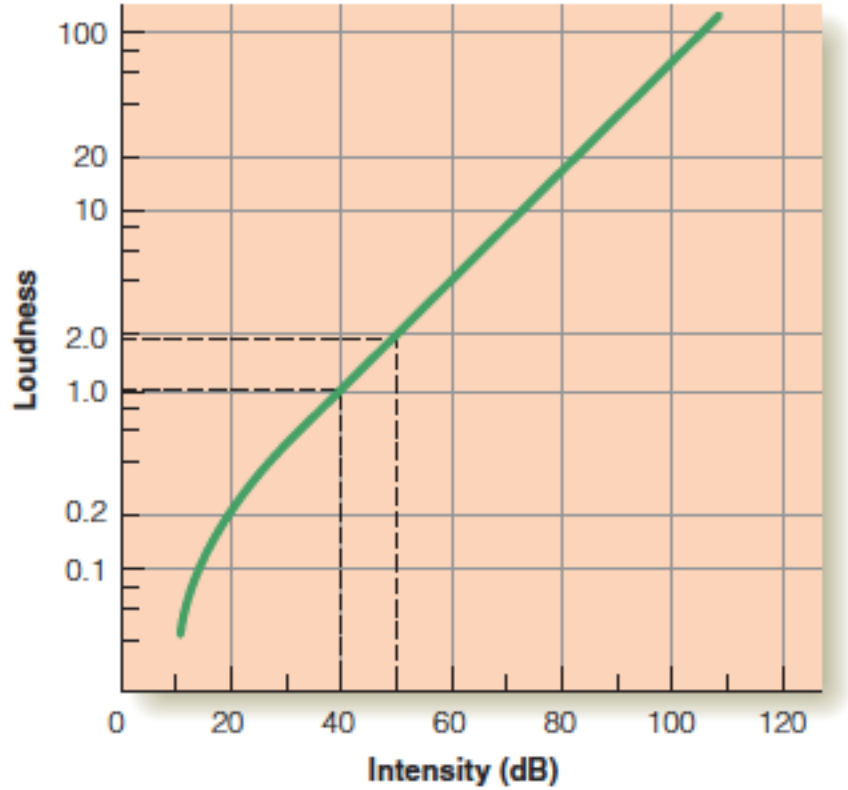
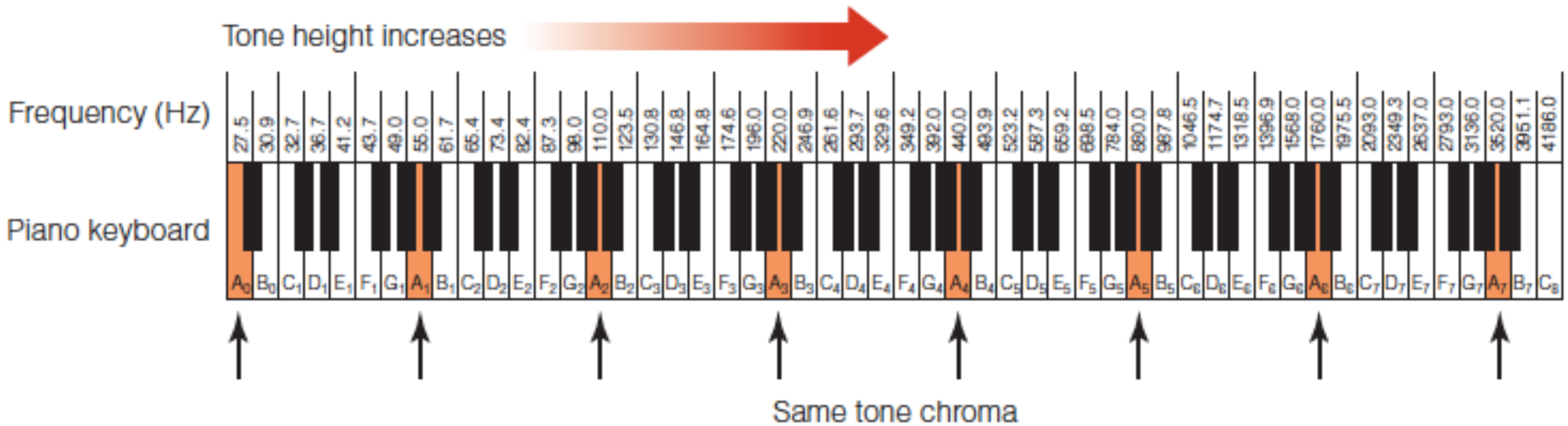


Figure 11.7 Loudness of a 1,000-Hz tone as a function of intensity, determined using magnitude estimation. The dashed lines show that increasing the intensity by 10 dB almost doubles the loudness. *(Adapted from Gulick, Gescheider, & Frisina, 1989.)*

Pitch

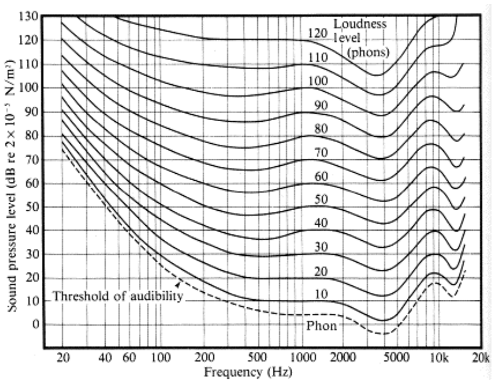
Pitch, the perceptual quality we describe as “high” or “low” is defined as *the attribute of auditory sensation in terms of which sounds may be ordered on a musical scale.*
(Bendor & Wang, 2005)

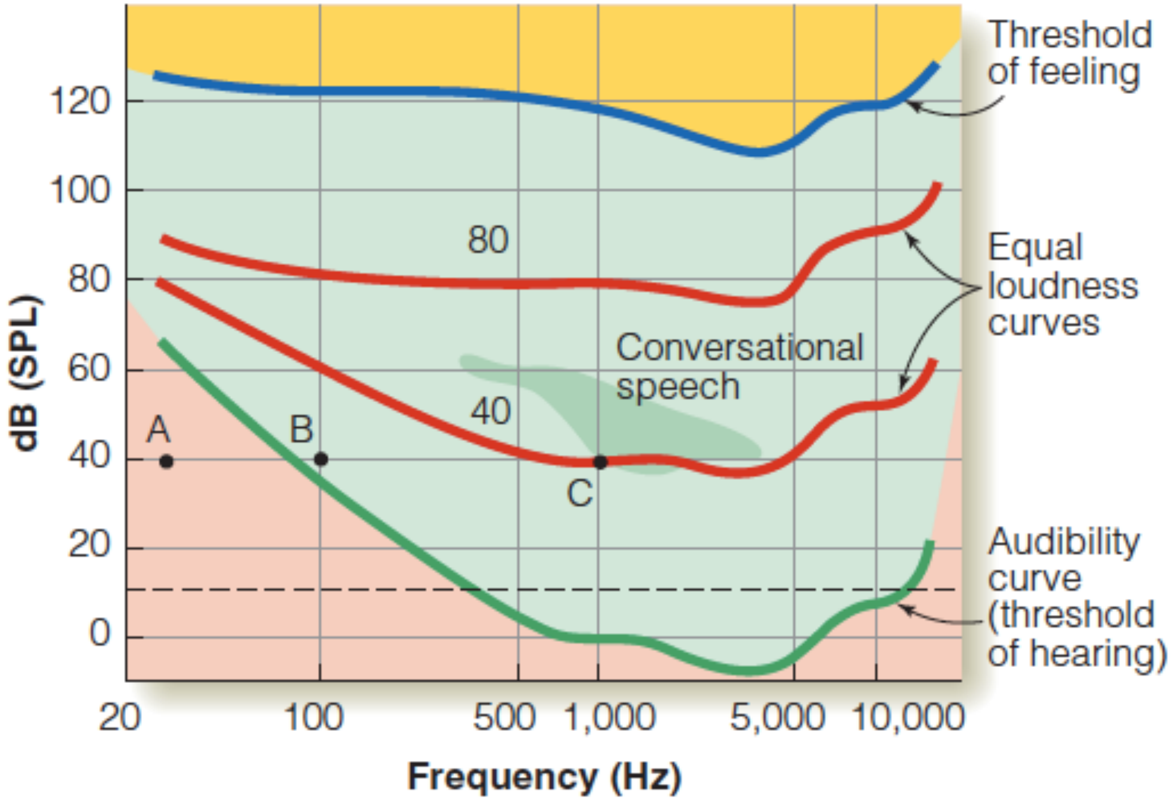


Audibility curve and auditory response area

Equal loudness curves
unit (phons)

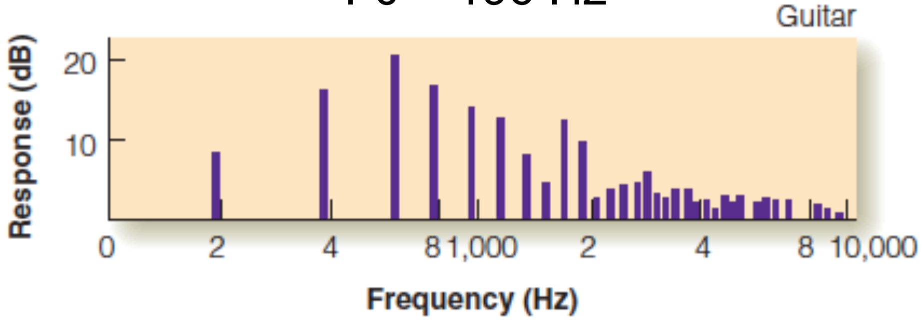
Phon = the level in dB of a 1kHz tone judged to be of the same loudness as the test stimulus

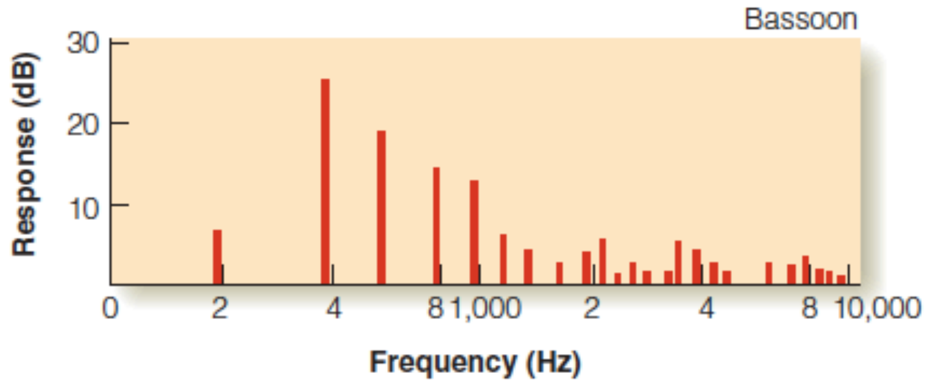


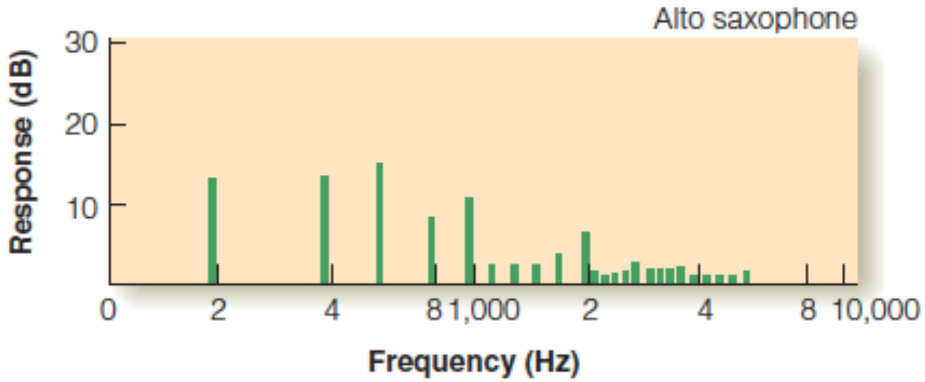


Timbre

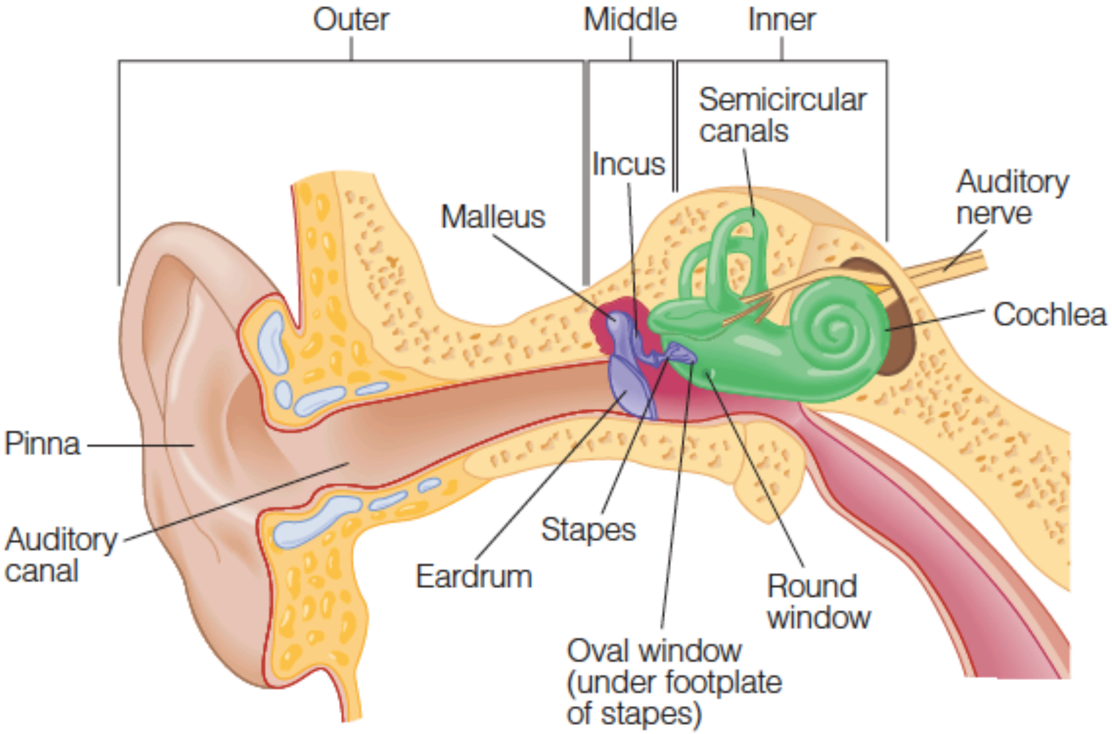
F0 = 196 Hz

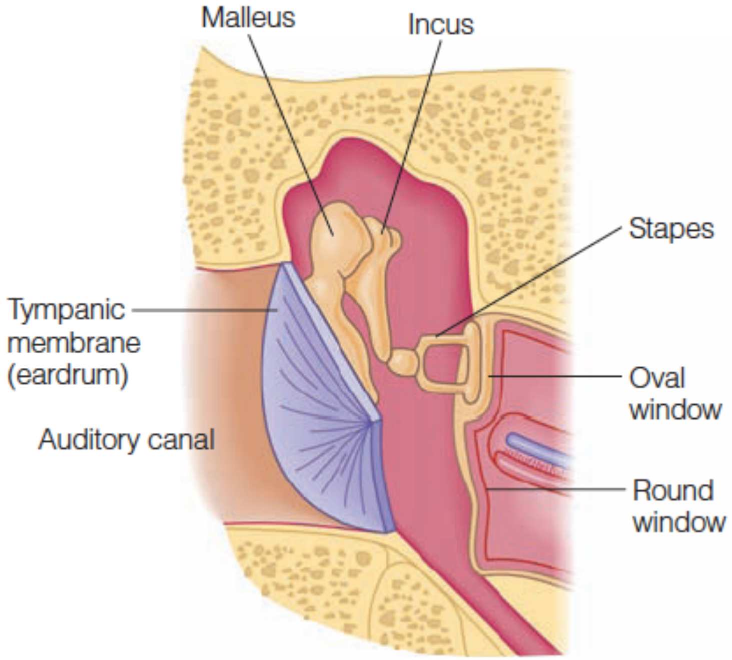






The ear

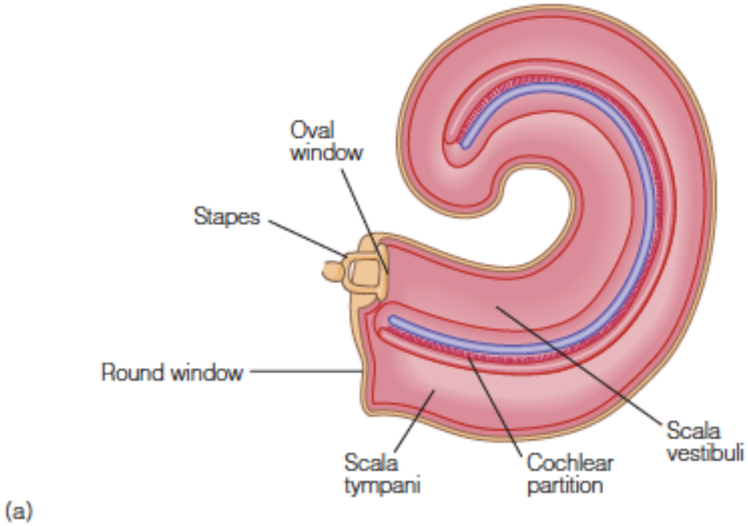


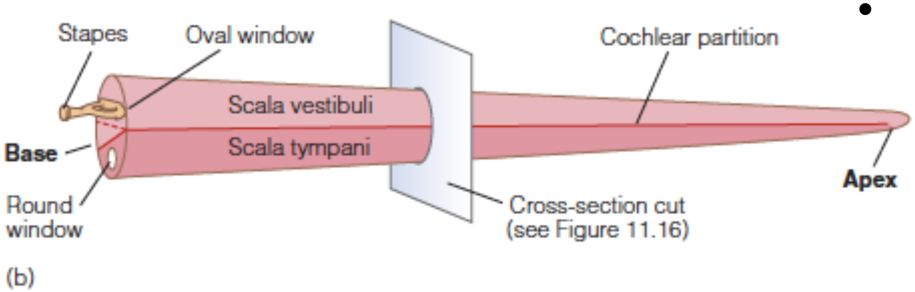


http://www.youtube.com/watch?v=dCyz8-eAs1I&feature=related

http://www.youtube.com/watch?NR=1&v=0jyxhozq89g&feature=fvwp

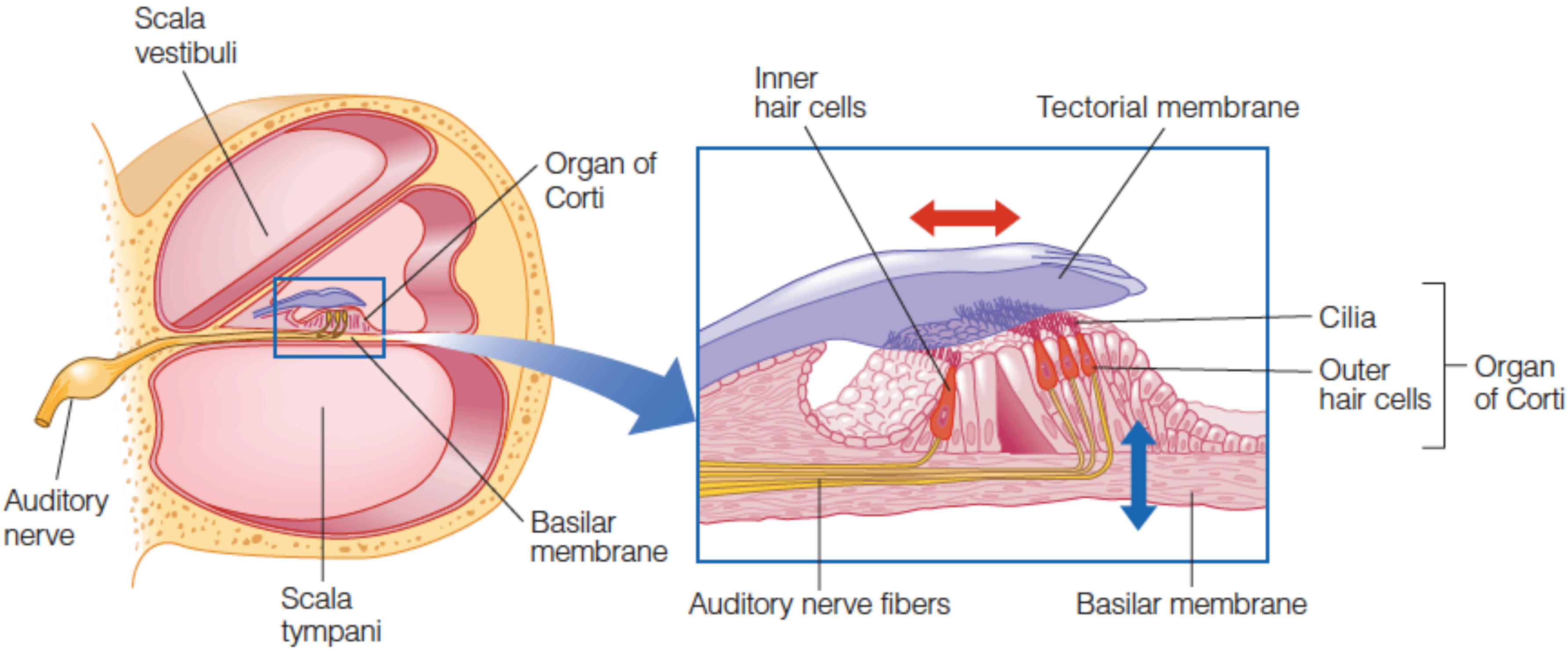
The cochlea



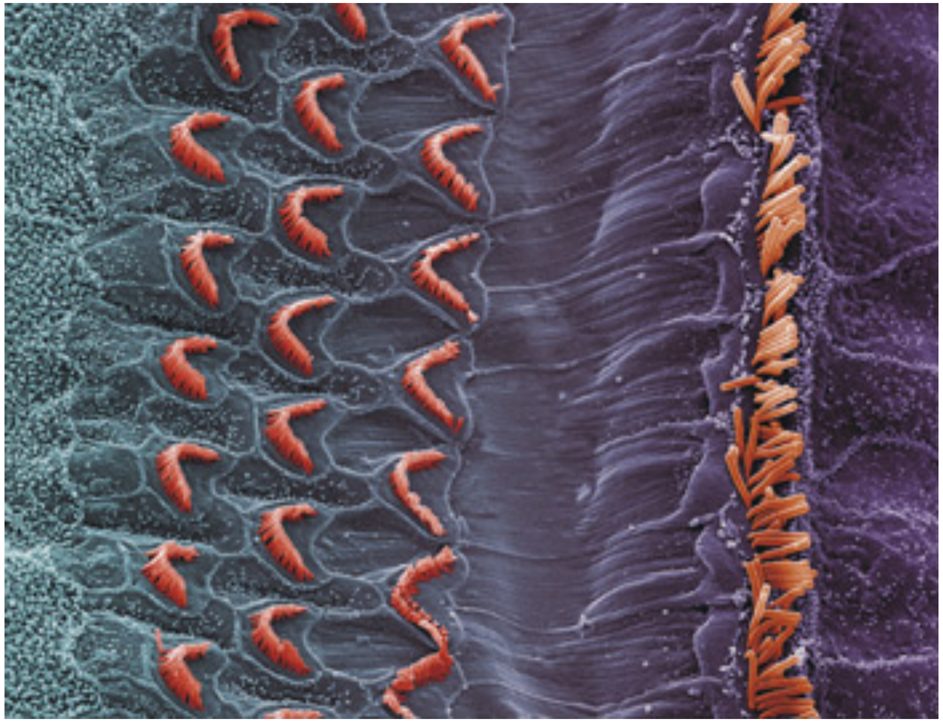


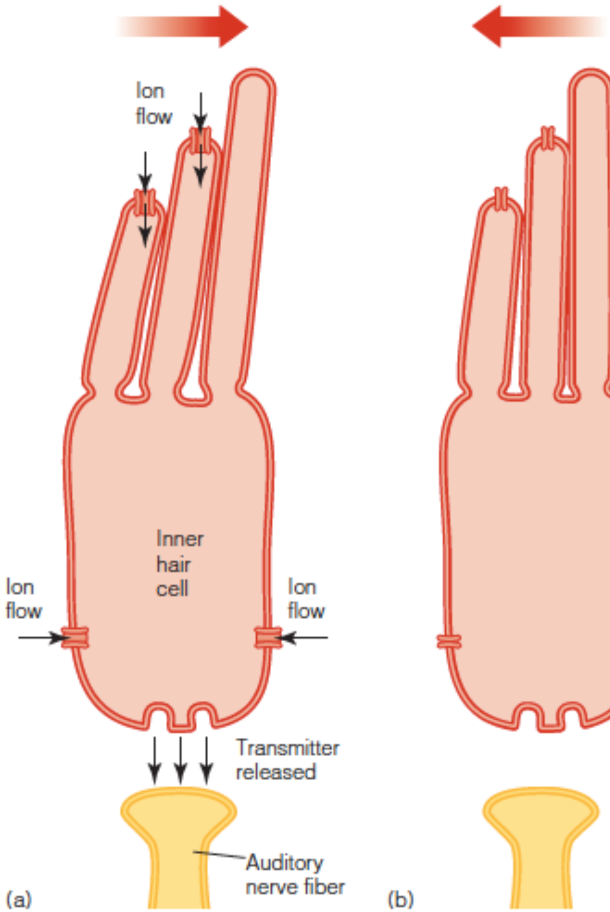
- The **cochlea** is a spiral tapered tube, 2¾ turn.
- The cochlea is divided into three parts along its length.
- **Helicotrema** is the part of the cochlear labyrinth where the scala tympani and the scala vestibuli meet.
- The **basilar membrane** is a vital part of the hearing process.
- As the cochlea becomes narrower, basilar membrane becomes wider. 0.1 mm at the oval window, 0.5 mm at helicotrema.

Basilar membrane

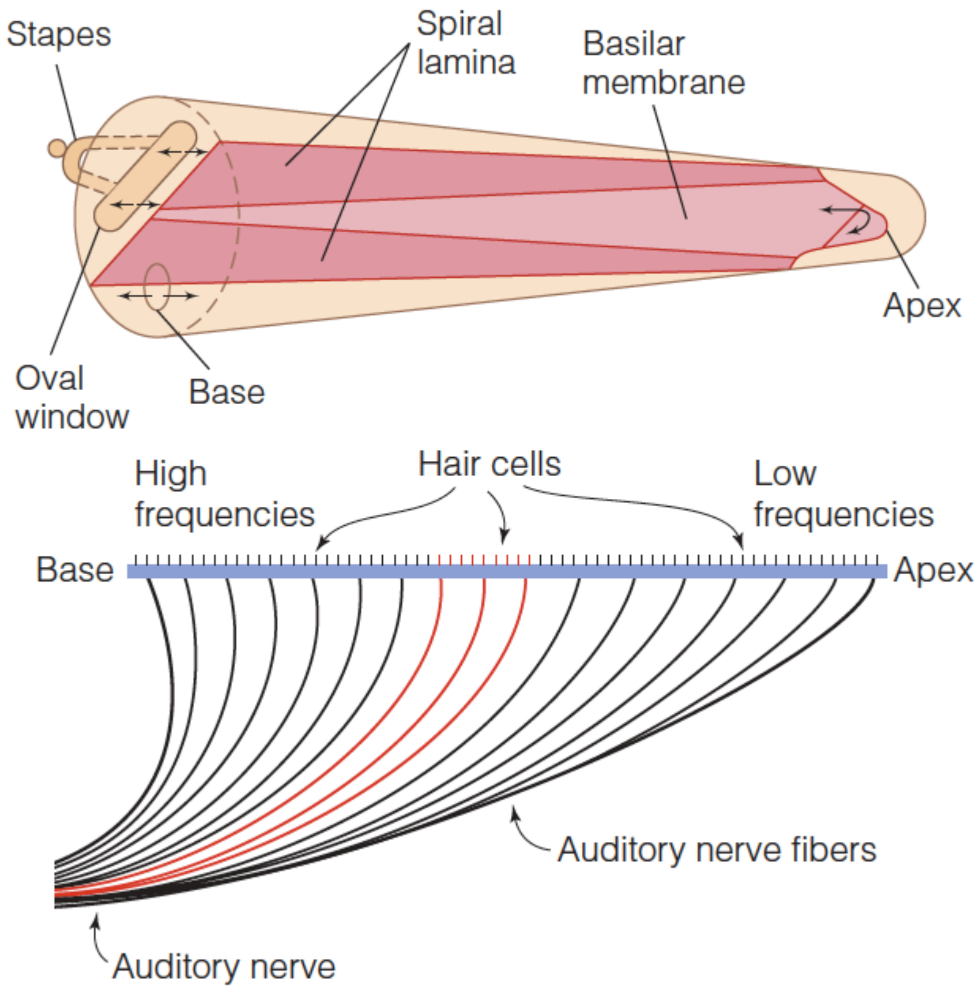


Inner and outer hairs cells





Frequencies



Sound frequency ➔ timing of firing

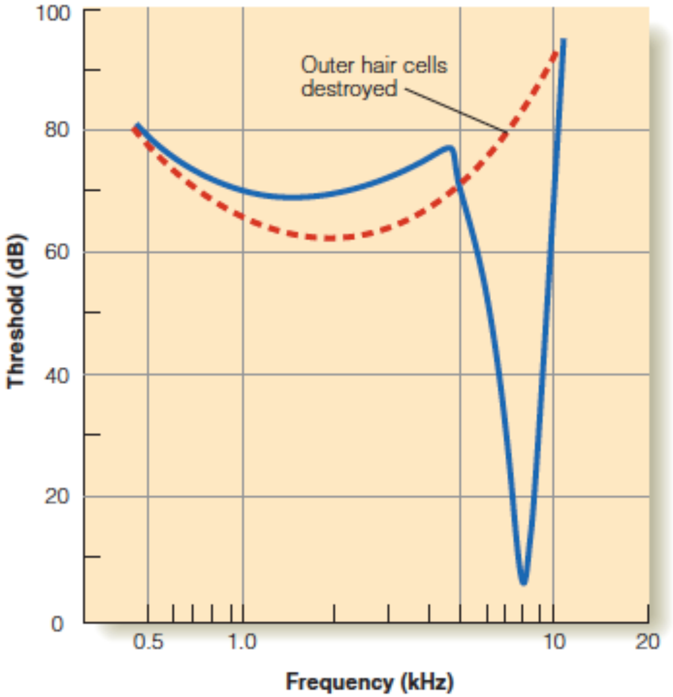
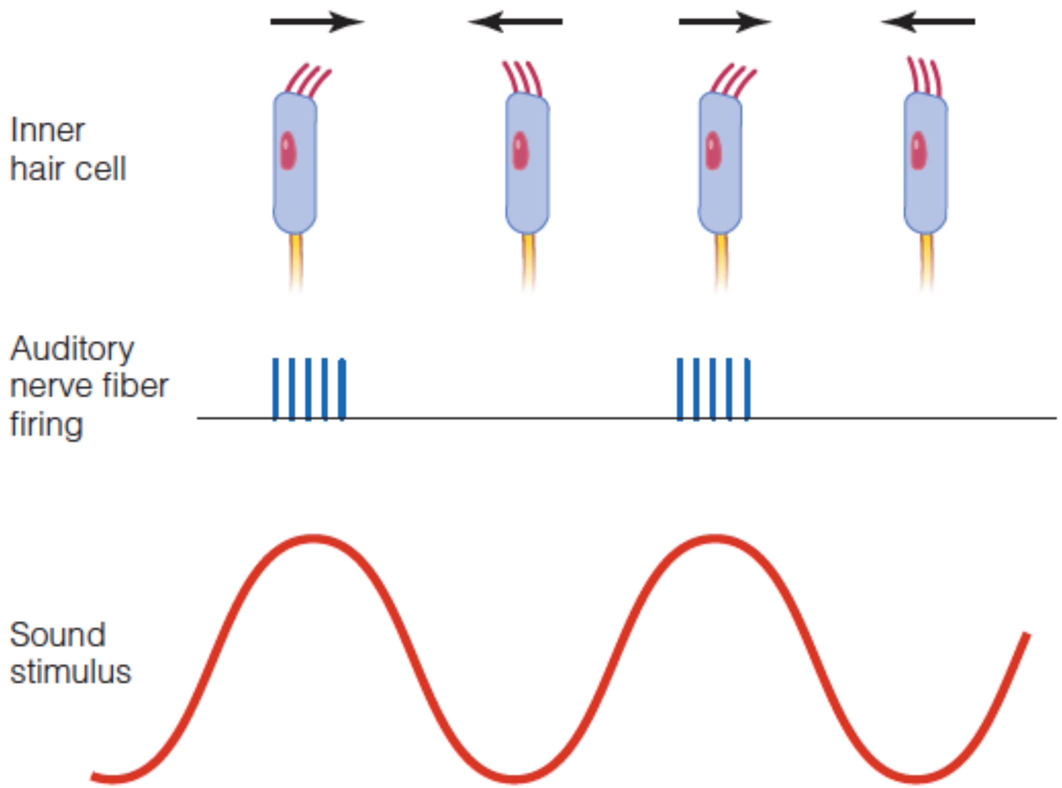
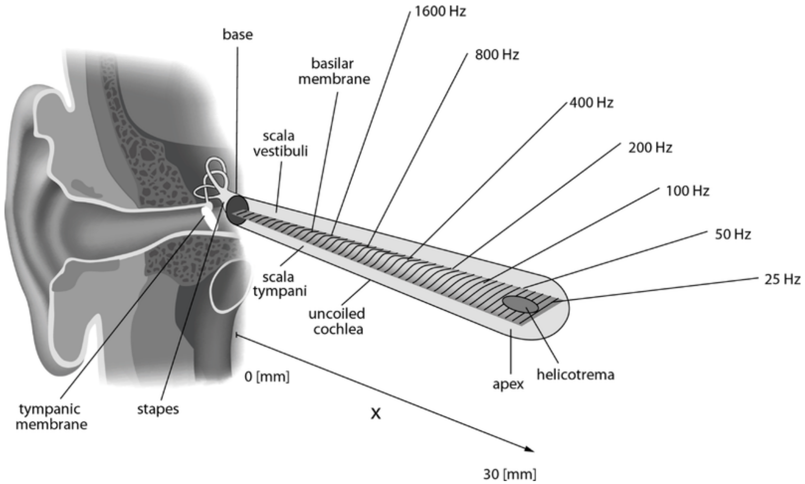
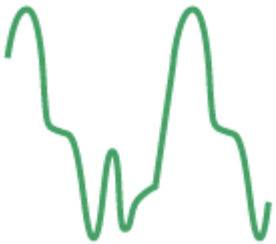


Figure 11.32 ▮ Effect of OHC damage on frequency tuning curve. The solid blue curve is the frequency tuning curve of a neuron with a characteristic frequency of about 8,000 Hz. The dashed red curve is the tuning curve for the same neuron after the outer hair cells were destroyed by injection of a chemical. *(Adapted from Fettiplace & Hackney, 2006.)*

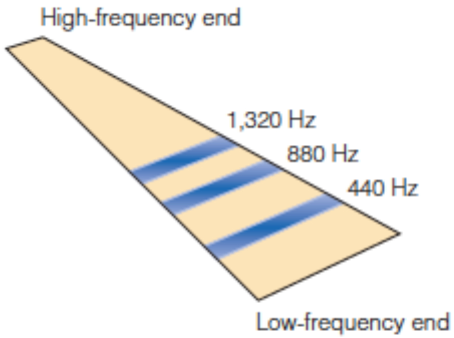


Basilar membrane: Vibration

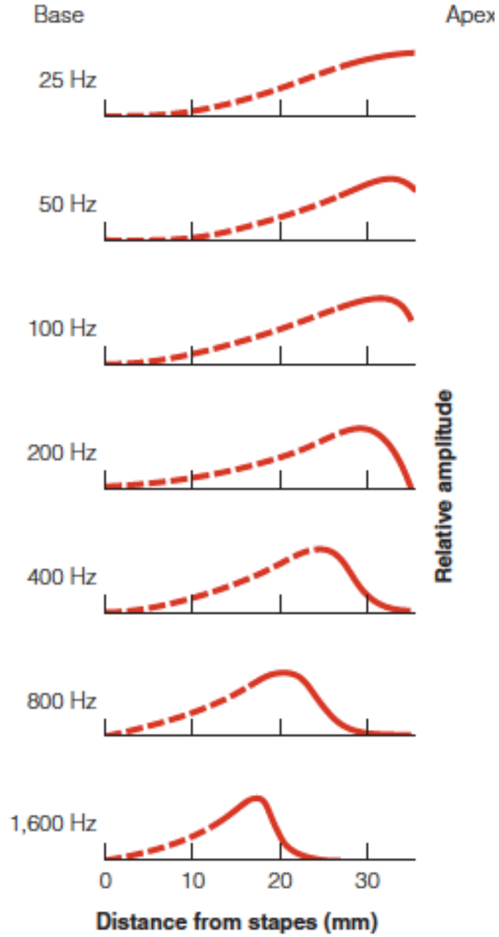




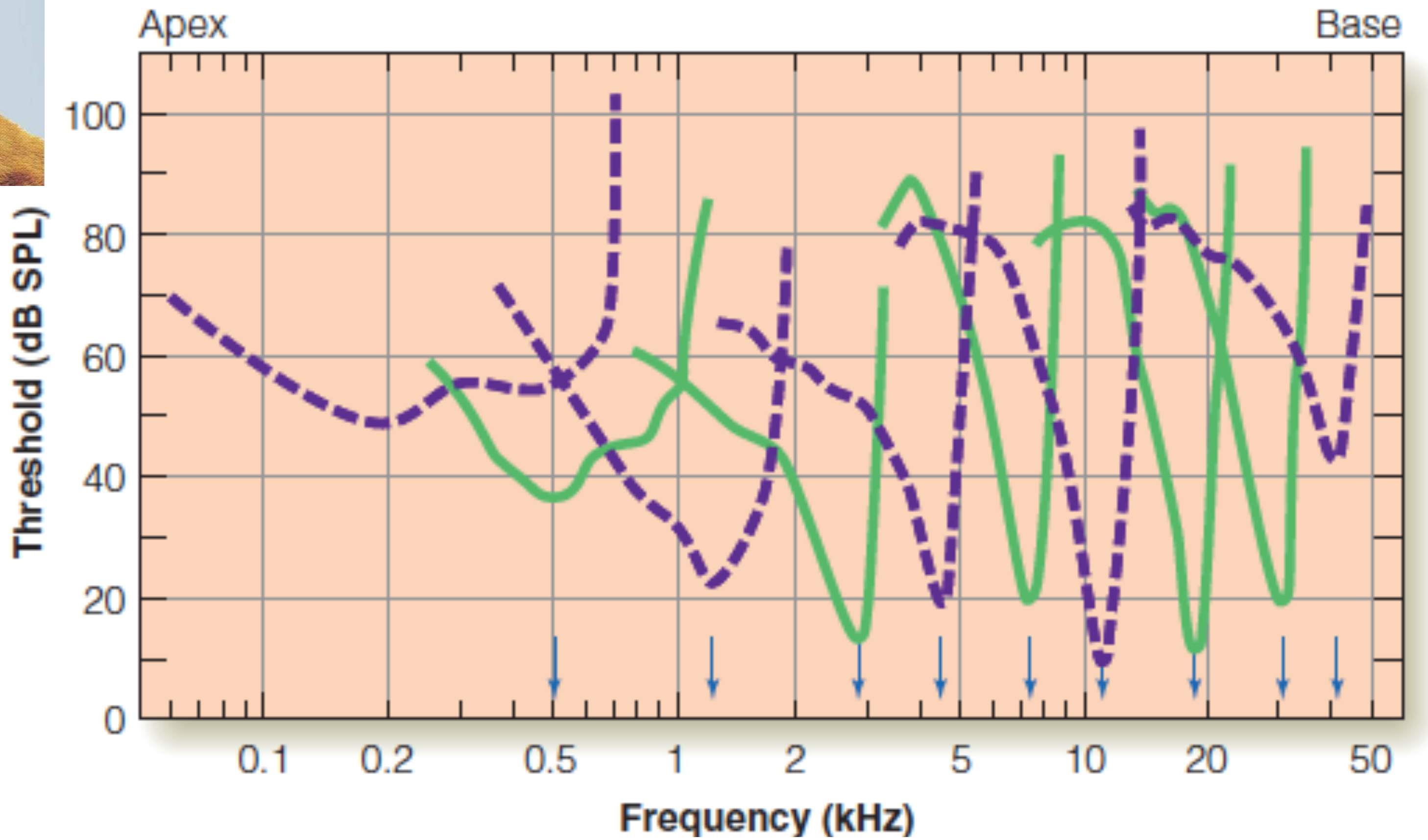
(a) Complex tone
(440, 880, 1,320 Hz harmonics)



(b) Basilar membrane



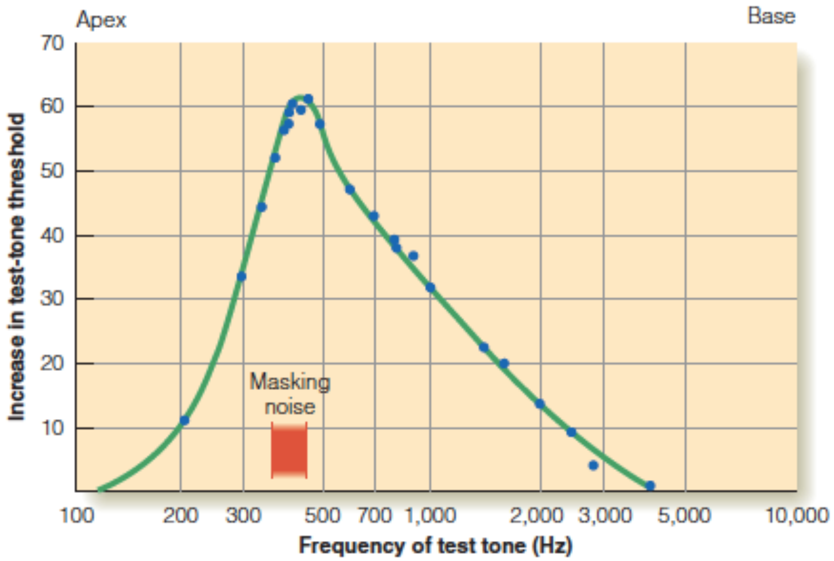
Tuning curves (cat auditory nerve fibers)

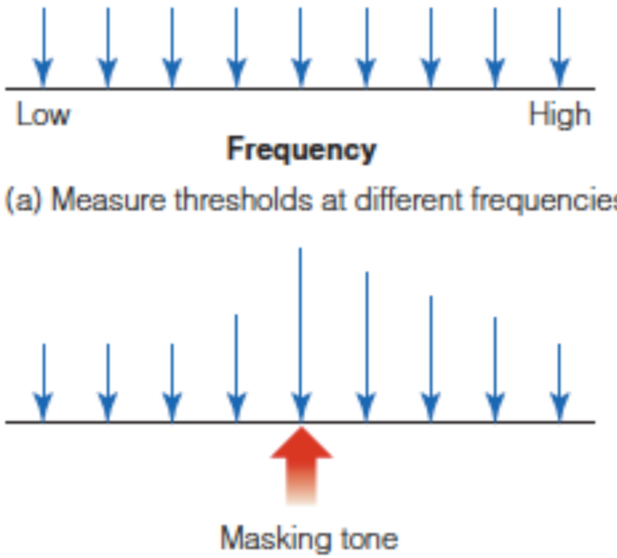


Masking

Low frequencies mask higher frequencies more than high frequencies mask lower frequencies

➔ The human auditory system is not sensitive to a detailed spectral structure, e.g. mp3.





(a) Measure thresholds at different frequencies (blue arrows)

(b) Remeasure thresholds with the masking tone present

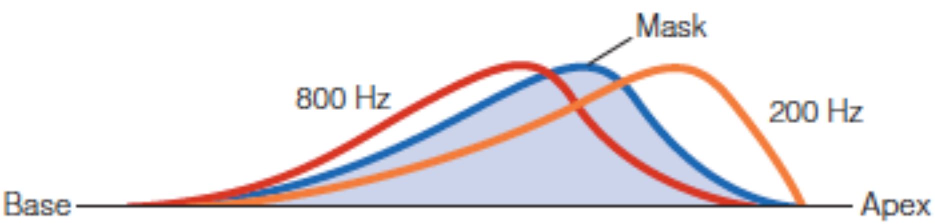


Figure 11.29 Vibration patterns caused by 200- and 800-Hz test tones, and the 400-Hz mask (shaded), taken from basilar membrane vibration patterns in Figure 11.24. Notice that the vibration caused by the masking tone overlaps the 800-Hz vibration more than the 200-Hz vibration.

Hearing loss

Conductive

- Deteriorated impedance conversion between the eardrum and the oval window
- Abnormalities at the eardrum, wax in the ear canal, injuries to the ossicles, inflammation in the middle ear

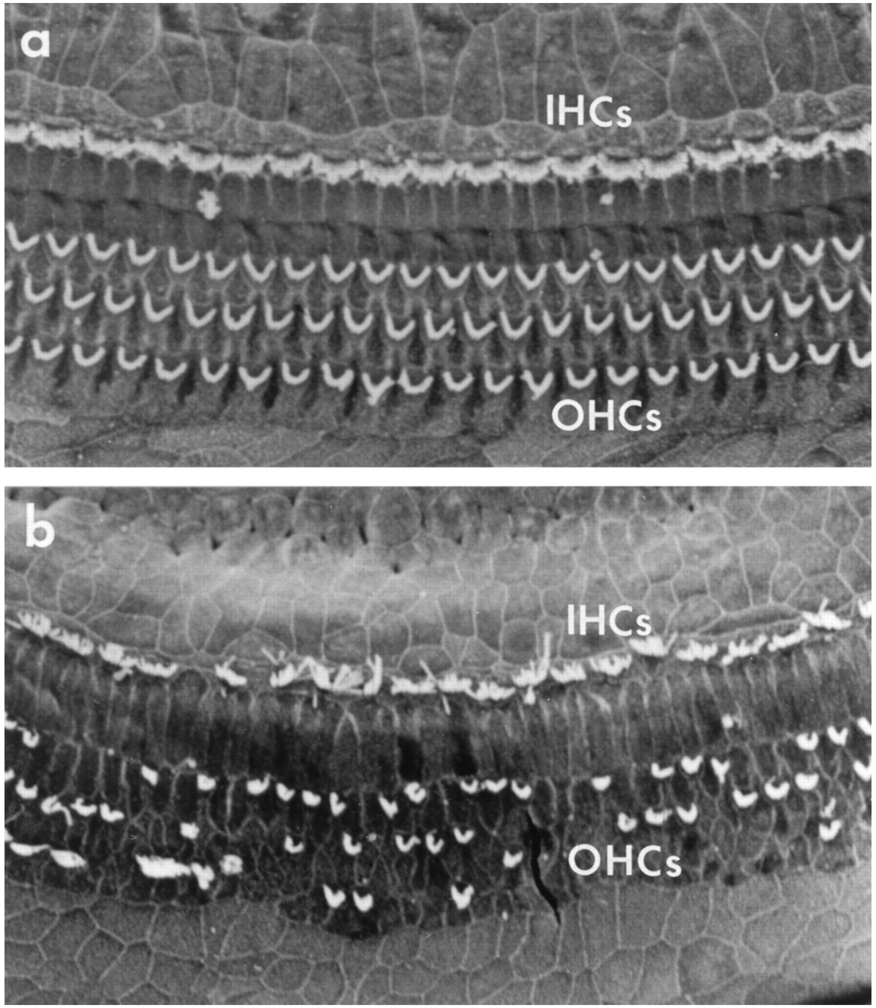
➔Sometimes possible to recover with surgery

Sensorineural

- Damage to the inner and outer hair cells
- Acoustic trauma, drugs, infection, congenital

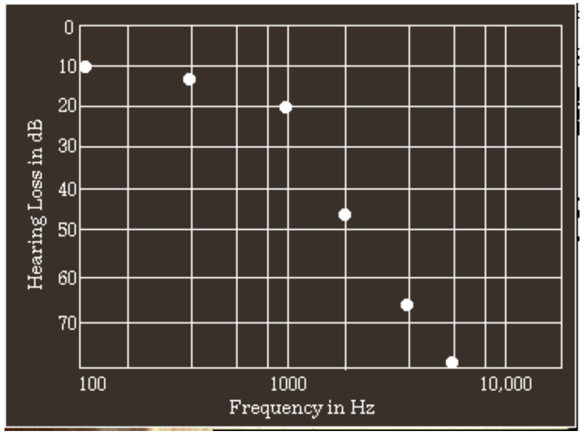
➔Usually permanent

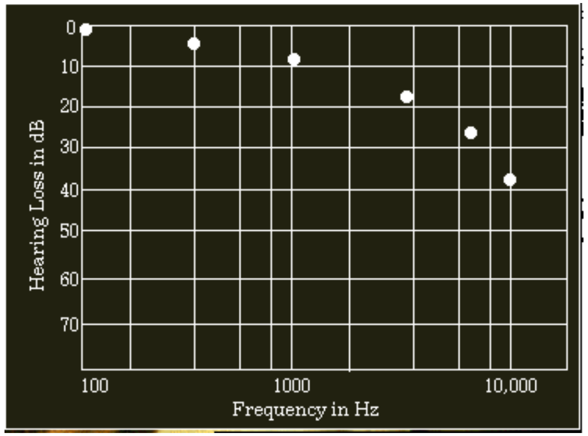
Sensorineural Hearing Loss

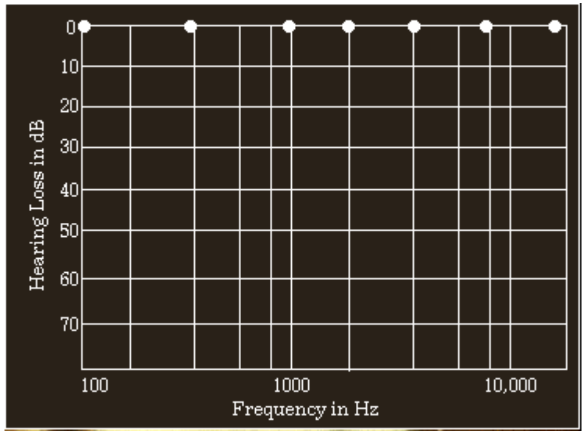


Hearing loss simulations

Severe	Moderate	Normal



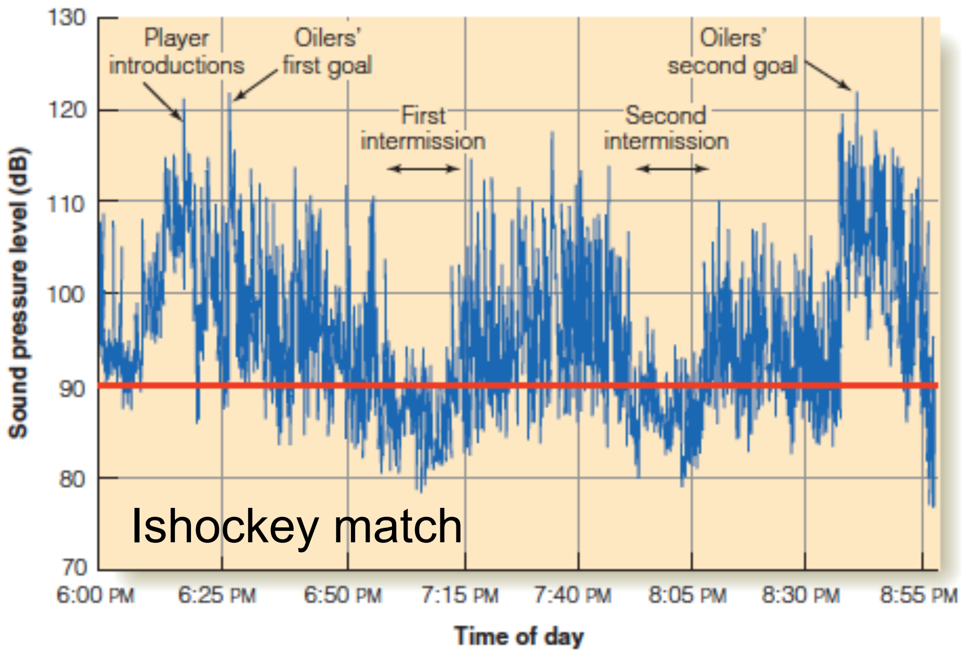


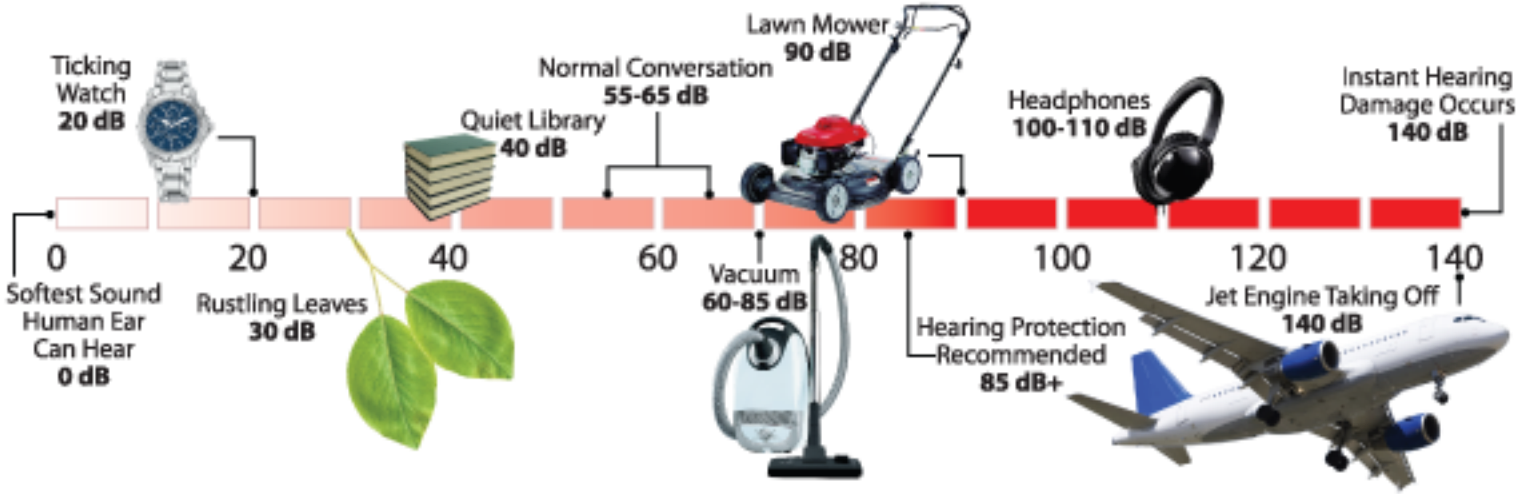


Hearing loss (Sweden)

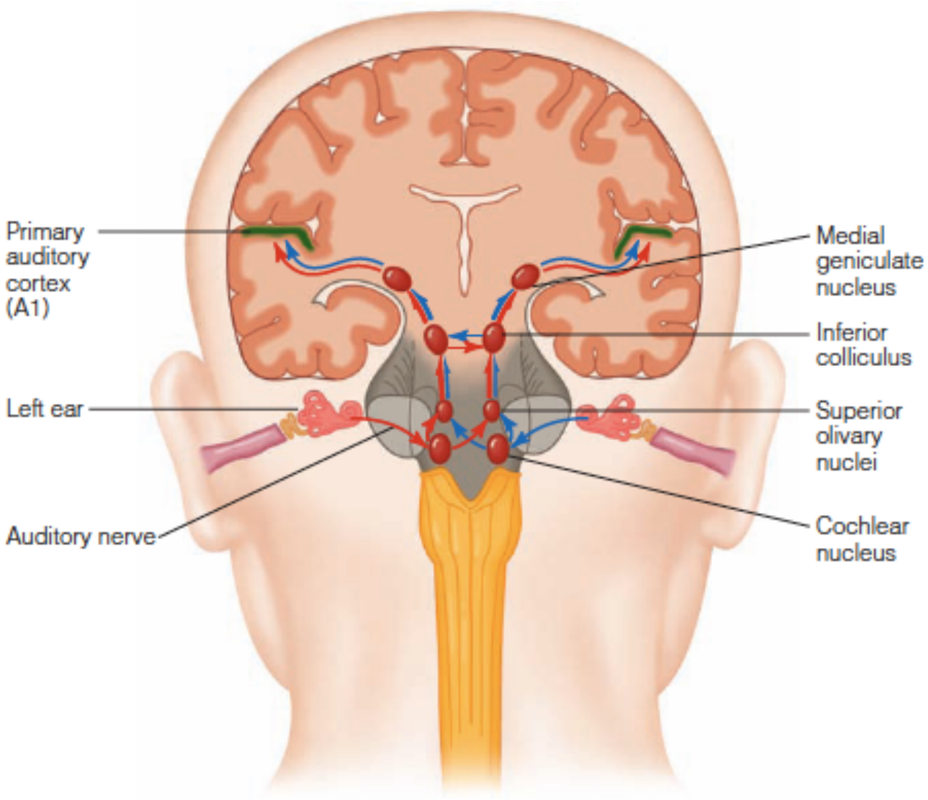
Åldersgrupp	Antal hsk	Andel hsk (procent)
16-24	27 664	2,9
25–34	62 913	5,4
35–44	103 522	8,2
45–54	139 250	11,8
55–64	230 900	20,0
65–74	173 445	22,7
75–84	201 522	34,6
Summa	**939 216**	**13,2**

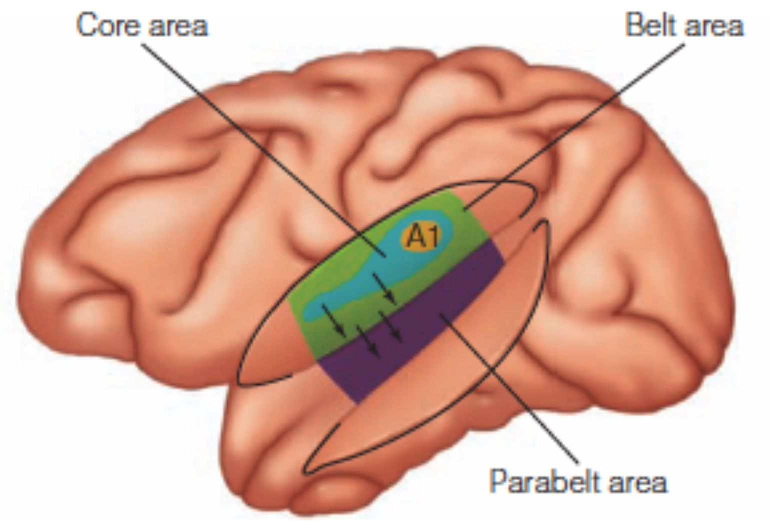
Safe sound level





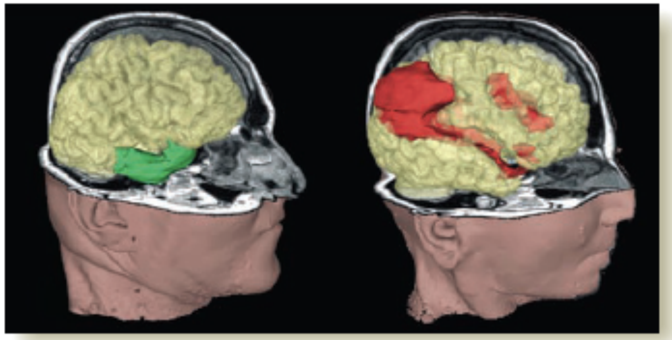
Auditory pathways



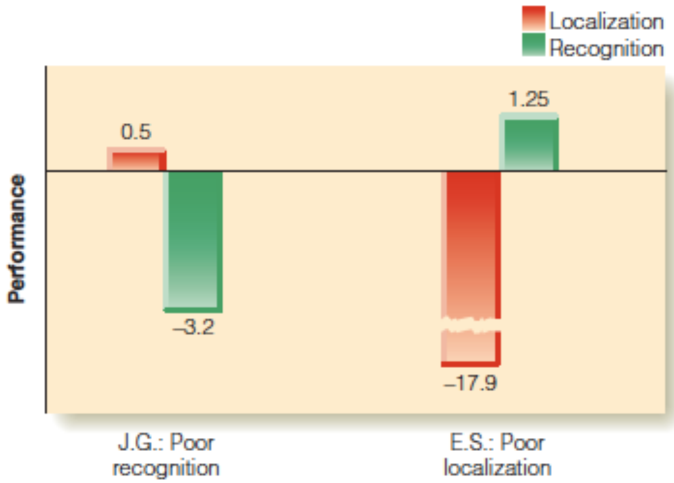


What and *where* streams

Areas with brain damage



(a)



(b)

Figure 11.40 Areas associated with *what* (green) and *where* (red) auditory functions as determined by brain imaging. *(Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory systems.* Proceedings of the National Academy of Sciences, 98, *12301–12306. Copyright 2001 National Academy of Sciences, U.S.A.)*

Auditory cortex is shaped by experience 1/2

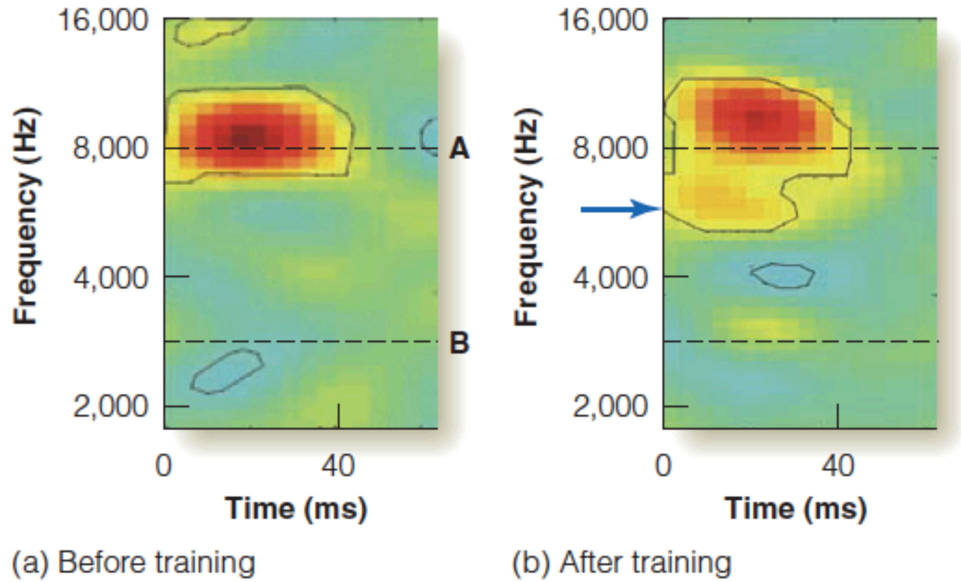


Figure 11.45 Response of a neuron in the ferret auditory cortex: (a) before training; (b) after training. See text for details. *(Reprinted by permission from Macmillan Publishers Ltd.: Fritz, J., Shamma, S., Elhilali, M., & Klein, D., Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex,* Nature Neuroscience, 6, *1216–1223. Copyright 2003.)*

Auditory cortex is shaped by experience 2/2

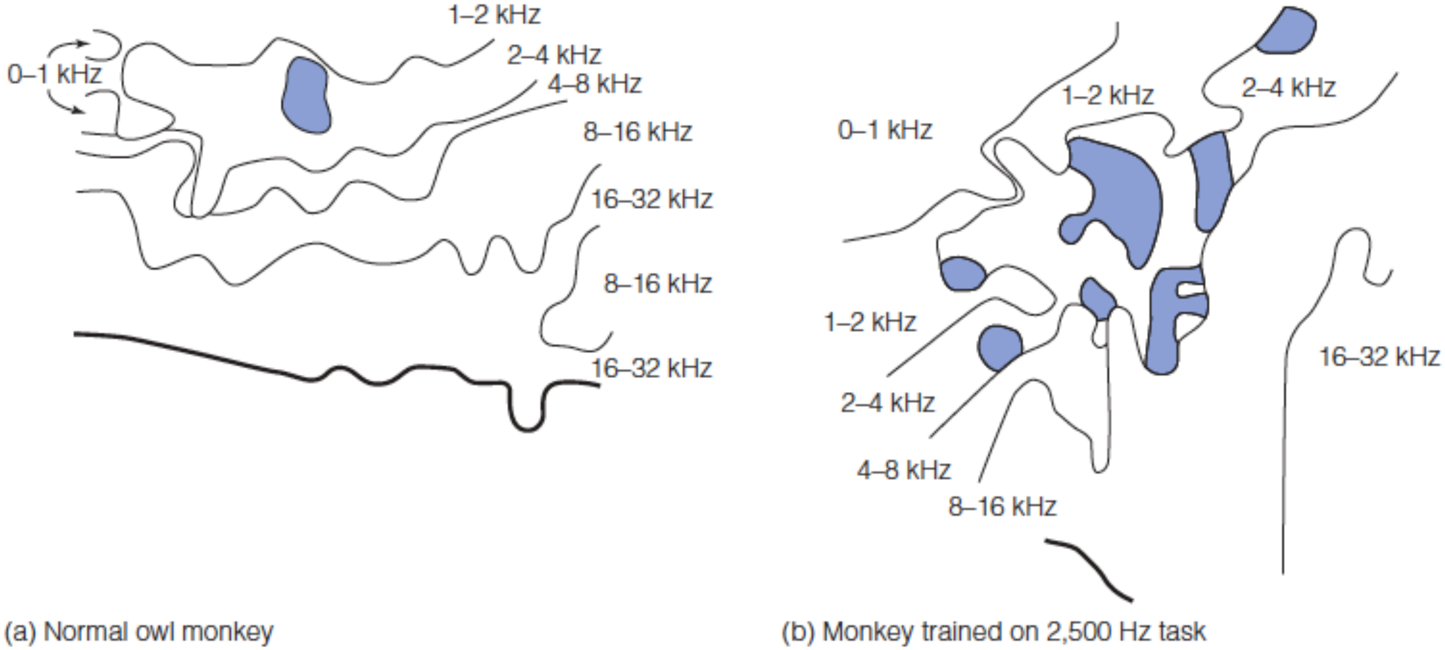
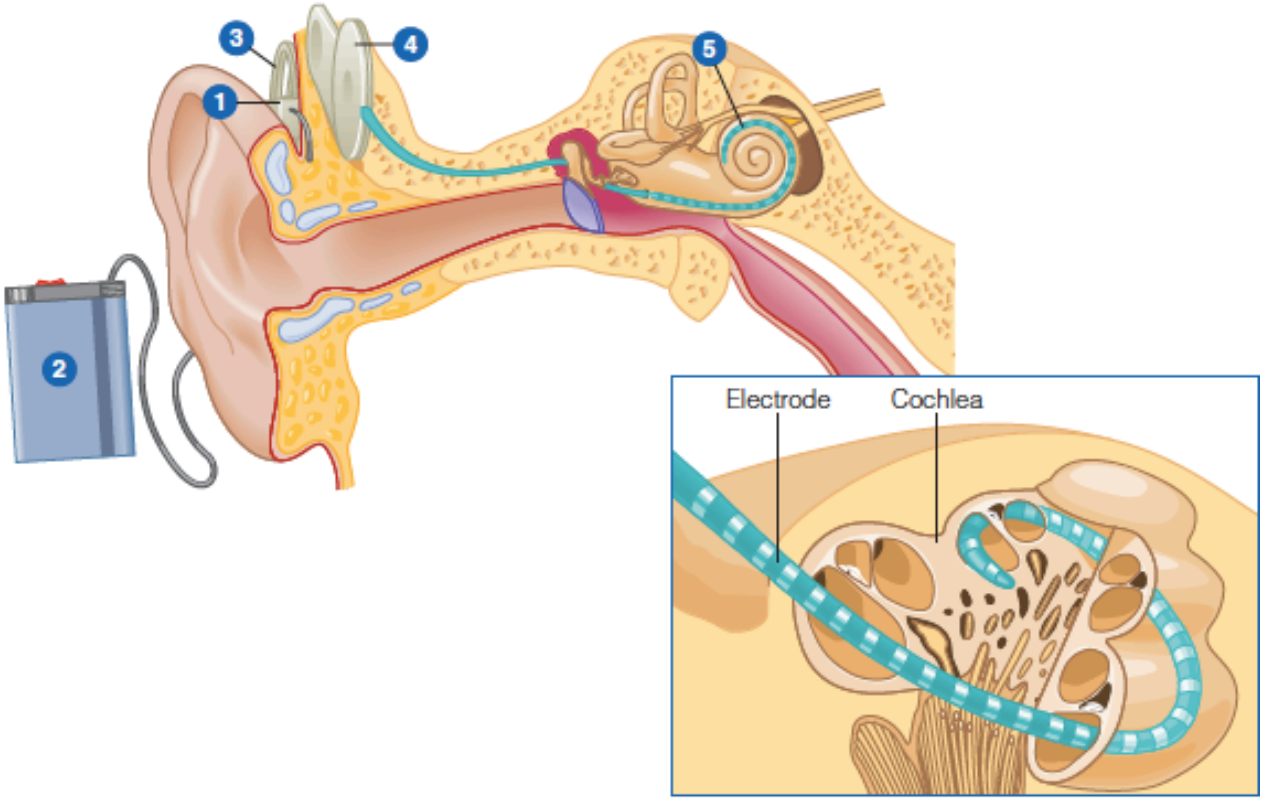


Figure 11.44 (a) Tonotopic map of the owl monkey's primary auditory receiving area (A1), showing areas that contain neurons with the characteristic frequencies indicated. The blue area contains neurons with CF = 2,500 Hz. (b) Tonotopic map of an owl monkey that was trained to discriminate between frequencies 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.)*

Cochlear implant (CI)

1. Microphone
2. Processor
3. Transmitter
4. Receiver
5. Electrodes

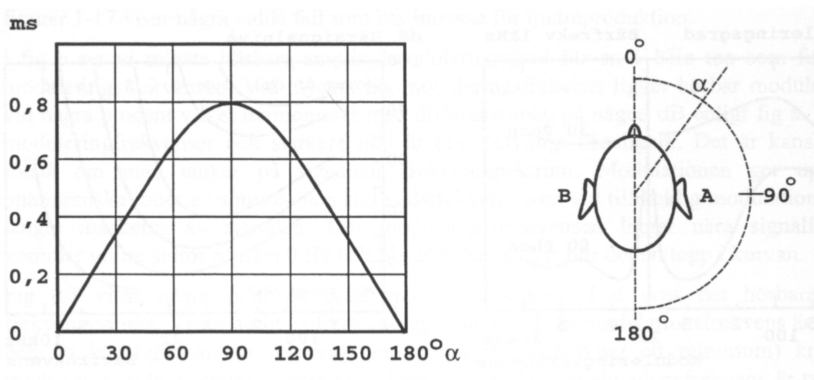


Sound localization – Horizontal plane

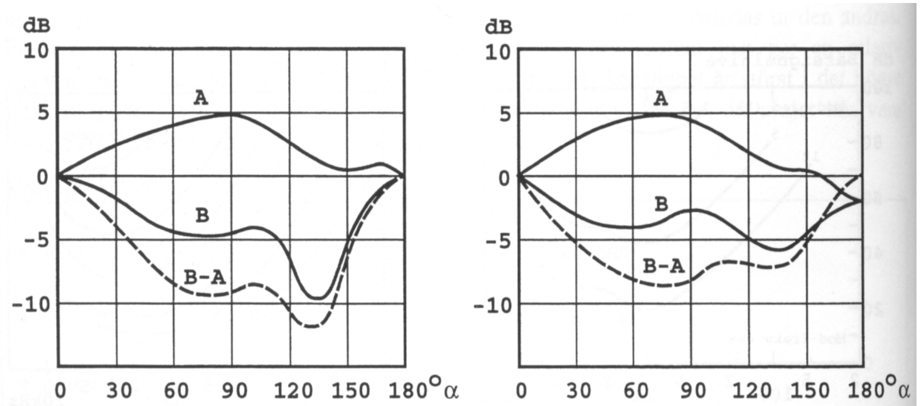
Time differences

Level differences

Time differences as function of angle



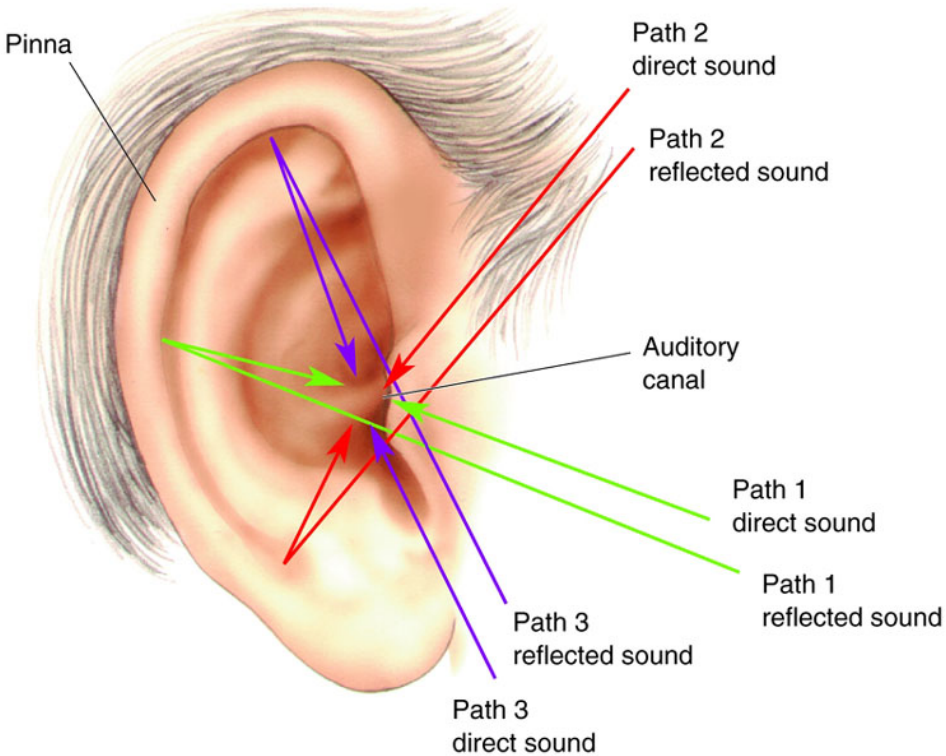
Level differences as function of angle



2 kHz tone

Speech

Sound localization – Vertical plane



Loud noises startle and get attention

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People. Chapter #48

From Deatherage (1972):

TABLE 48.1 How To Get Attention With Sounds

Audio alarm	Intensity	Attention-getting ability
Foghorn	Very high	Good, but not if there is a lot of other low frequency noise
Regular horn	High	Good
Whistle	High	Good, but only if intermittent
Siren	High	Good if pitch rises and falls
Bell	Medium	Good when there is other low-frequency noise
Buzzer	Low to medium	Good
Chimes or gong	Low to medium	Fair

Loud noises startle and get attention

Weinschenk, S.M. (2011). 100 Things Every Designer Needs to Know About People. Chapter #48

People habituate to stimuli

Examples: computer fan, church bells, clock

Takeaways

- If you're designing an application, you may have control over the sounds that occur when a people take certain actions, for example, making a mistake, reaching a goal, or donating money.
- Pick a sound that is appropriate to the amount of attention you need. Save the high-attention sounds for when it's really important, for example, if people are about to format their hard drives, or take an action that can't be undone.
- If you use sounds to get attention, then consider changing them so that people will not habituate and the sounds will continue to be attention-getting.