Psychoacoustics

The study of how humans perceive sound

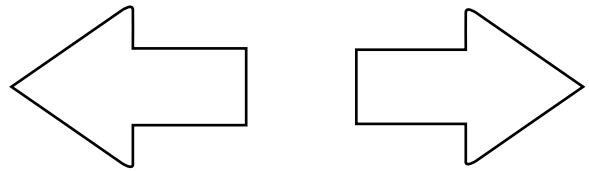
Objective vs Subjective qualities of sound

Objective qualities

-Frequency

-Level

-Spectrum

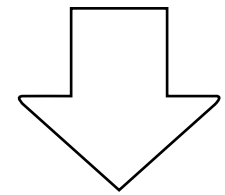


Subjective qualitites

-Pitch

-Loudness

-Timbre

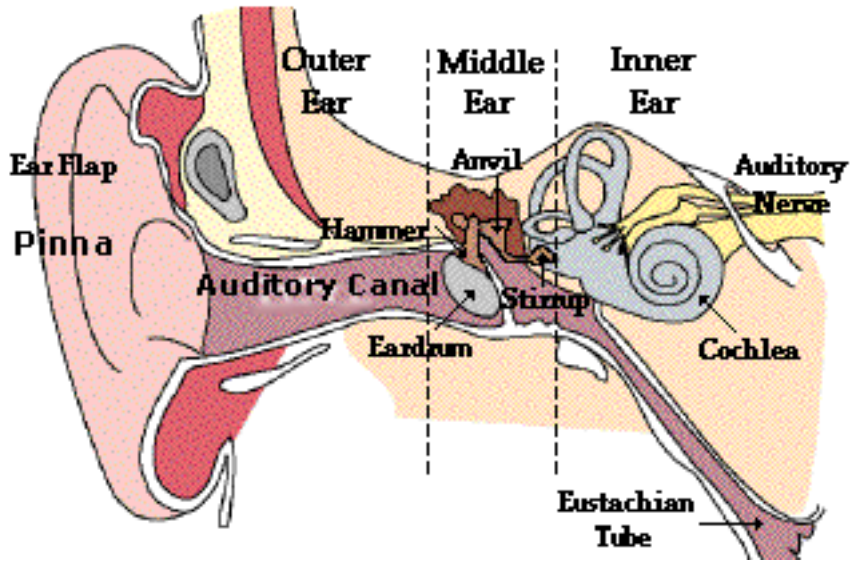


Psychoacoustics, a relatively new field, has elements of psychology, acoustics, physics, and biology.

The Human Ear

-Has 3 different parts with different functions.

-Together, they create the complicated window through which we capture pressure vibrations in our environment.



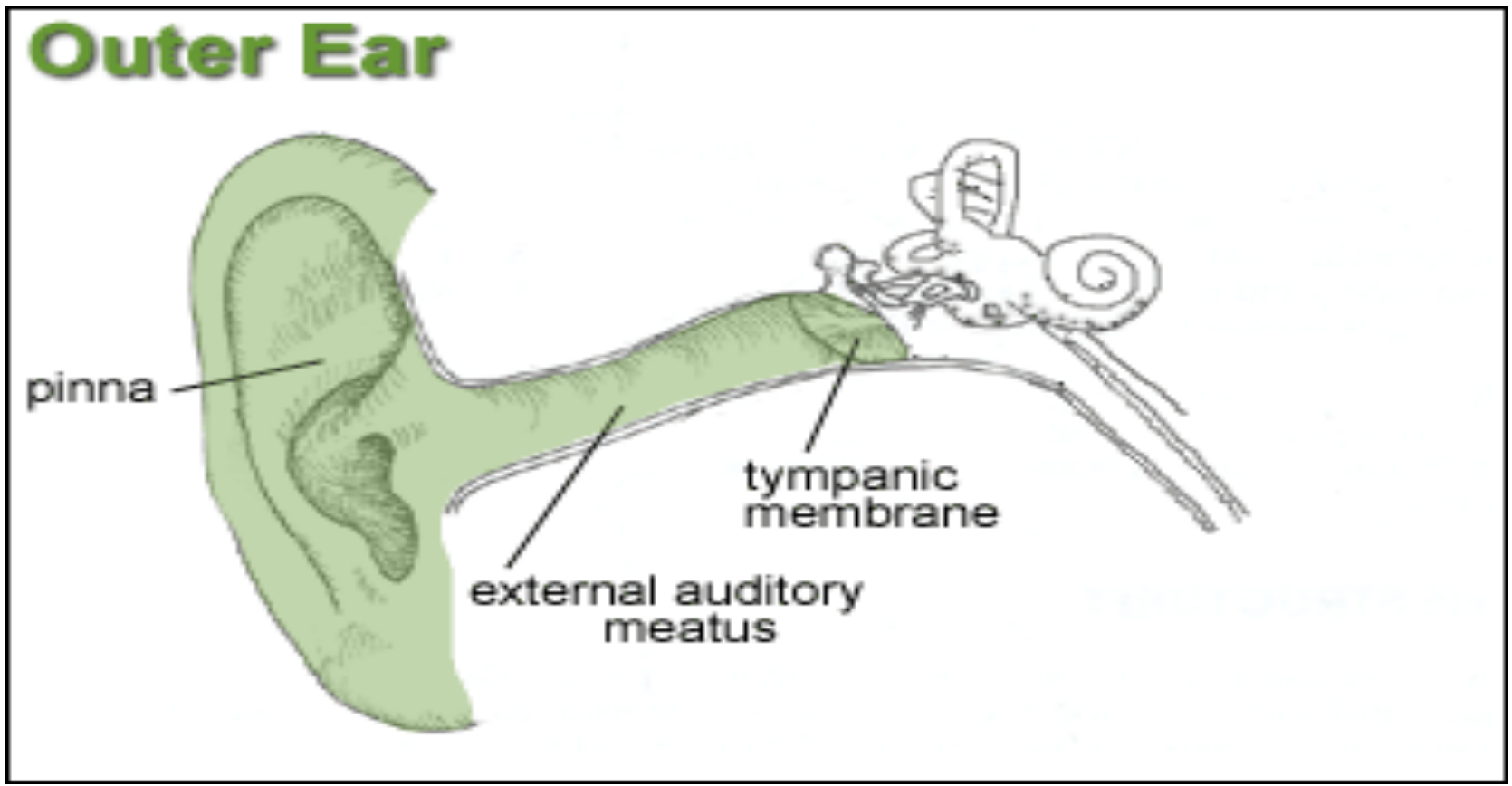
The Outer Ear

-Outer Ear: for capturing sound

Pinna: captures and amplifies sound, also used for localization and filtering

Temporal Bones: protection for ear, also contributes to hearing

Ear Canal: diameter 5-9 mm, length 2.5cm, resonant frequency of 3k, as much as 20dB of gain between 2.5-7k, hearing loss usually worse at 4k, terminates at the ear drum

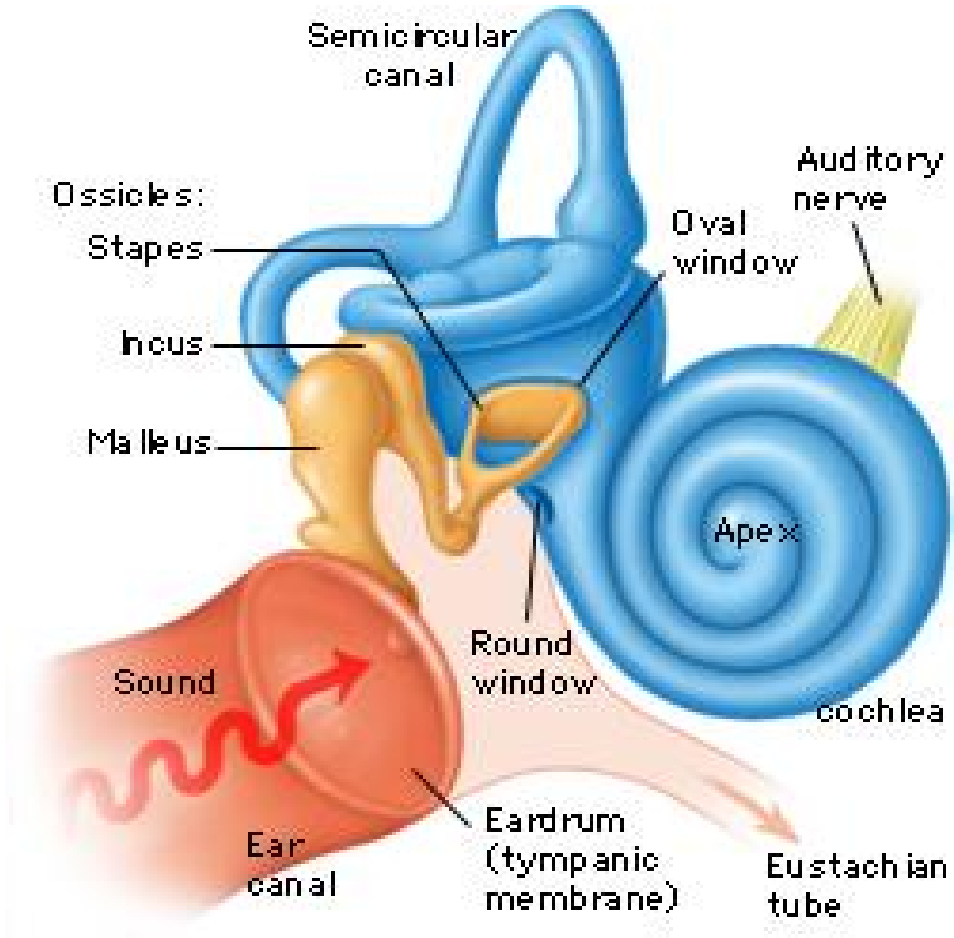


The middle ear

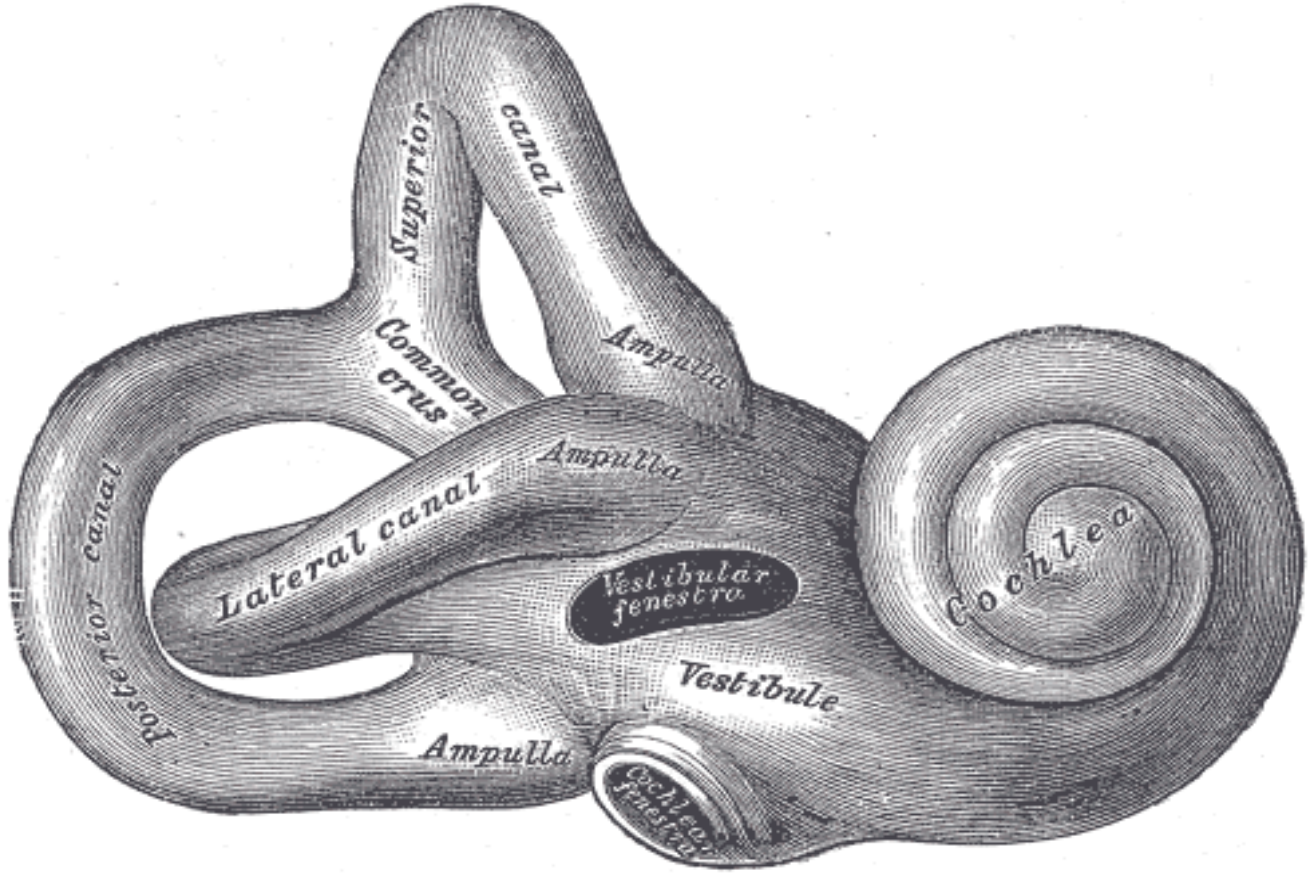
-Middle Ear: for matching impedance between outer and inner ear

Three Ossicles: fixed to ear drum, exhibit various motions to transmit acoustic energy to the inner ear

Acoustic Reflex: muscles attached to the ossicles tense to protect the ear from dangerously high sound pressure levels(only works with slow onset and LF sounds)



-Inner Ear: for analyzing and processing sound, has 2 systems



Vestibular System: for balance
Auditory System: for, well, hearing

Cochlea: snail shaped, has different membranes w/ different fluids, the Organ of Corti has hairs that feel movement and convert it to electrical impulses in the auditory nerve.
Inner Hair Cells: 1 row, about 1500 hairs, act like microphones, tuned to specific frequencies depending on their position
Outer Hair Cells: 3 rows, about 3500 hairs, they are active and change shape to add gain to the IHCs. Because they are active, and consume a lot of energy they are usually damaged first due to loud sound or toxins. Thus, just adding gain with a hearing aid won't work.

Nonlinearity of the Ear

-In a linear system, the frequency content of the input and output will be the same, though relative amplitudes and phases might be altered due to filtering.

-In a nonlinear system, output will include new frequencies not present in the original signal. Because our auditory system has mechanisms such as acoustic reflex in the middle ear and active processes in the inner ear, it is nonlinear. Because of this, even in a completely quiet environment, the inner ear might generate tones. This is a sign of a healthy, functioning inner ear, unlike *tinnitus* , which results from exposure to dangerously high SPL.

Perception of Phase

-Phase recognition is quite important for the perception of timbre. For example, an impulse tone and white noise have identical amplitude spectrum. The only difference is in the phase spectrum. Also, with speech, scrambling the phase spectrum while leaving the amplitude spectrum unchanged will make the speech unintelligible.

-The auditory nerve fires at about 5kHz. This means that our ears can only differentiate pitch up to 5k.

-The signals from our ears are combined in the brain, and the interaural phase difference (IPD) is important for spatial hearing.

Auditory Neurons

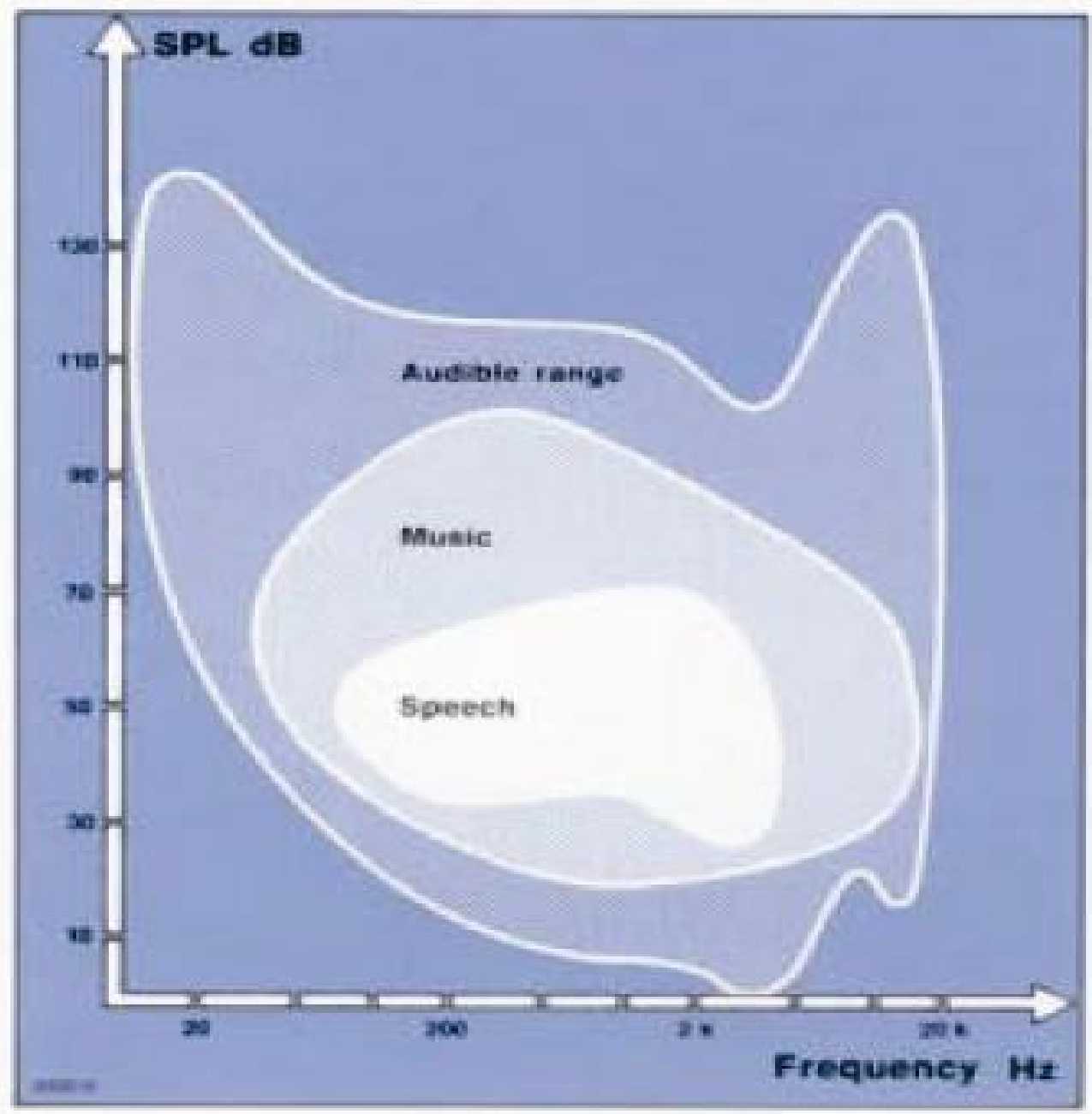
-As level increases, frequency selectivity drops. Since the brain perceives pitch based on neuron input, louder tones will tend to sound sharp.

-Masking: In the presence of a louder "masking" signal, a quieter signal will will not be heard until it passes a certain threshold. This threshold depends mainly on the bandwidth of the masker and the maskee. The masking signal can played played at the same time, before (up to 100ms), or even after! (up to 20ms) the quieter signal. This has very useful applications in digital audio encoding such as mp3 conversion.

-Auditory Filters/Critical Bands: Level perception depends on the bandwidth of the signal. With constant SPL, only adjusting a signal's bandwidth, loudness will be perceived as constant up to a certain critical threshold, above which we will perceive the sound to be getting louder, even though the SPL does not change. This led to the idea of auditory filters, a mechanism for listening to sound through a set of averaging filters instead of to every hair individually.

Auditory Area

The part of the auditory spectrum that we can hear. It is bounded by the threshold of hearing, or the softest sounds we can hear, and the threshold of pain, the point at which sound becomes painful to us.

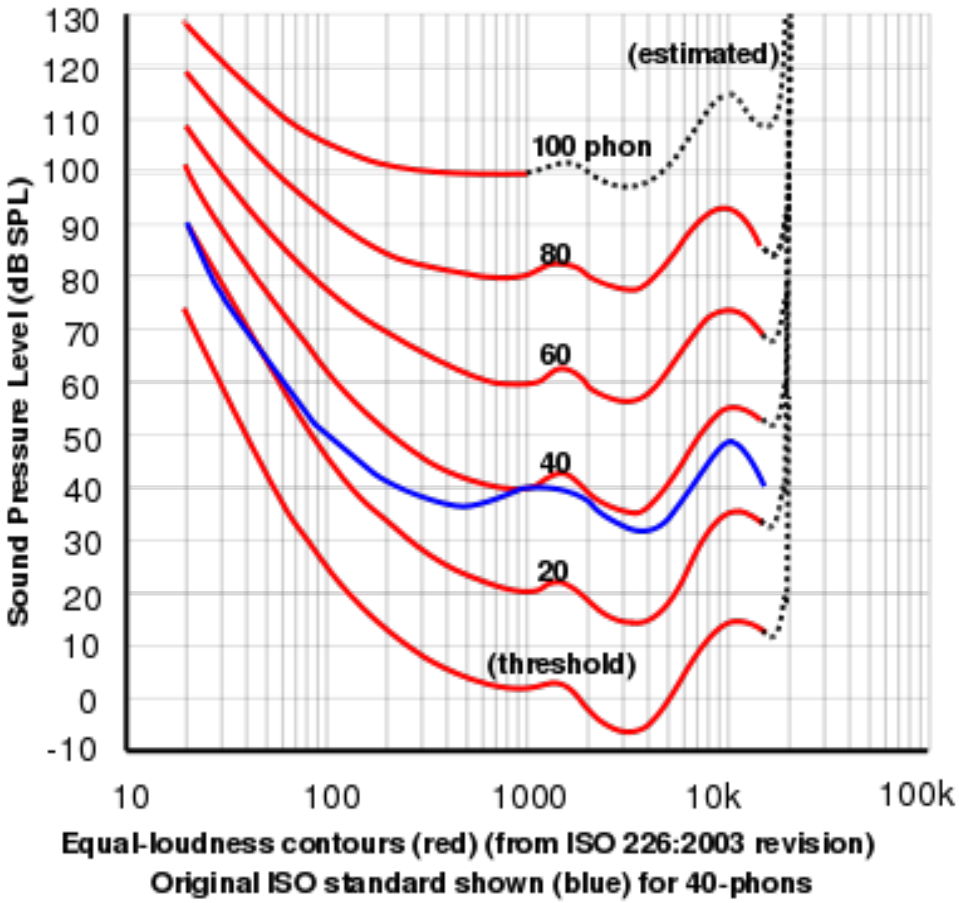


Level Intensity vs. Loudness

-Level is an objective measurement, while loudness is a subjective perception that is a function of our ear and the given signal.

-The Fletcher-Munson Curves show how different frequencies are perceived at different loudness levels by our ears. Notice the most sensitive frequency band is in the range of human speech.

-Weber's Law: states that the JND (just noticable difference) in level intensity is proportional to the overall intensity of a signal. A signal that is more than 30dB has a JND of .4dB. Below 30dB overall level, the JND is 3dB.



Pitch

-A subjective quality we assign to sound that depends on the frequency, SPL, and the waveform of a signal. To perceive pitch, a sound must be able to be matched by a listener with a "pure tone". A small drum might sound higher than a big drum, but one cannot match it to a pure tone, which is why you wouldn't say it has a pitch.

-Pitch is also binaural, meaning that it is an addition of what both of your ears hear. If you play a pure tone in one ear that is the 3rd harmonic of a given fundamental, and a pure tone in the other ear that is the 5th harmonic of that fundamental, you will hear the fundamental tone, not two distinct tones.

Pitch (continued)

-Range of Pitch: we can only perceive pitch from 20Hz-5kHz, even though our hearing extends up to 15-20kHz. This is because our auditory nerve fires at a rate of 5kHz, so higher frequencies can be heard, but not translated into a discernable pitch. This property of our ear that ends at 5kHz is called "phase-locking".

-Dependance on level: "Steven's Rule" states that below 1k, pitch decreases as level increases. Above 3k, pitch increases as level increases. In between 1-3k, level has little effect on pitch.

-A tone must have a certain duration for a pitch to be perceived. Otherwise it is heard as a click.

Timbre (pronounced tam-ber), means "sound color"

-Depends mainly on the spectrum of a signal, but also on waveform, SPL, frequency, the location of the spectrum, and temporal characteristics.

-Both amplitude and phase spectrum must be taken into account when analyzing tambre.

Why 2 Ears?

-Obviously, 1 acts as a backup should anything happen to the other.

-Also, having 2 ears aids us in sound localization, makes conversational speech more intelligible, and gives us immunity to background noise.

-The ITD and ILD between ears play a huge roll in sound perception and localization.

-The Hass Effect: a click played in the left ear and the right ear at exactly the same time will be perceived as coming from directly in front of you.

Precedence Effect

-Precedence Effect: two clicks, delayed up to 5ms, will be perceived as one sound, but our ears will localize the sound from only the first click. With voice or instrument of any other large bandwidth signal, the delay can be extended up to 40ms. This has very practical applications. Imagine a large auditorium with side-fill speakers toward the back. The precedence effect can be utilized in order to create the illusion that the sound is coming from the stage, even though in actuality, most of the sound that the people in back are hearing is coming from the side-fills.

Because of the precedence effect, our ears will lock onto a sound and localize from it, even if that initial sound ceases. Imagine two loudspeakers in a room. One begins to play a constant pure tone, while the other remains silent. Then, the first speaker is rolled off while the second begins to play the tone. Even though the first speaker is off, and can even be unplugged, the sound will still be perceived as coming from it, and not the second speaker. This effect can last several minutes.

Cocktail Party Effect

Our ears can filter background conversational noise and lock on to certain voices in the mix of sound.

Distance Perception

-High frequencies roll off over long distances, which is why thunder is low and rumbly from far away, and much sharper when close. However, this effect requires large distances, and is insignificant in the distances usually dealt with in sound.