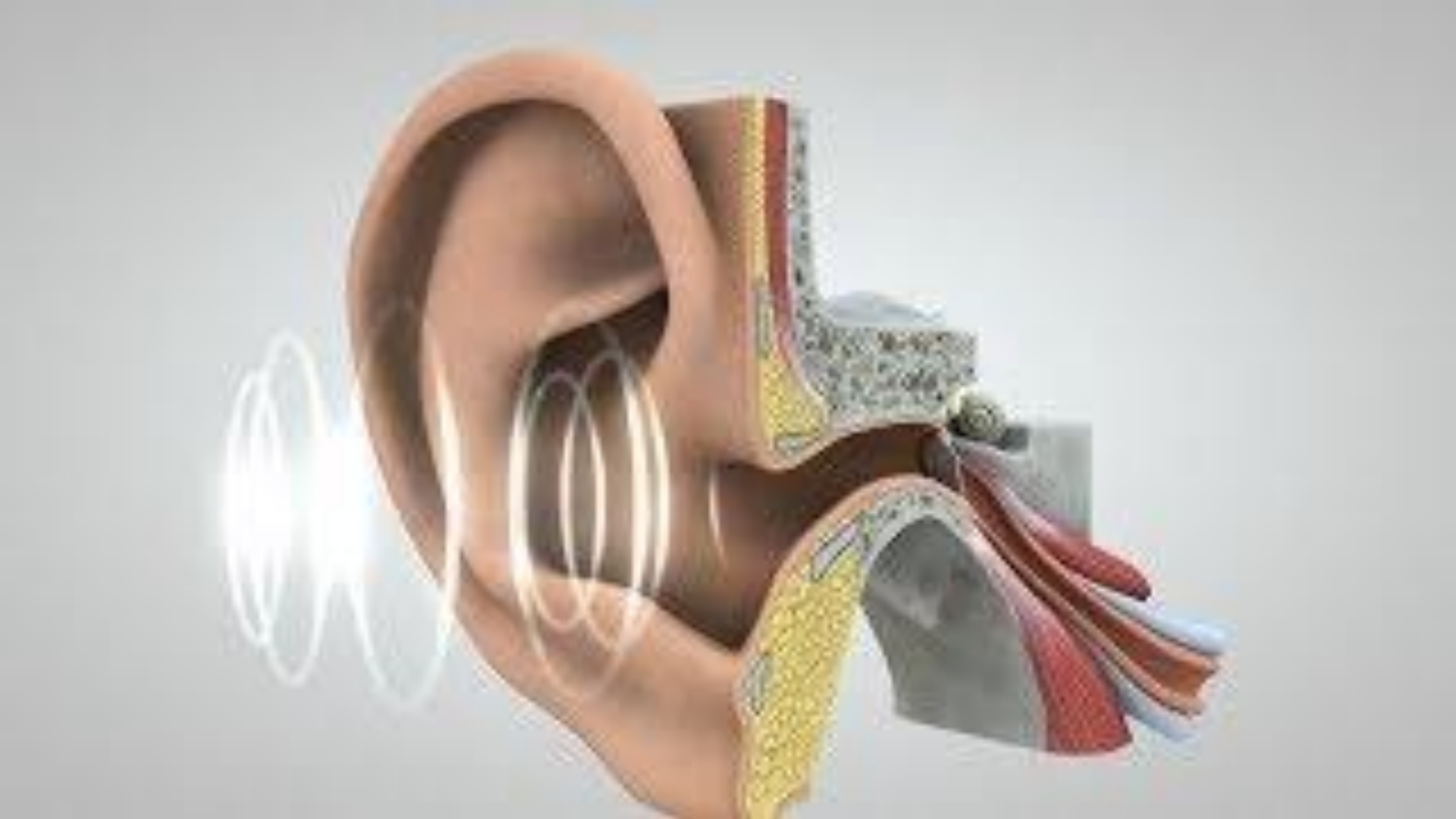
MUS 7: Auditory Processing

Instructor: Jingwei Liu

S124 UC San Diego



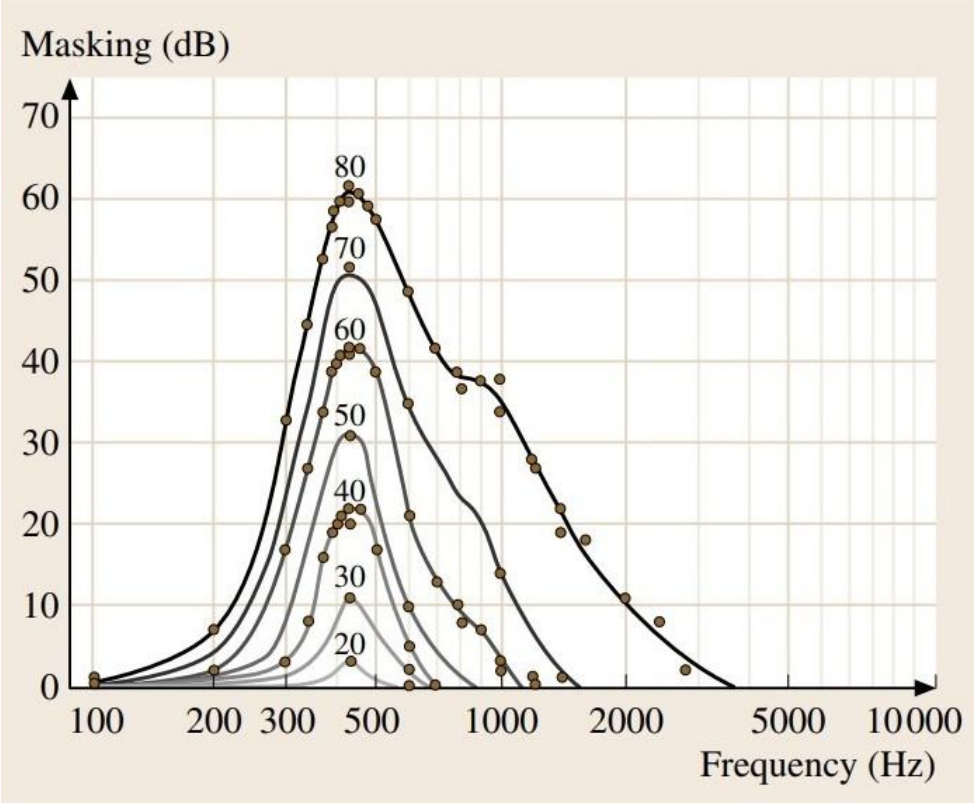

Auditory processing
Khan Academy

Auditory Filter

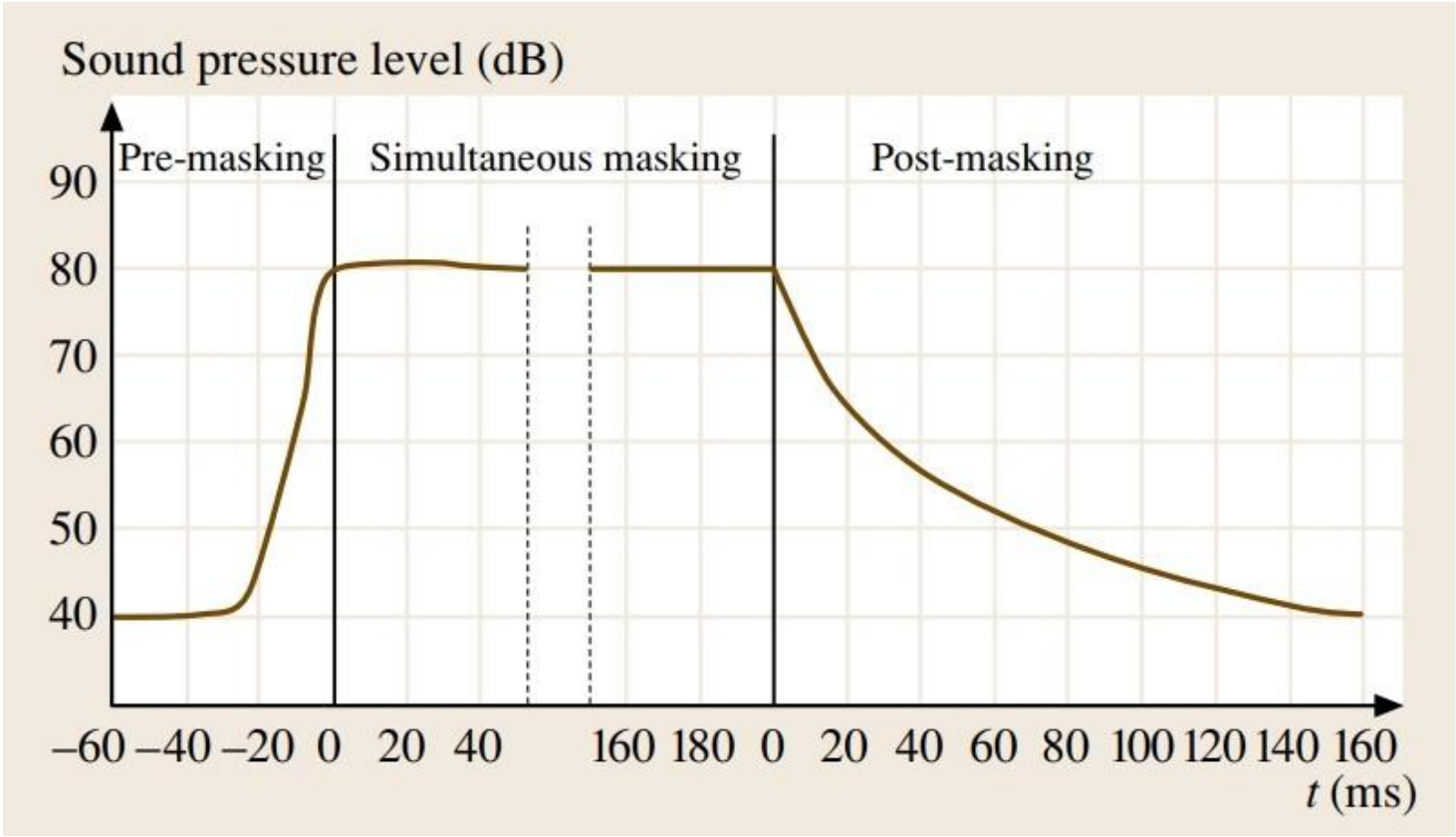
- Our ability to separate the components of a complex sound depends, at least in part, on the frequency analysis that takes place on the basilar membrane.
- Fletcher and Helmholtz suggested that the peripheral auditory system behaves as if it contains a bank of bandpass filters, with overlapping passbands. These filters are now called the *auditory filters*.
- Fletcher thought that the basilar membrane provided the basis for the auditory filters. Each location on the basilar membrane responds to a limited range of frequencies, so each different point corresponds to a filter with a different center frequency.

Auditory Masking

- The process by which the threshold of audibility for one sound is raised by the presence of another (masking) sound.
- A signal is most easily masked by a sound having frequency components close to, or the same as, those of the signal.



Temporal Masking



Critical Band

- In audiology and psychoacoustics the concept of critical bands, introduced by Harvey Fletcher, describes the frequency bandwidth of the "auditory filter" created by the cochlea.
- Roughly, the critical band is the band of audio frequencies within which a second tone will interfere with the perception of the first tone by auditory masking (beating, auditory roughness).
- They are non-linear, level-dependent and the bandwidth decreases from the base to apex of the cochlea as the tuning on the basilar membrane changes from high to low frequency.

Mel Scale

MIDI-Frequency Conversion

$$f(n) = 440 \times 2^{(n-69)/12}$$

Mel-frequency cepstrum

$$f_{mel} = 1000 \log(1 + \frac{f}{1000}) / \log 2$$

Mel-Spectrogram

Loudness

Sound Intensity (watts/m^2)

$$I = \frac{P}{4\pi r^2}$$

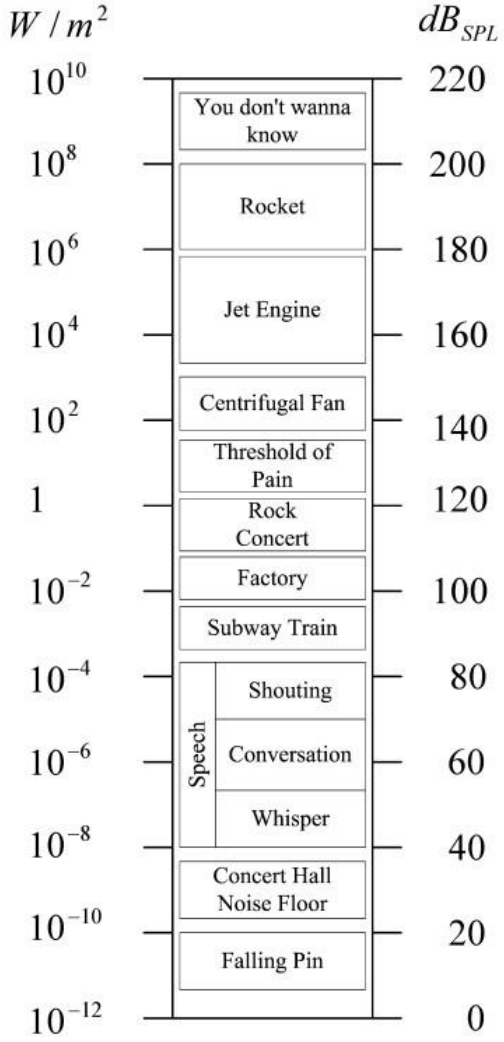
Sound Intensity Level (dB)

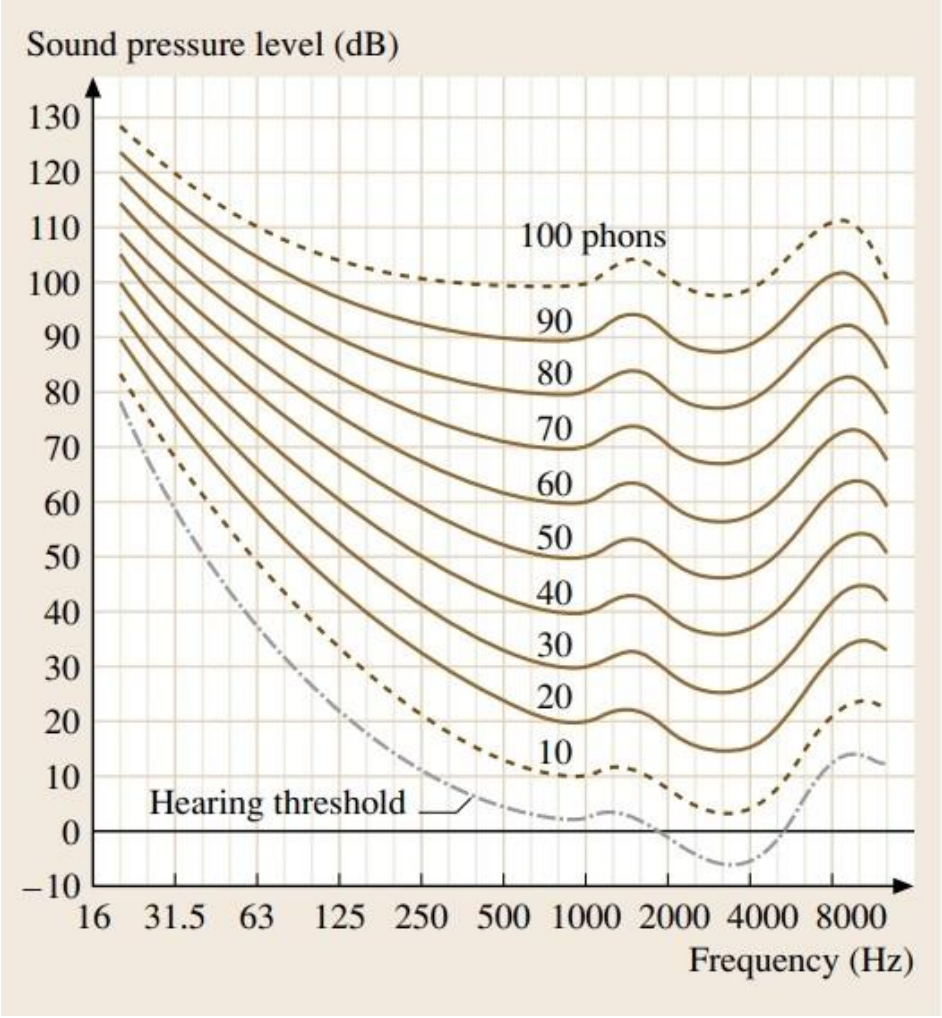
$$dB_{SIL} = 10 \log_{10}(I/I_0)$$

I_0 is the threshold of hearing. Sound pressure, aka. amplitude A has $I \propto A^2$

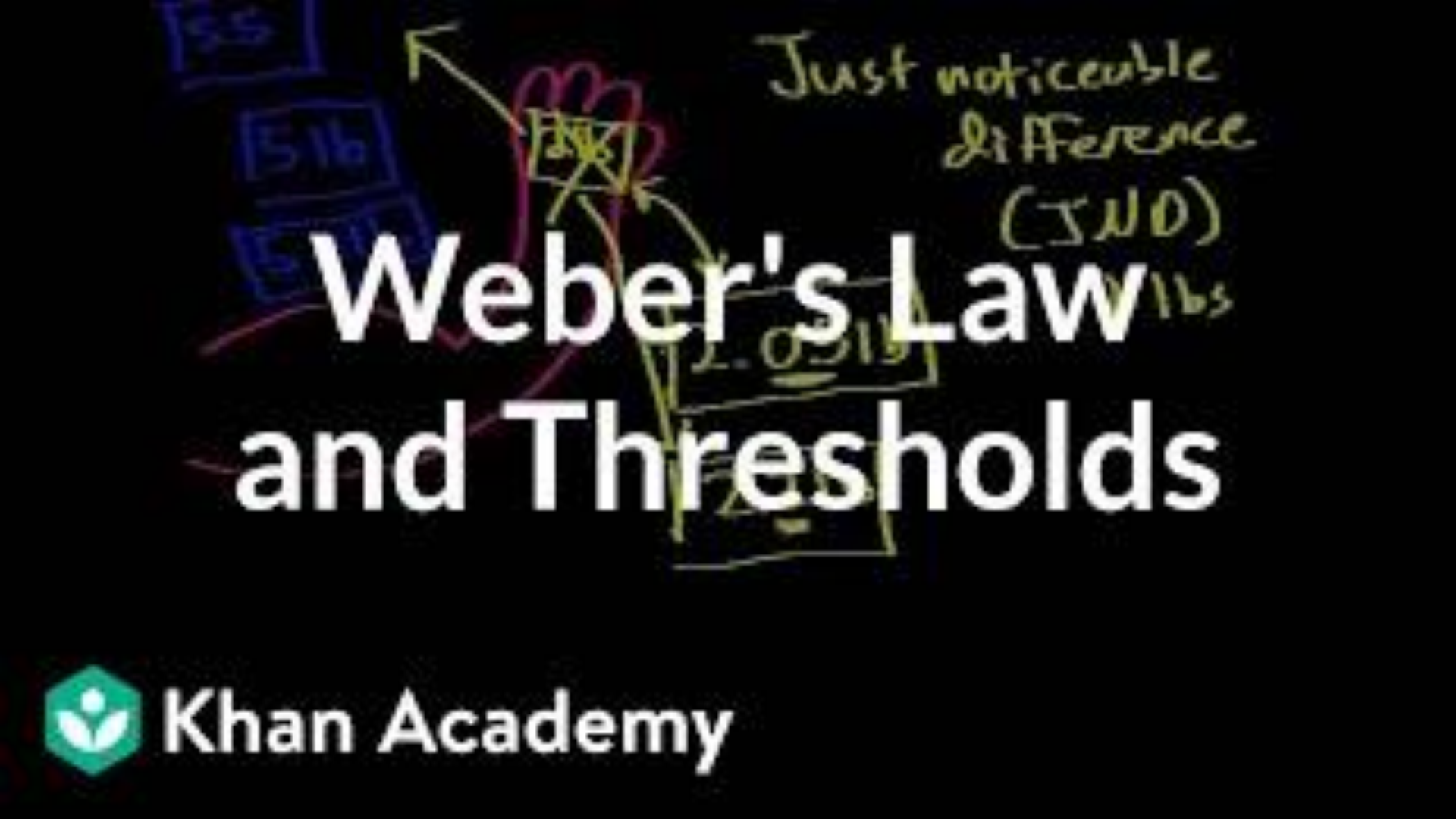
Sound Pressure Level (dB)

$$dB_{SPL} = 20 \log_{10}(A/A_0)$$



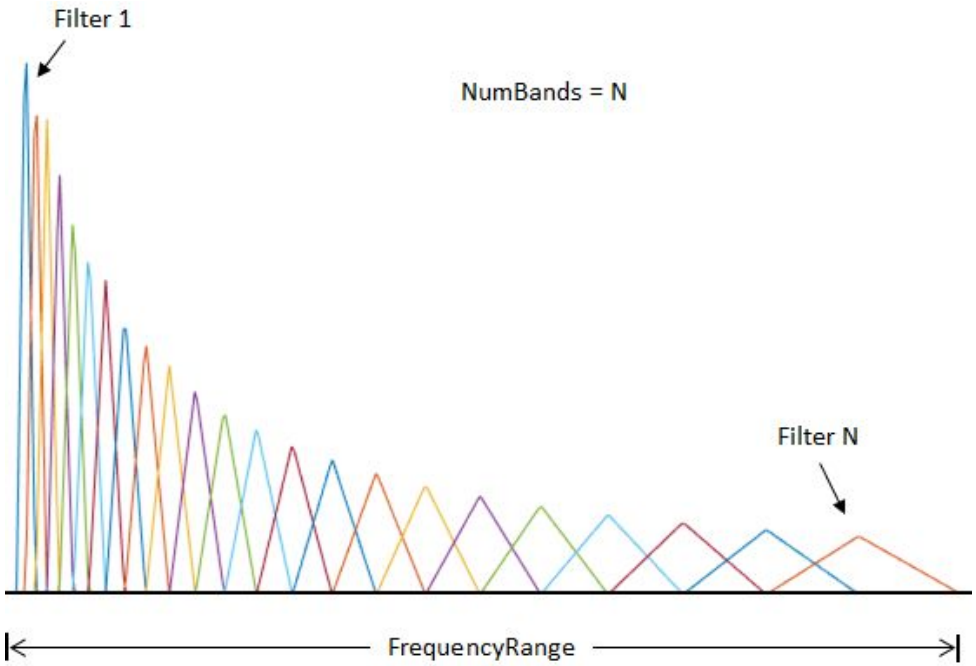


Equal Loudness Curves


Just noticeable difference (JND)
Weber's Law and Thresholds
Khan Academy

Perceptual Resolutions

- The **amplitude resolution** of the ear is generally taken to be about 0.25 dB under best-case conditions, although for some situations it is considered to be slightly larger, on the order of 0.5–1.0 dB.
- **Frequency resolution**: auditory filters; critical band.



Math & Musical Tuning

$$\sqrt[4]{5} \quad \frac{3}{2}$$

$$\sqrt[12]{2} \quad \frac{\sqrt{5}}{2}$$

Auditory Scene Analysis

- As discussed earlier, the peripheral auditory system acts as a frequency analyzer, separating the different frequency components in a complex sound. Somewhere in the brain, the internal representations of these frequency components have to be assigned to their appropriate sources.
- If the input comes from two sources, A and B, then the frequency components must be split into two groups; the components emanating from source A should be assigned to one source and the components emanating from source B should be assigned to another. The process of doing this is often called *perceptual grouping*.
- The process of separating the elements arising from two or more different sources is sometimes called *segregation*.

Auditory Objects

- An auditory object can be defined as the percept of a group of successive and/or simultaneous sound elements as a coherent whole, appearing to emanate from a single source.
- *Simultaneous grouping*: the grouping together of all the simultaneous frequency components that emanate from a single source at a given moment.
- *Sequential grouping*: the connecting over time of the changing frequencies that a single source produces from one moment to the next.

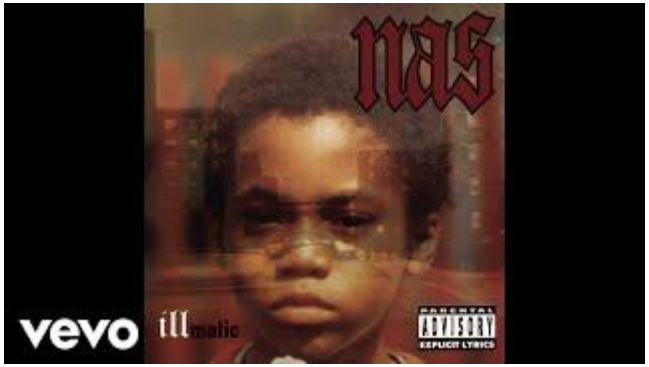
Information Used to Separate Auditory Objects

- Fundamental Frequency and Spectral Regularity
- Onset and Offset Disparities
- Contrast with Previous Sounds
- Correlated Changes in Amplitude or Frequency
- Sound Location

Changes in Auditory Stimuli

- Unrelated auditory streams rarely start or stop at exactly the same time.
- The features of a single auditory stream (frequency, amplitude, and timbre) tend to change slowly and gradually over time. An abrupt change often signals a new auditory stream.
- All the frequency components of a single auditory stream tend to change in the same way at the same time (e.g. by growing louder as the sound source approaches the listener).

Source Separation (Computational Auditory Scene Analysis)



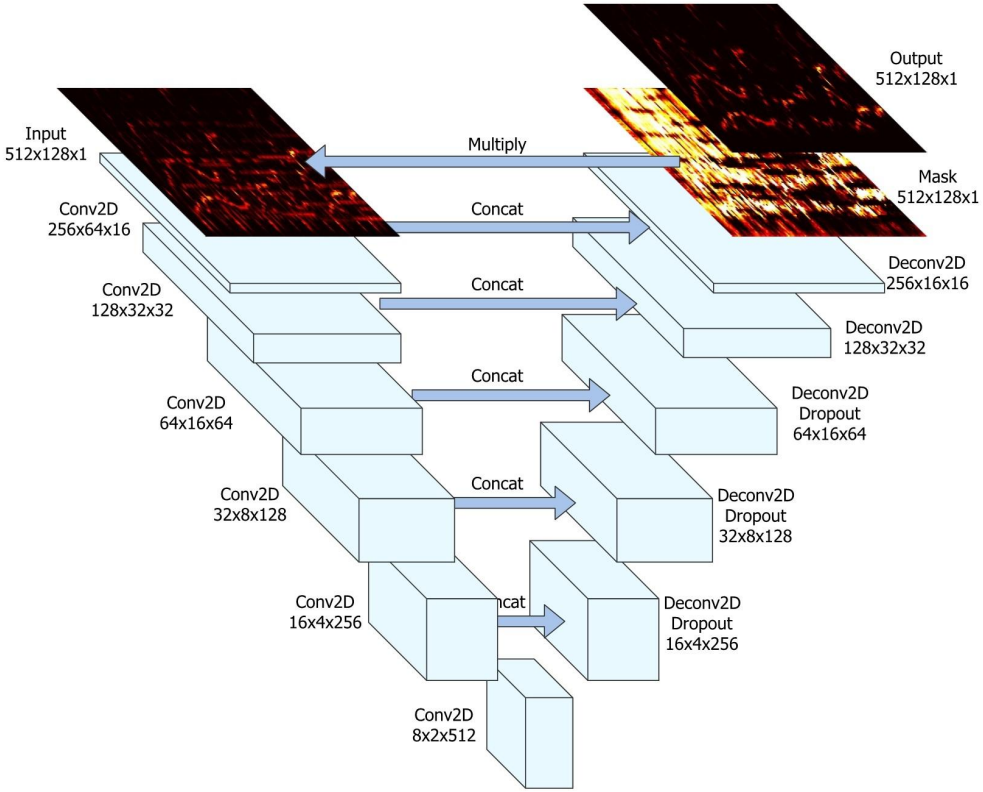
Vocal

Accompaniment (Beat)

Drums

Bass

Other



Deep U-net Convolutional Neural Network

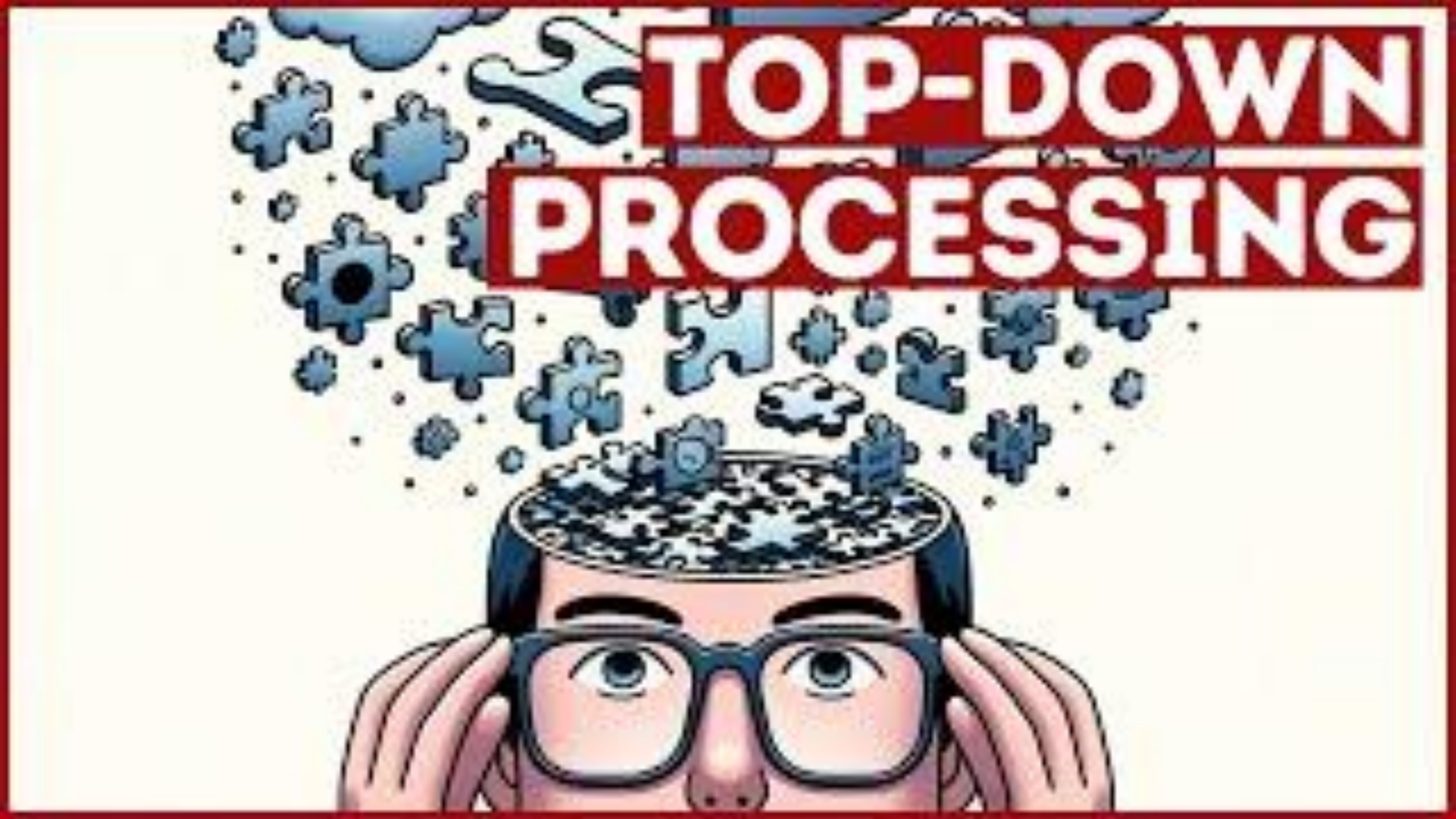
Brain, Perception, and Music


BT
Your Brain On Music


Broadbent's Early Selection Theory
Theories of selective attention
Khan Academy

Bottom-up vs. top-down processing

Khan Academy


TOP-DOWN
PROCESSING

Auditory Illusions