Multimedia Systems

II. Introduction to Sound

2.2. The Auditory System

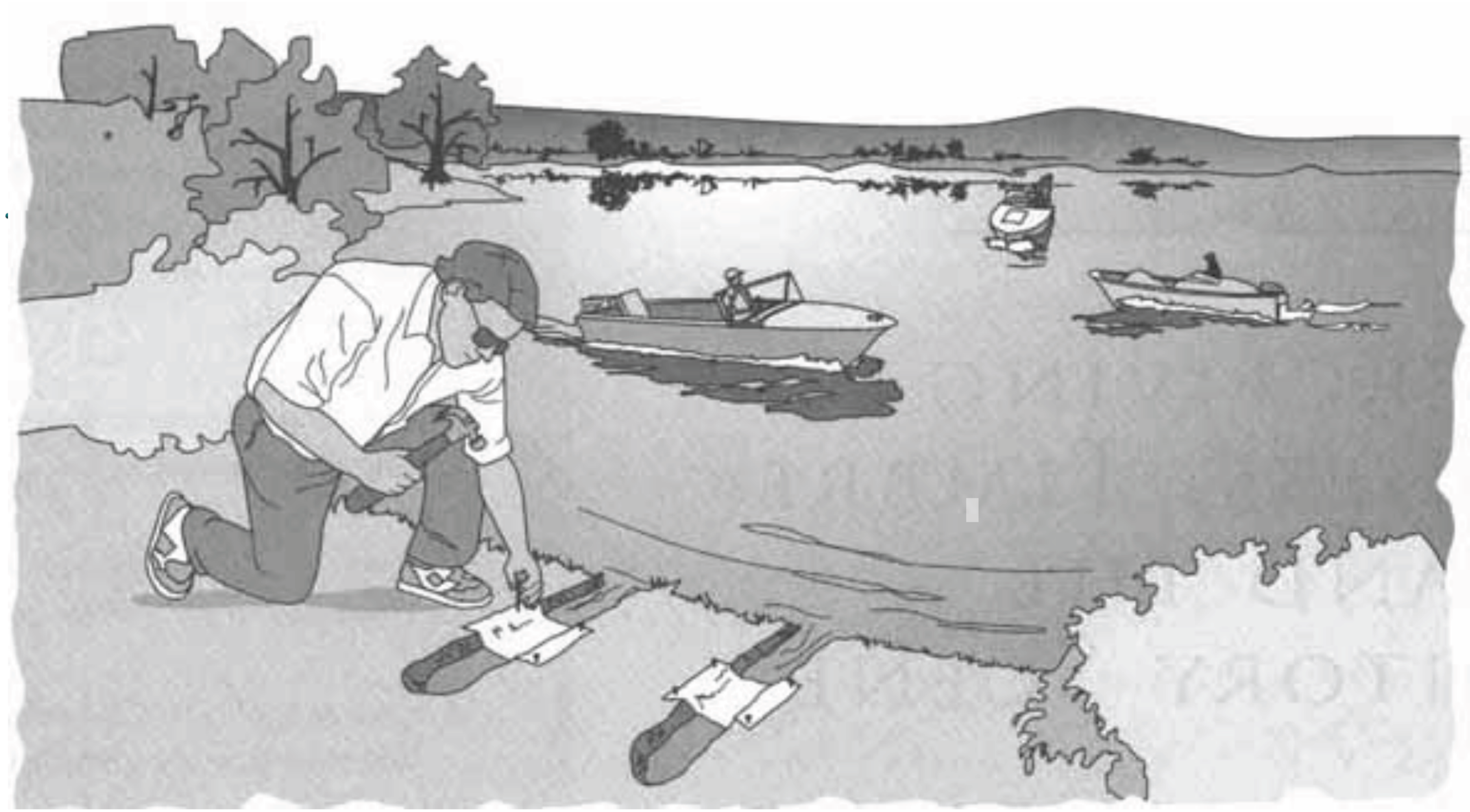
António Sá Pinto

FEUP – 2019/2020

Agenda

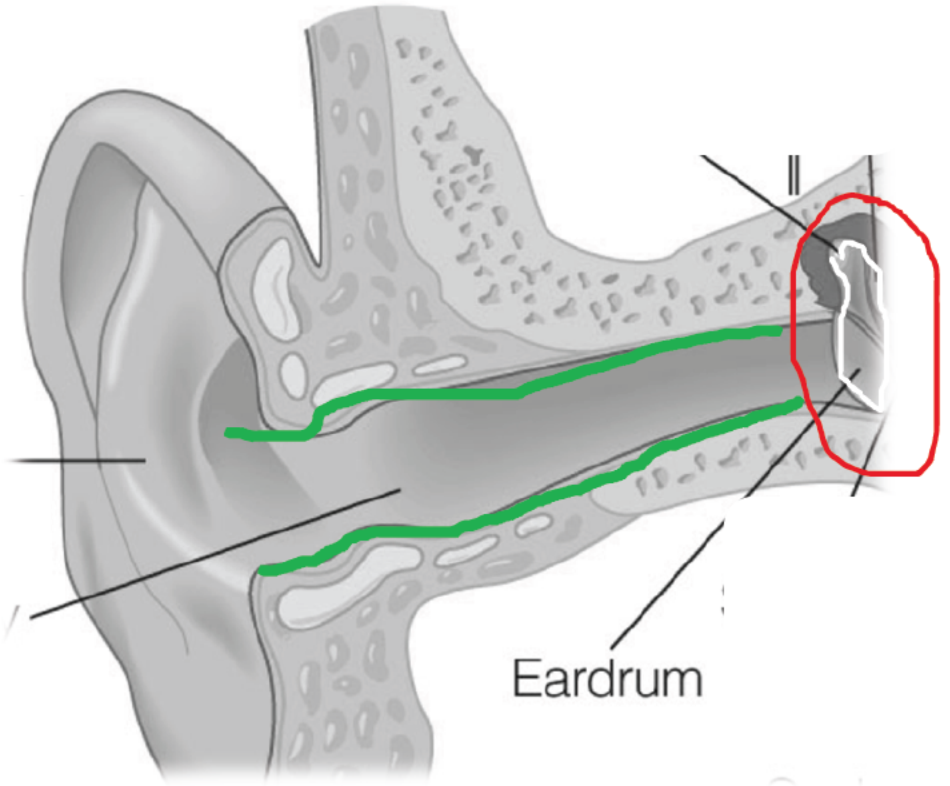
- Physiology of the Auditory System
- Psychoacoustics
- Sound Perception
- Human Auditory Specifications
- Subjective Evaluation of Sound Perception

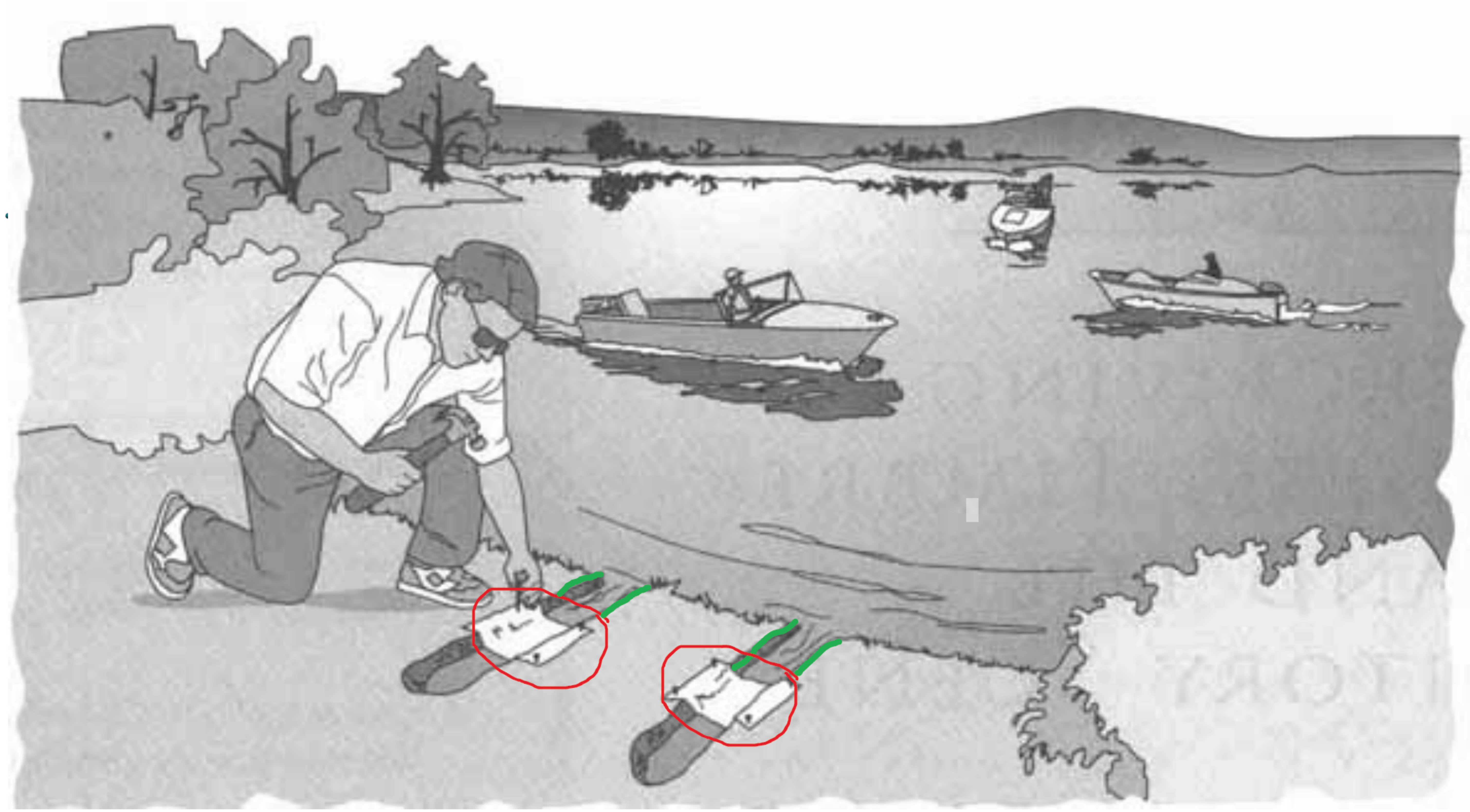
The Auditory System



The Lake Analogy from: (Bregman A.S.) Auditory Scene Analysis - The Perceptual Organization of Sound. *MIT Press, 1990*

The Auditory System



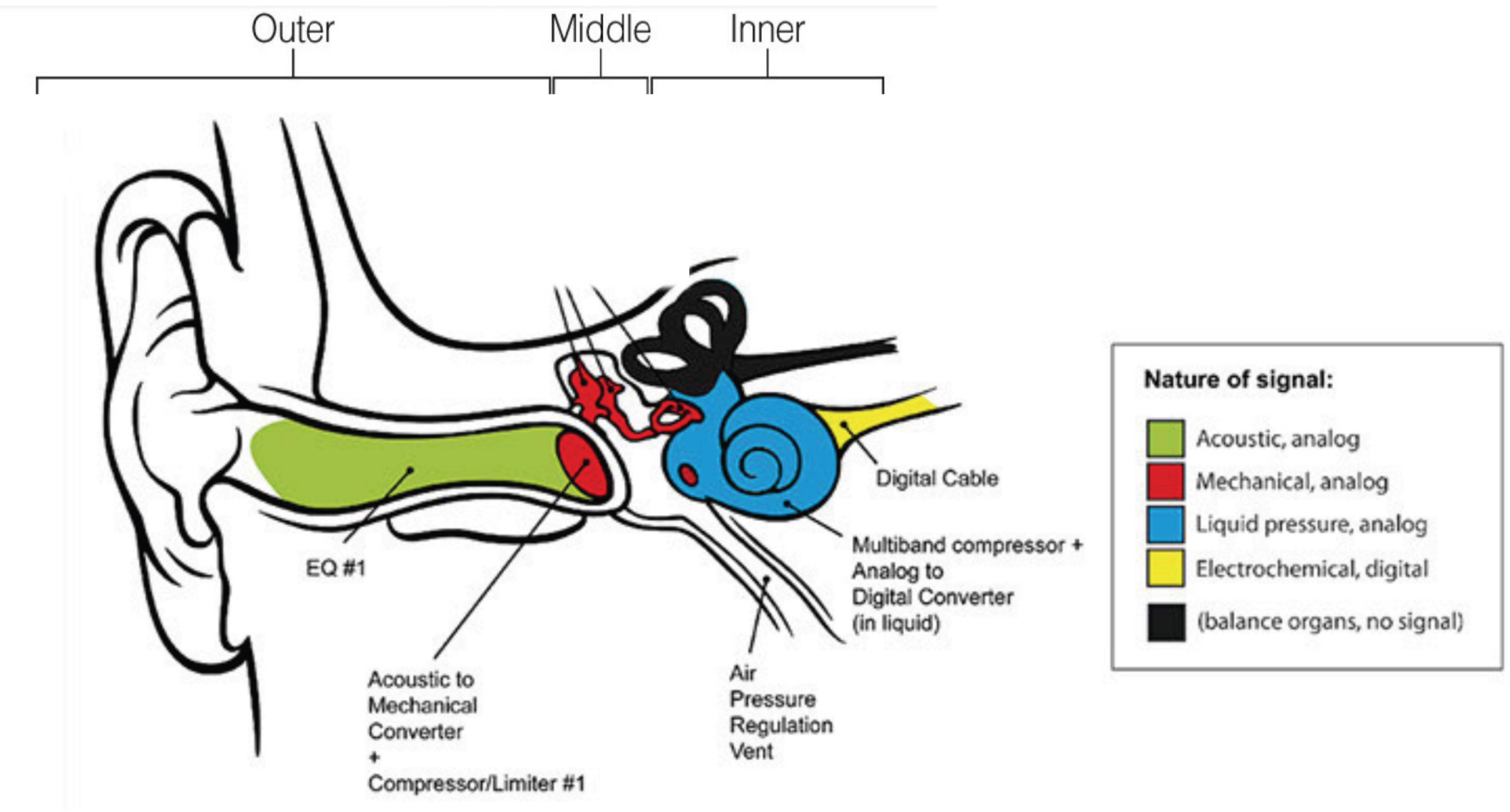


The Lake Analogy from: (Bregman A.S.) Auditory Scene Analysis - The Perceptual Organization of Sound. *MIT Press, 1990*

The Auditory System

Peripheral Auditory System

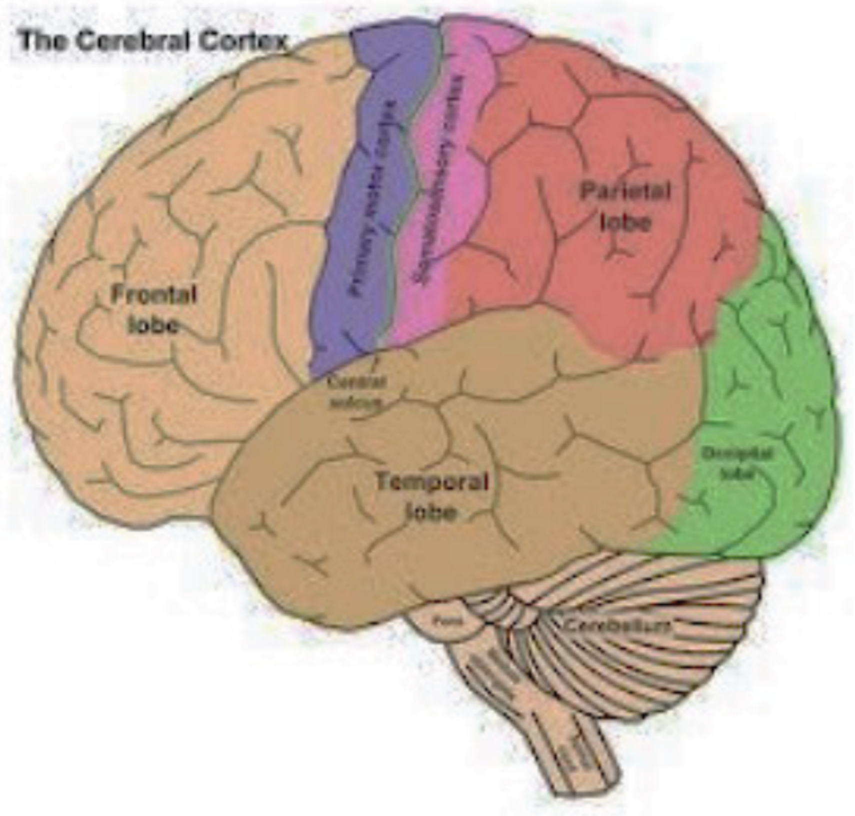
- Outer ear
- Middle ear
- Inner ear



from: https://www.uaudio.com/blog/how-the-ear-works/

Central

- Auditory nerve→Brain



from: Wikipedia

Auditory Transduction

https://www.youtube.com/watch?v=46aNGGNPm7s

Psychoacoustics

The way sounds are perceived. (Not to be confused with the physics of sound - acoustics).

It is very relevant to:

- Audio codification;
- Audio compression;
- Music perception;
- 3D Sound;
- Noise measurements/cancellation;
- Hearing aids;
- Etc.

Sound Perception

Definition, Scales of Measurement and Physiological Basis

- Loudness
- Pitch
- Timbre
- Localisation

Loudness

Definition

That attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud. *(ANSI)*

The (perceived) **loudness** of an acoustic sound is related to its **intensity** (physical property), but depends on the frequency and distribution of spectral energy.

Loudness

Scales

dB

- The quietest sound that most people can hear has a sound pressure of $2x10^{-5}$ Pa, so this pressure is called the **threshold of hearing**. For convenience, in practice, SPL is usually represented in decibels (dB) relative to the threshold of hearing P_r.

$$dB(SPL) = 20\log\left(\frac{P_m}{P_r}\right), \text{ with } P_m\text{: measured SPL, } P_r\text{: reference SPL}$$

Phon

Sone

(Weighted) dB-A, dB-C, dB-Z

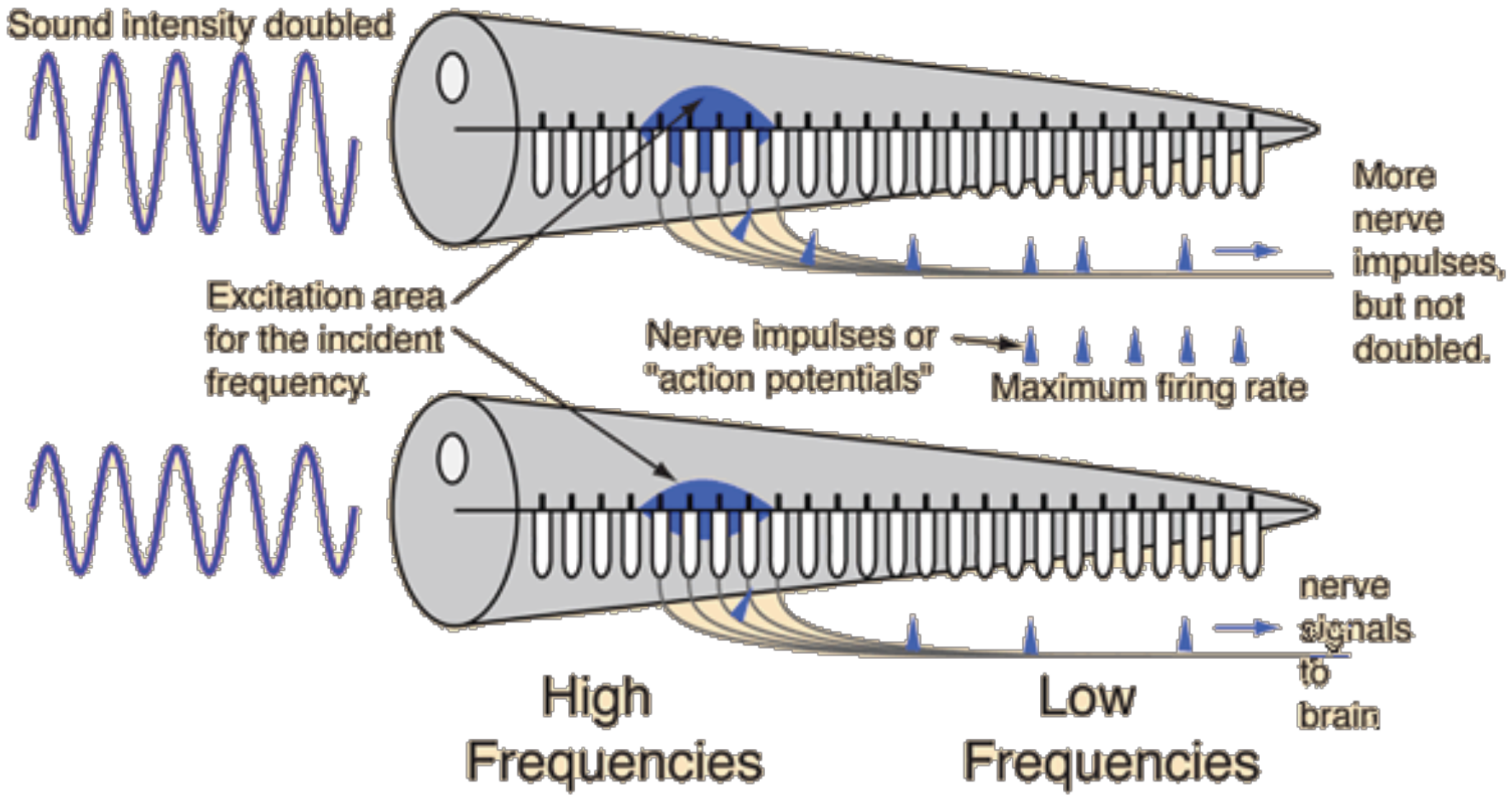
Loudness

Theory of Perception

Loudness is a measure of the subjective response to the amplitude of vibration of hair cells in the basilar membrane.

At <u>low</u> intensity levels:
- Increased intensity = increased firing rate

At <u>high</u> intensity levels:
- Information is carried by several nerve fibers



from: http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/soucon.html

Pitch

Definition

The attribute of auditory sensation in terms of which sounds may be ordered on a musical scale extending from low to high. *(ANSI)*

- is a "subjective" attribute, and cannot be measured directly. Therefore, the measurement of pitch requires a human listener (the "subject") to make a perceptual judgement.
- In general, sounds having a periodic acoustic pressure variation with time are perceived as pitched sounds, for non-periodic acoustic pressure waveform, as non-pitched sounds.

The perceived pitch is related to the repetition rate of the waveform of a sound, therefore it (generally) corresponds to the frequency of a pure tone and the fundamental frequency of a complex tone.

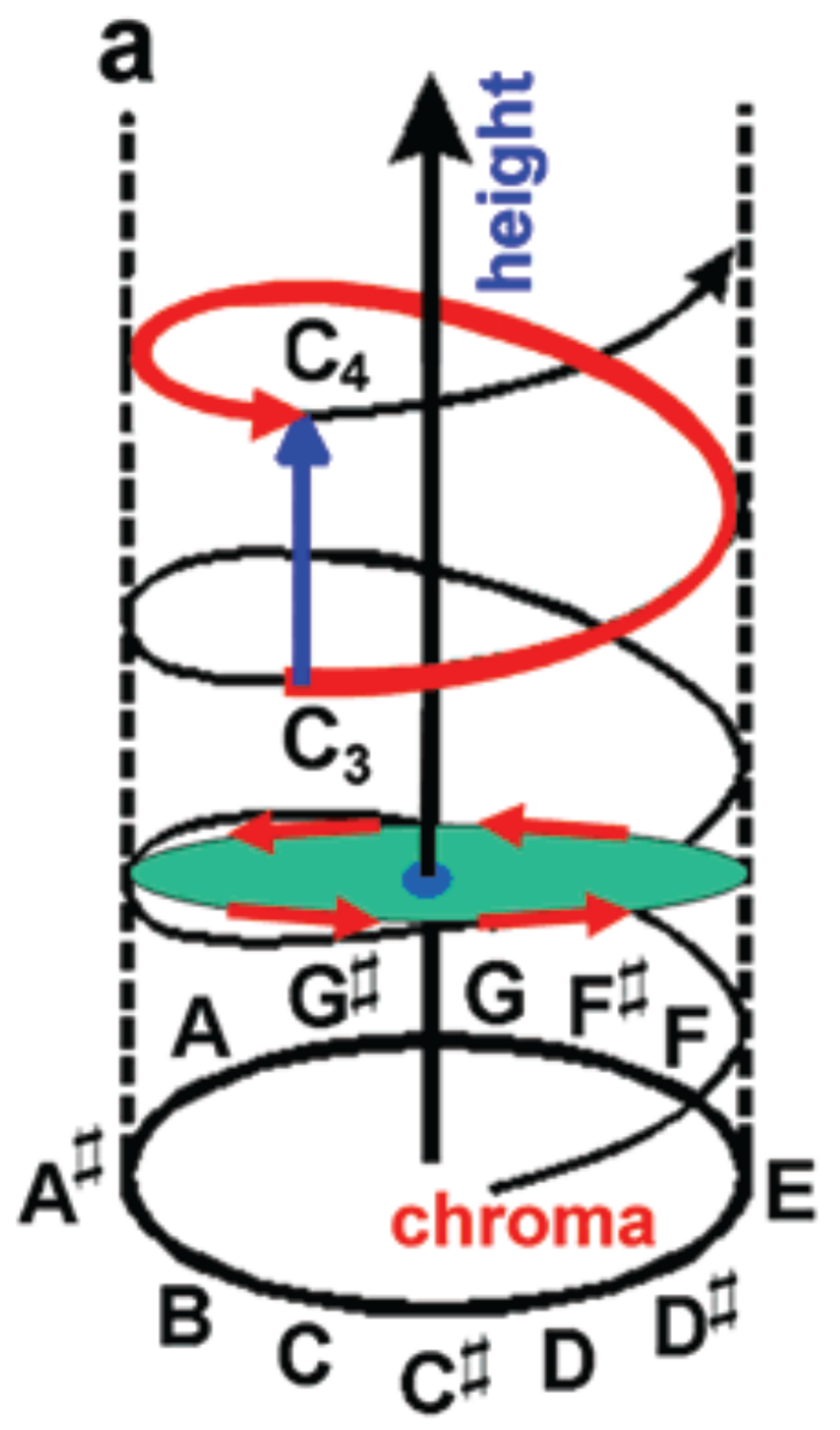
Pitch Helix

The Helical Model

The pitch helix is a pitch space where linear pitch is wrapped around a cylinder, thus modelling the special relationship that exists between octave intervals. Shepard (1964)

The model is a function of 2-dimensions:

- **Height**: naturally organizes pitches from low to high;
- **Chroma** (Class): represents the inherent circularity of pitch;



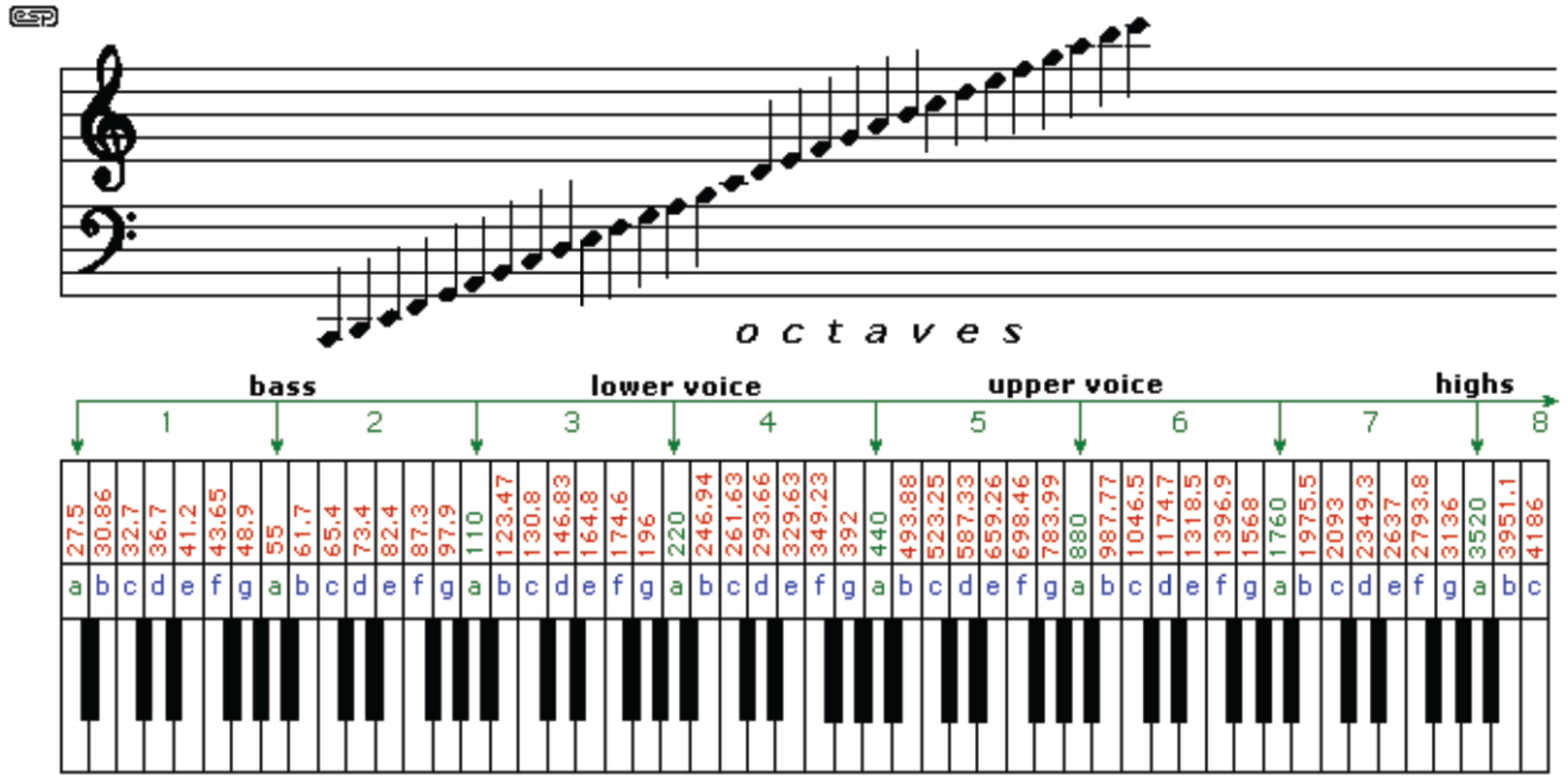
This model accounts for the perceived similarity of pitches that are separated by octaves: A1-A2-A3; C2-C3-C4.
In Western notation, musical notes are identified first by their position in the octave (**chroma**) and then by the octave in which they are placed (e.g. G4, F#3).

Shepard

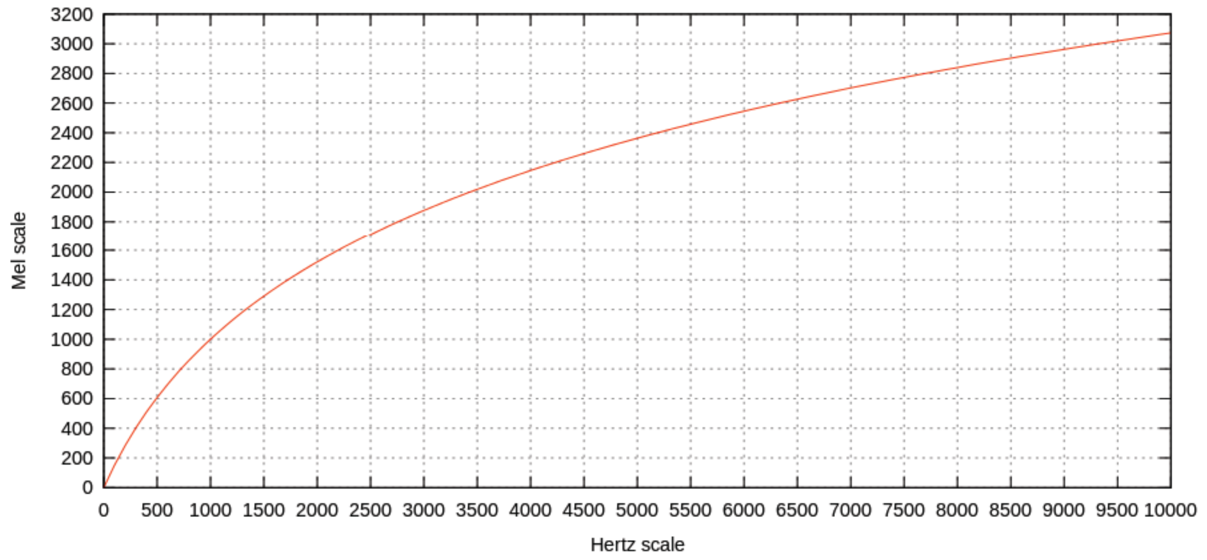
Pitch

Scales

Musical Scales capture both pitch height and pitch chroma, but they differ across cultures.



Mel scale is is a perceptual scale of pitches judged by listeners to be equal in distance from one another. It is “universal”, but doesn’t capture pitch chroma. (In practice, it is rarely used)

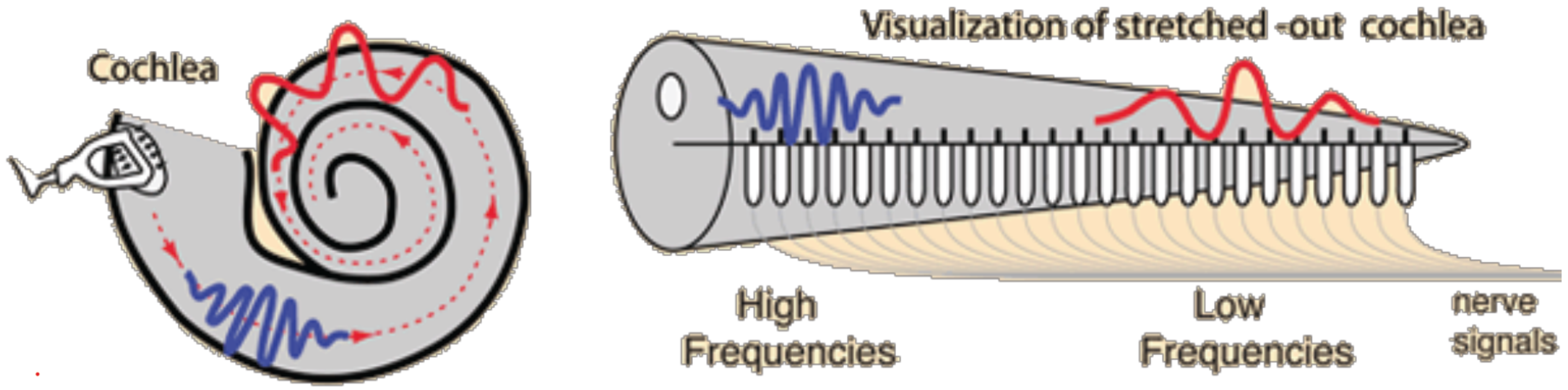


Pitch

Theory of Perception

‘Place’ theory

- Spectral analysis is performed on the stimulus in the inner ear, different frequency components of the input sound excite different places or positions along the basilar membrane, and hence neurones with different centre frequencies.



from: http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/soucon.html

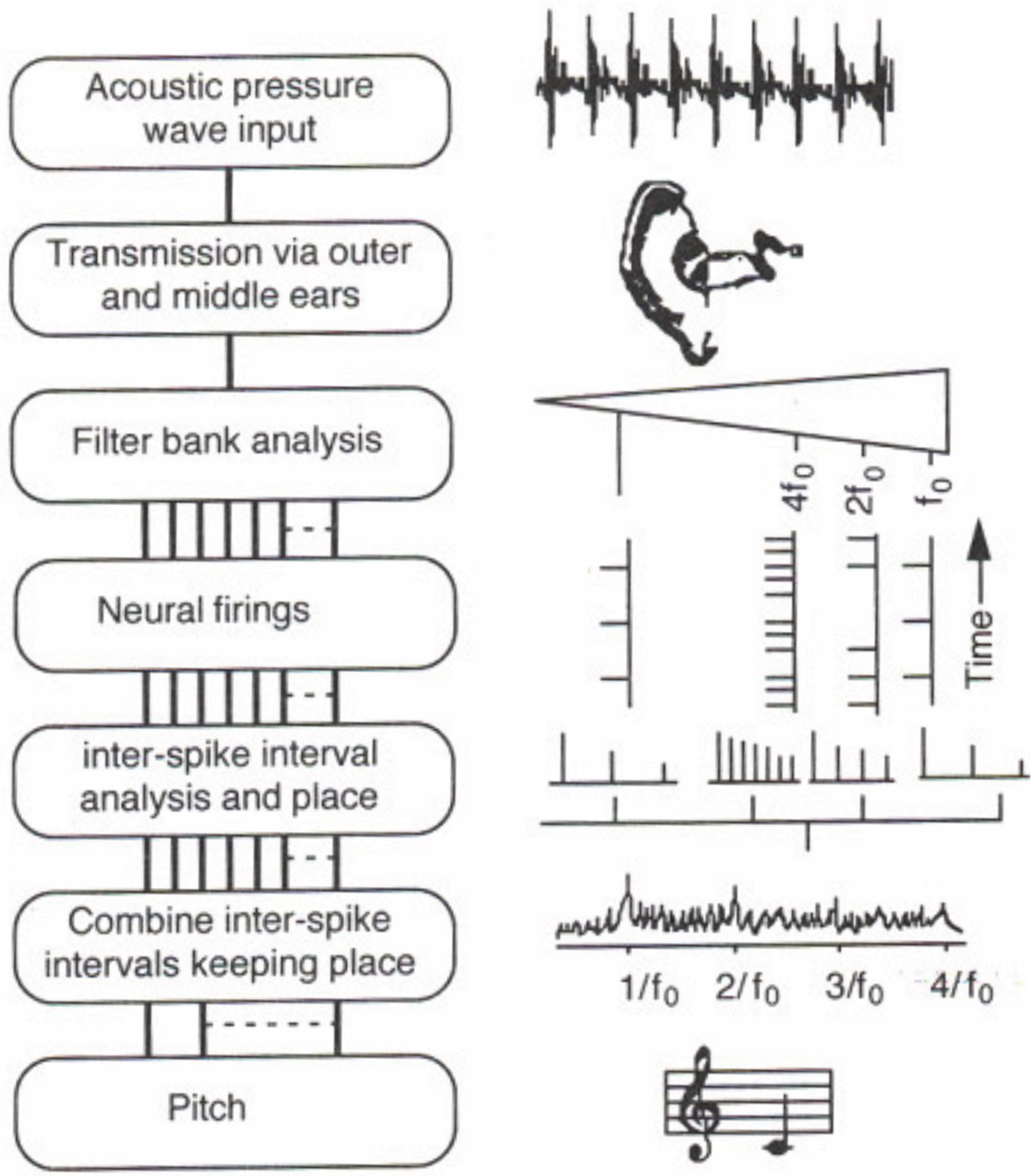
‘Temporal’ theory

- Pitch corresponds to the time pattern of the neural impulses evoked by that stimulus. Nerve firings tend to occur at a particular phase of the stimulating waveform, and thus the intervals between successive neural impulses approximate integral multiples of the period of the stimulating waveform.

Pitch

Theory of Perception

'Contemporary' theory



from: (Moore) http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/soucon.html

Timbre

Definition

The attribute of sensation in terms of which a listener can judge that two sounds have the same loudness and pitch are dissimilar. *(ANSI)*

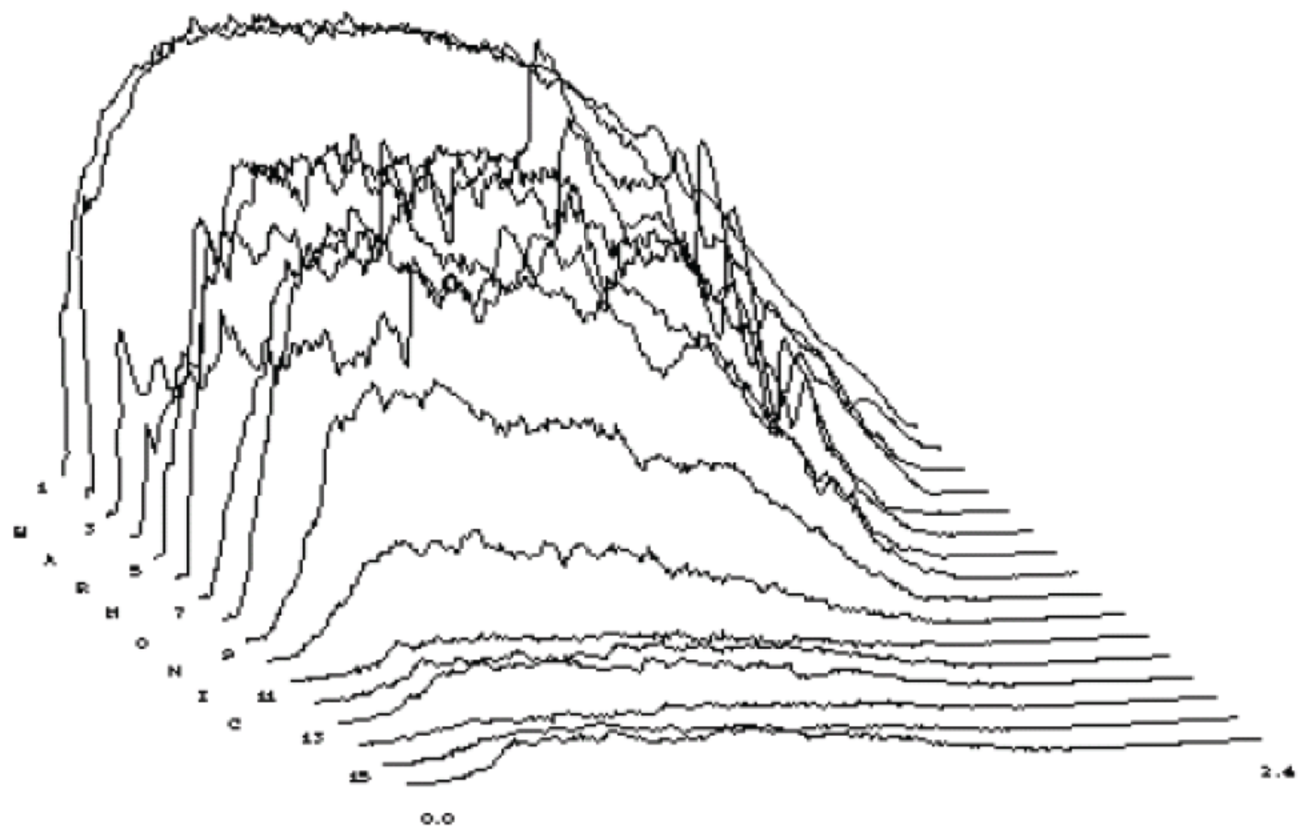
- In other words, everything that is not loudness or pitch! *...the psychoacoustician's multidimensional waste-basket category for everything that cannot be labeled pitch or loudness."* (McAdams and Bregman 1979)

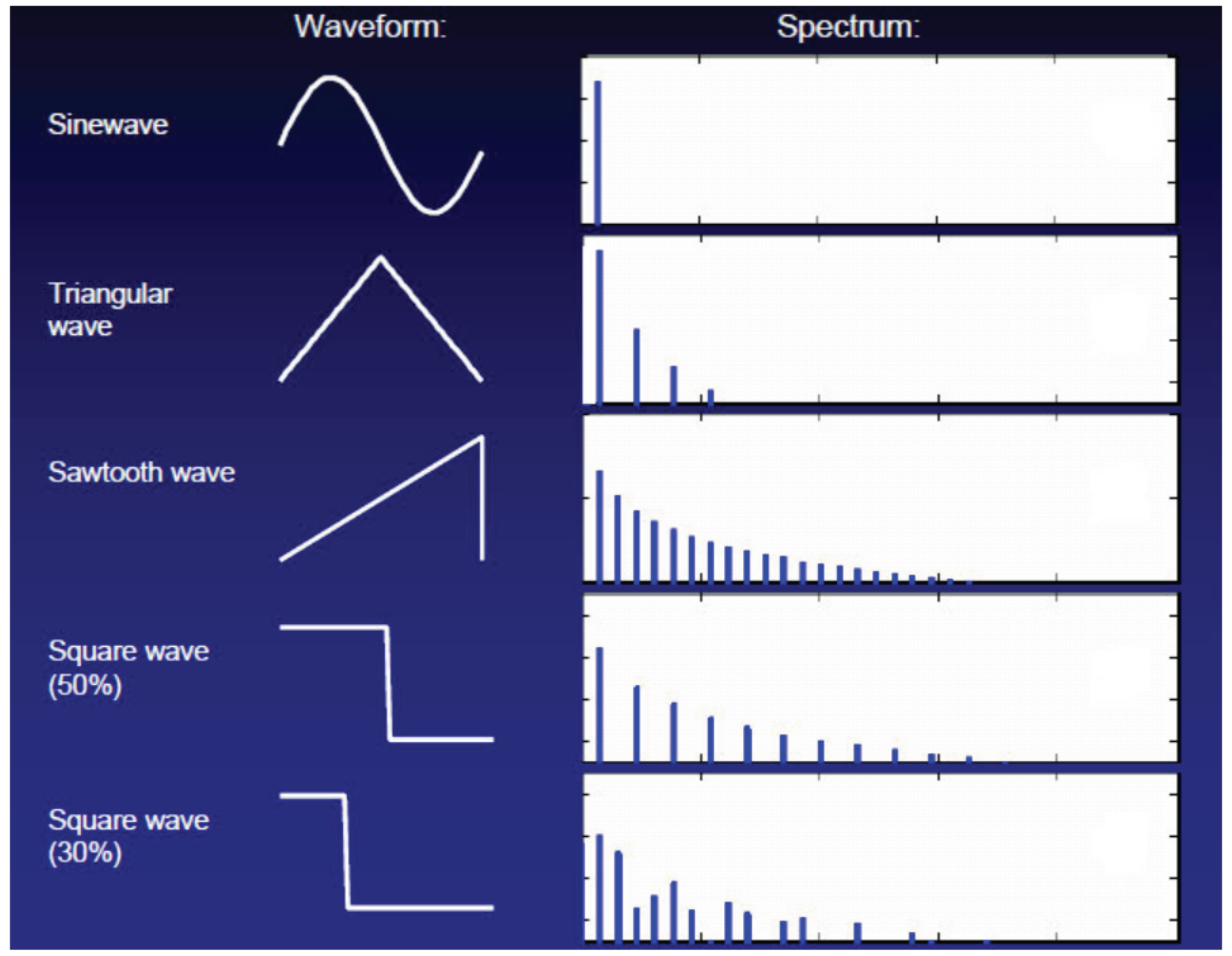
- Musically, it is essential to distinguish different types of musical instruments, e.g. to distinguish a piano playing a 220Hz note and a trumpet playing the same 220Hz note, at equal loudness.

Timbre

Definition

Timbre relates to: static spectrum; spectral envelope; time envelope (ADSR); dynamic spectrum (time-evolving); phase; etc,…





- Not ”measurable”, given its perceptual nature;
- Not even ”estimated”, given its multidimensional nature;

Localisation

Definition

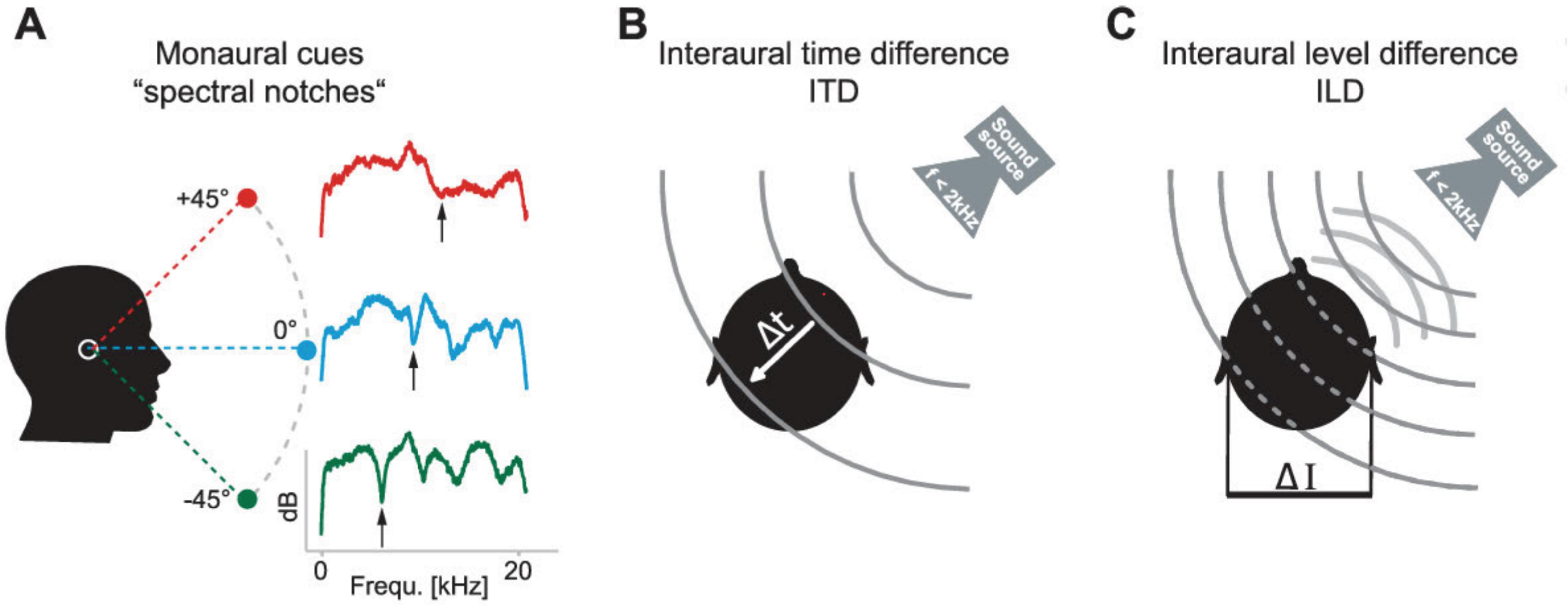
Sound localisation refers to judgements of the direction and distance of a sound source.

Humans localise sound sources based on the monaural cues (derived from a single ear, i.e. single channel sound), and/or the binaural cues (obtained from both ears).
The monaural cues come from the interaction between the sound source and the human auditory system, where the temporal/spectral shape of the original sound source is modified by the auditory before it reaches the human brain.
The binaural cues include the interaural time difference (ITD) and interaural intensity difference (IID).

Localisation

Theory of Perception

- Spectral Information
- Interaural time difference (ITD);
- Interaural intensity difference (IID);



from: Grothe et al. 2010 Mechanisms of Sound Localization in Mammals

Auditory Specifications

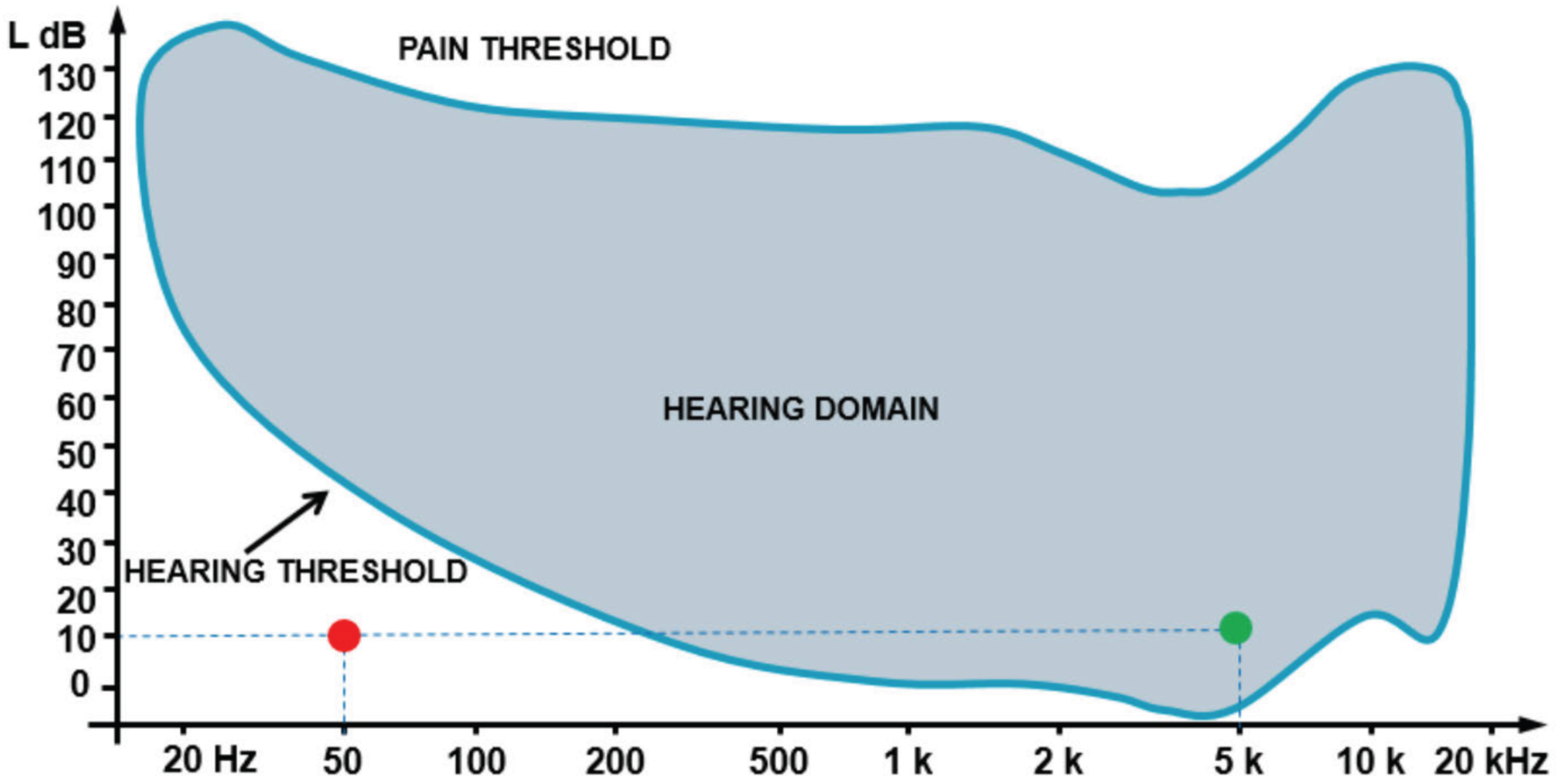
Frequency range
20Hz to 20,000Hz (textbook); 50Hz to 13,000 Hz (effective); most psychophysics is done between 100-5000Hz (range where one obtains interpretable data).

Intensity range
Extends over many orders of magnitude (depending on frequency); at the sweet spot (~1000Hz-3000Hz) about a 120dB dynamic range.

Sensitivity
Just Noticeable Difference (JND) for frequency: ~0.2%
JNDs for loudness discrimination: ~1dB
JNDs for Timing differences:
- few microseconds in spatial hearing (JND for azimuthal ~1deg)
- 2 msec (gap thresholds);
- 25 miliseconds (order threshold);

Auditory Specifications

Hearing Range

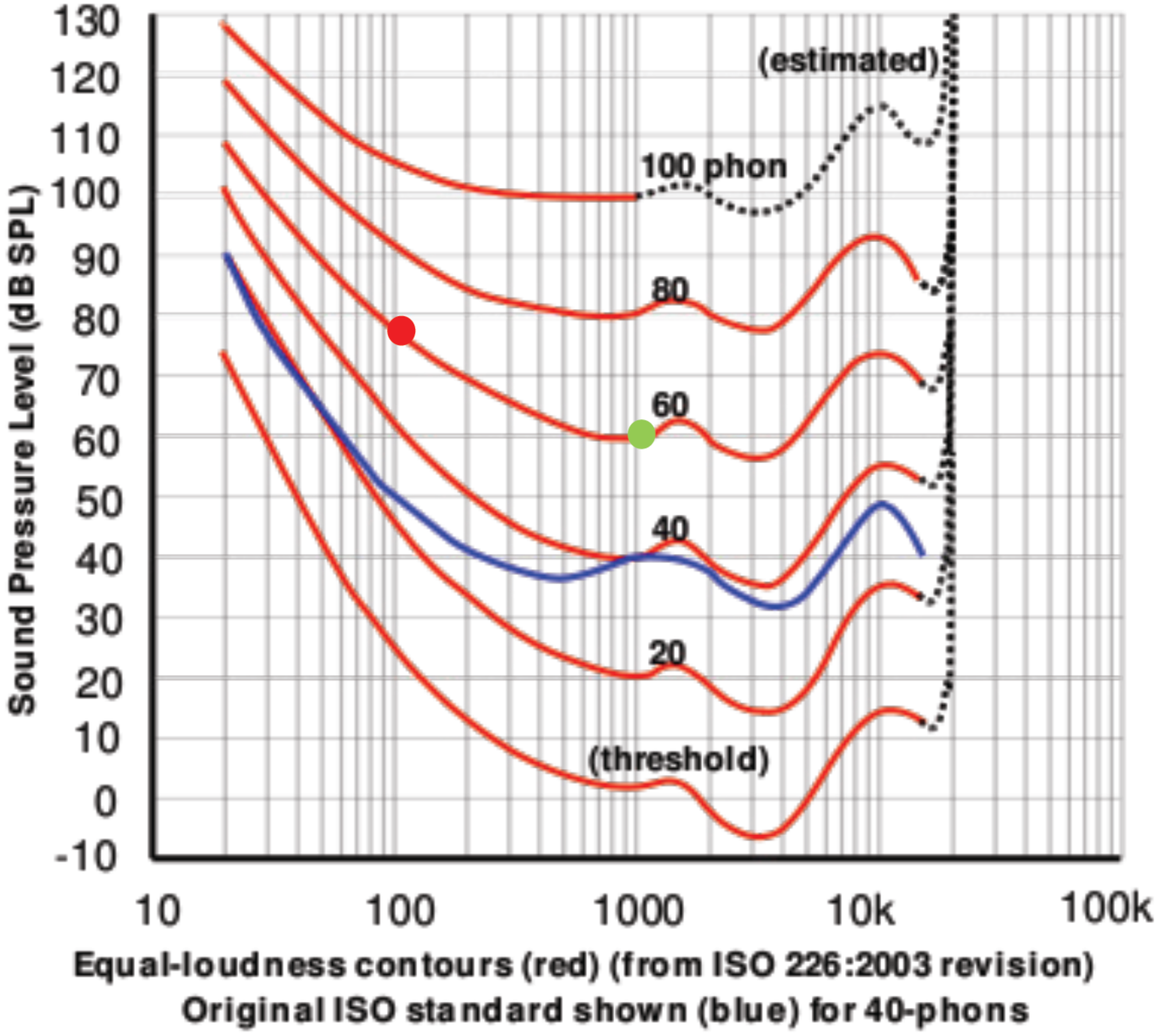


from: https://community.sw.siemens.com/s/article/sound-quality-metrics-loudness-and-sones

A tone at 50Hz and 10dB is inaudible. A tone at 5000Hz and 10dB is audible.

Auditory Specifications

Equal Loudness Curves (Fletcher-Munson 1933)



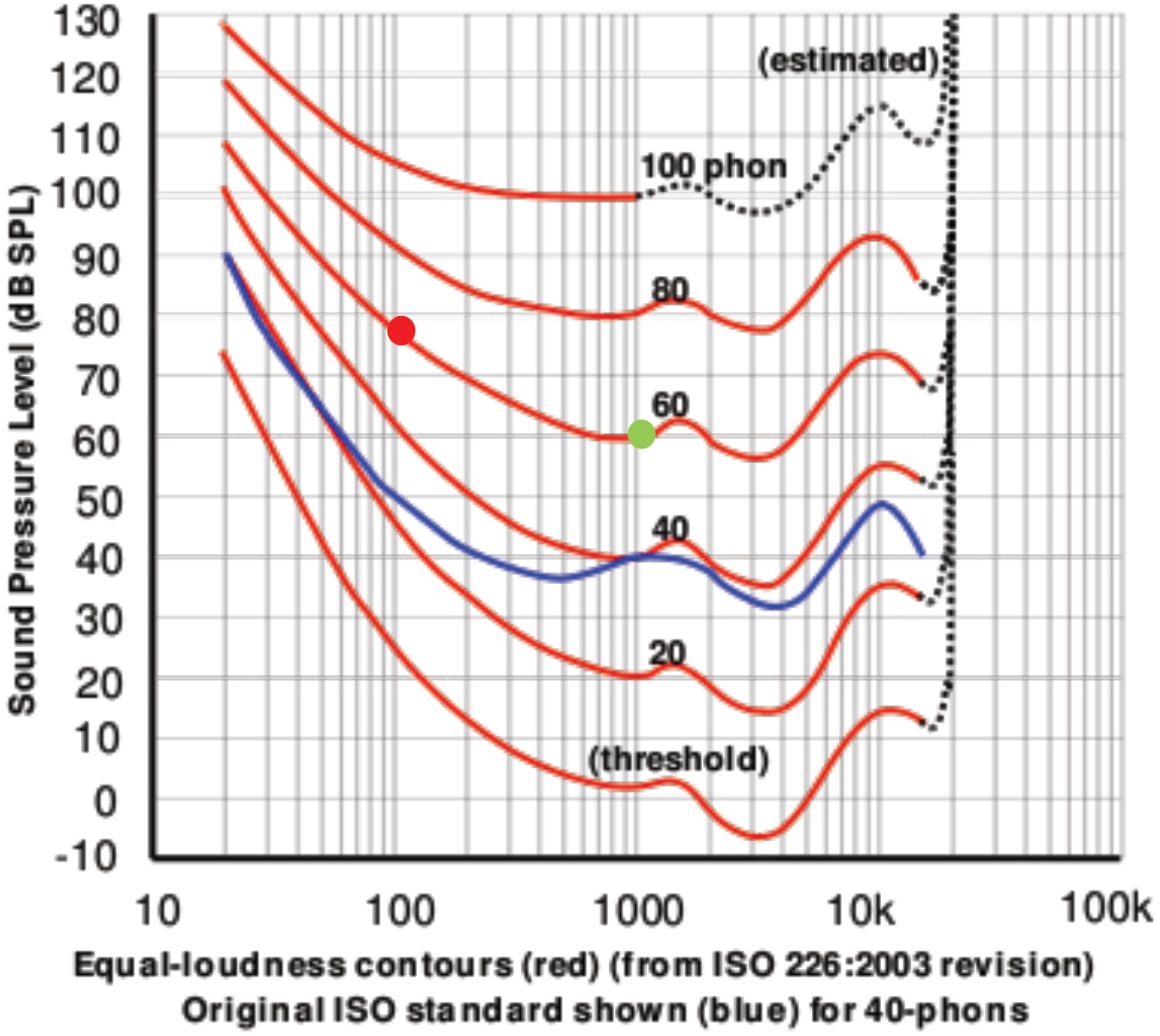
from: Wikipedia

A tone at 100Hz and 78dB has the same (perceived) loudness as a tone at 1000Hz and 60dB.

Auditory Specifications

Scales of Loudness

dB
Phon
Sone
(Weighted) dB-A, dB-C, dB-Z



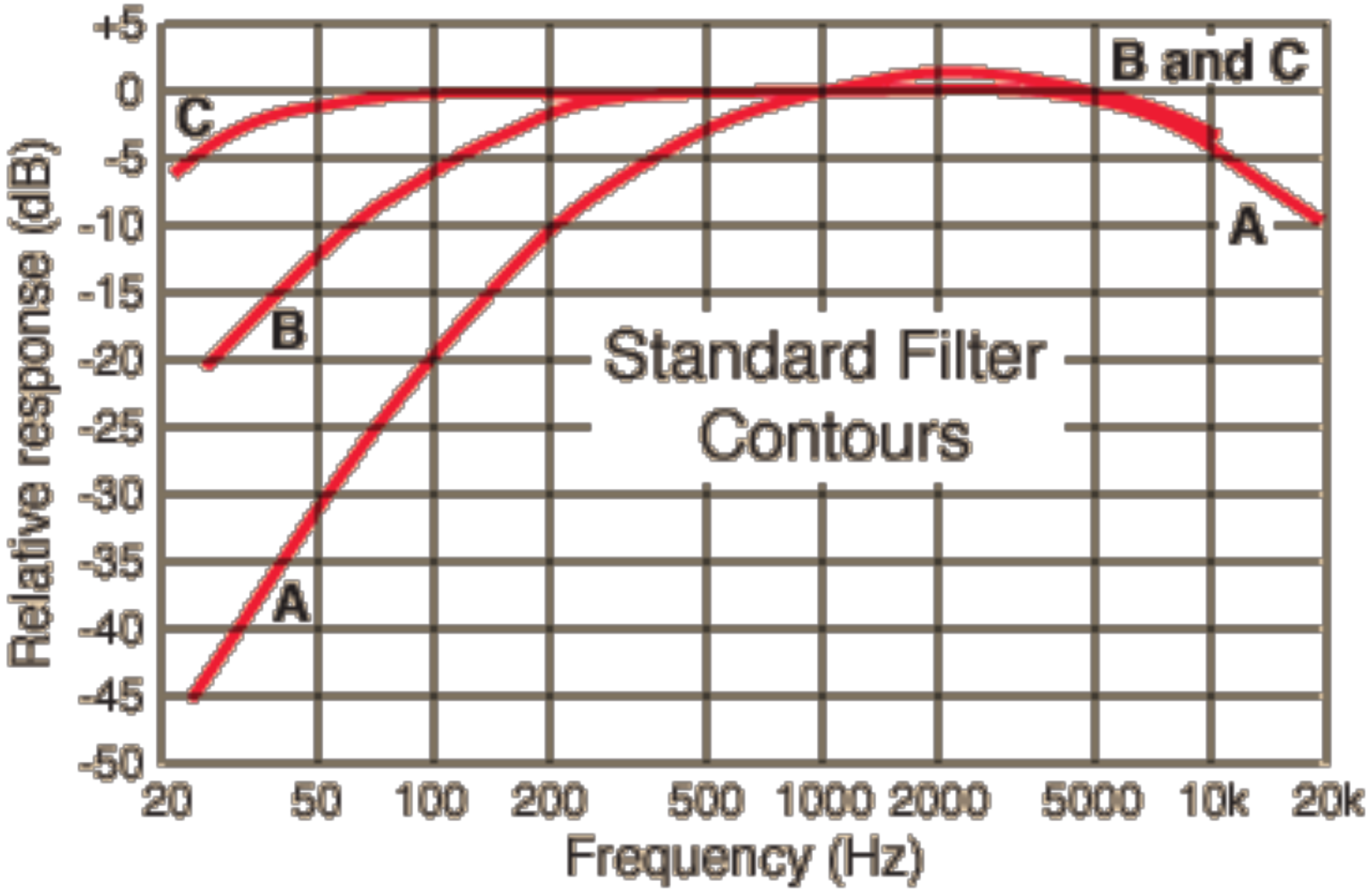
from: Wikipedia

Auditory Specifications

Scales of Loudness

dB
Phon
Sone
(Weighted) dB-A, dB-C, dB-Z

- The unit dB implies that all frequencies in the audible frequency range are treated equally - something very different from what the ear does.
- Standard filter contours are used to make the instrument more nearly approximate the normal human ear. The different contours were intended to match the ear at different sound intensities.
- Measurements of the sound level environmental/industrial sound levels.

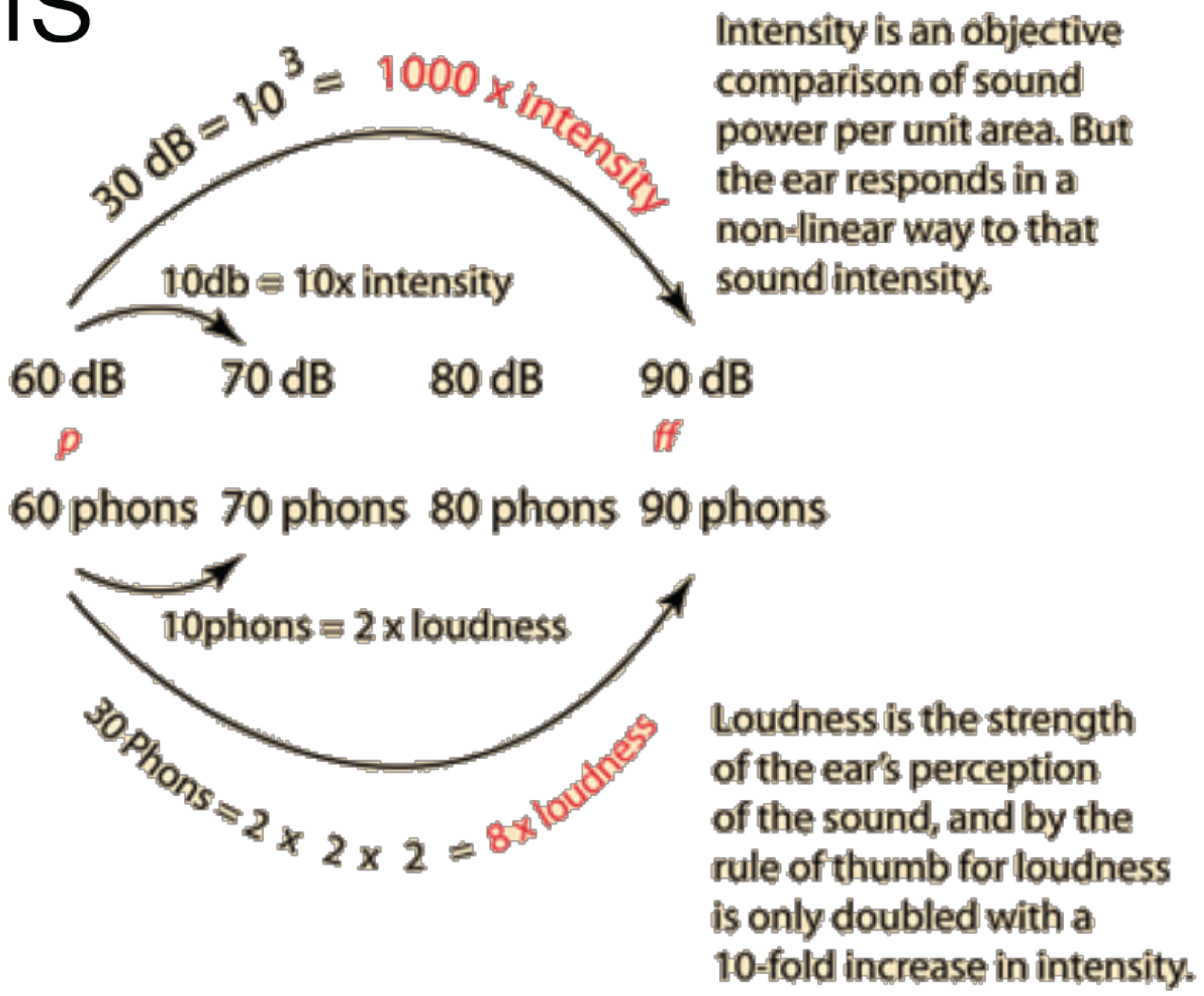


from: Wikipedia

Auditory Specifications

Scales of Loudness

dB
Phon
Sone
(Weighted) dB-A, dB-C, dB-Z



from: http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/

sone	1	2	4	8	16	32	64	128	256	512	1024
phon	40	50	60	70	80	90	100	110	120	130	140

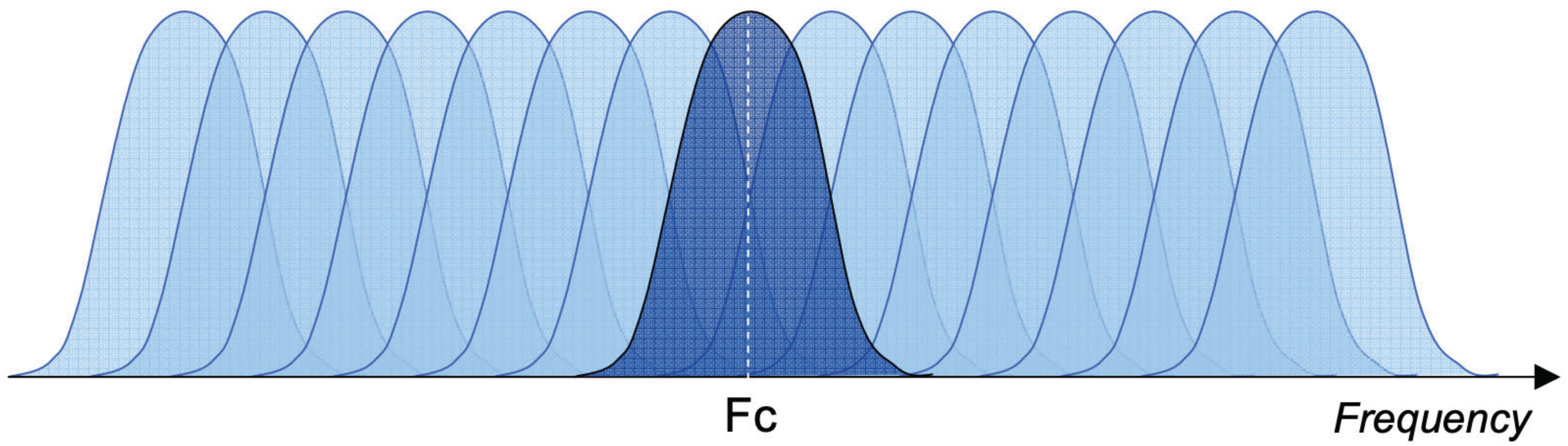
from: Wikipedia

Auditory Specifications

Critical Bands

The auditory system behaves like a bank of overlapping bandpass filters. These filters are termed "auditory filters".

Fletcher H. (1940). Auditory patterns. Rev. Mod. Phys. 12, 47-65.



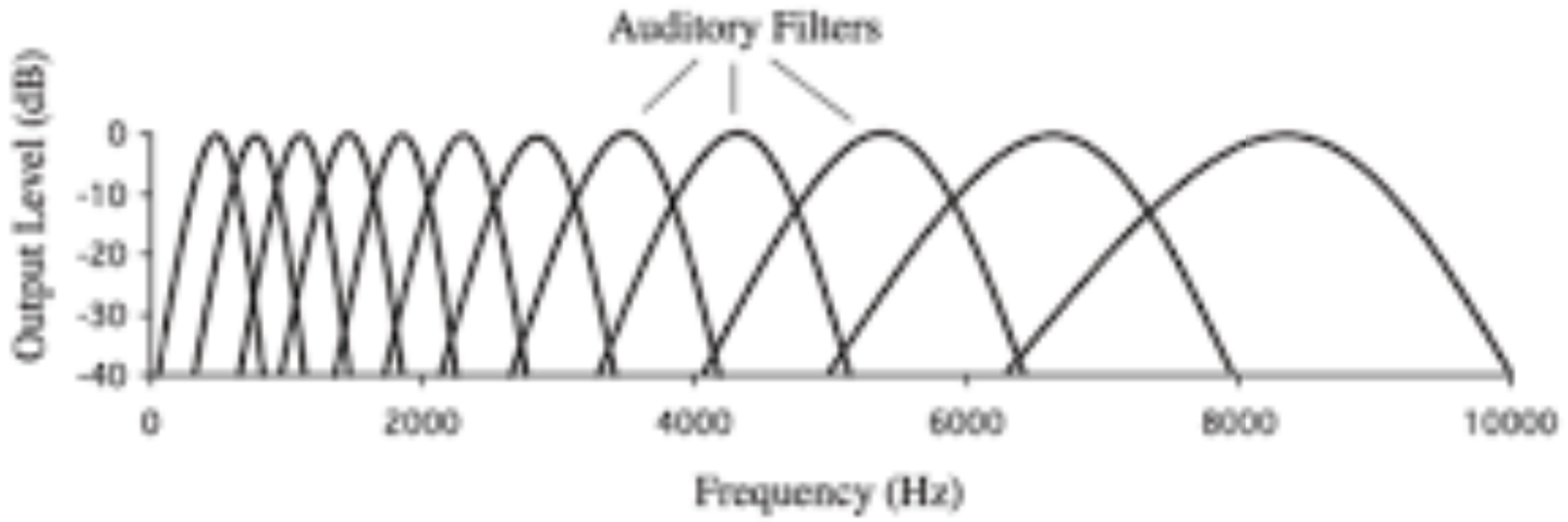
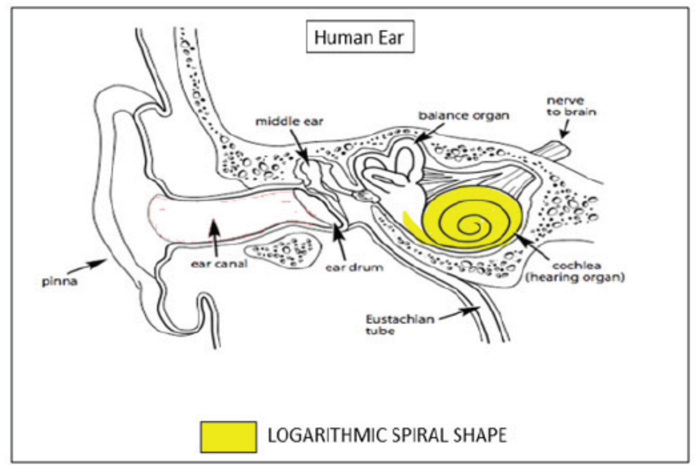
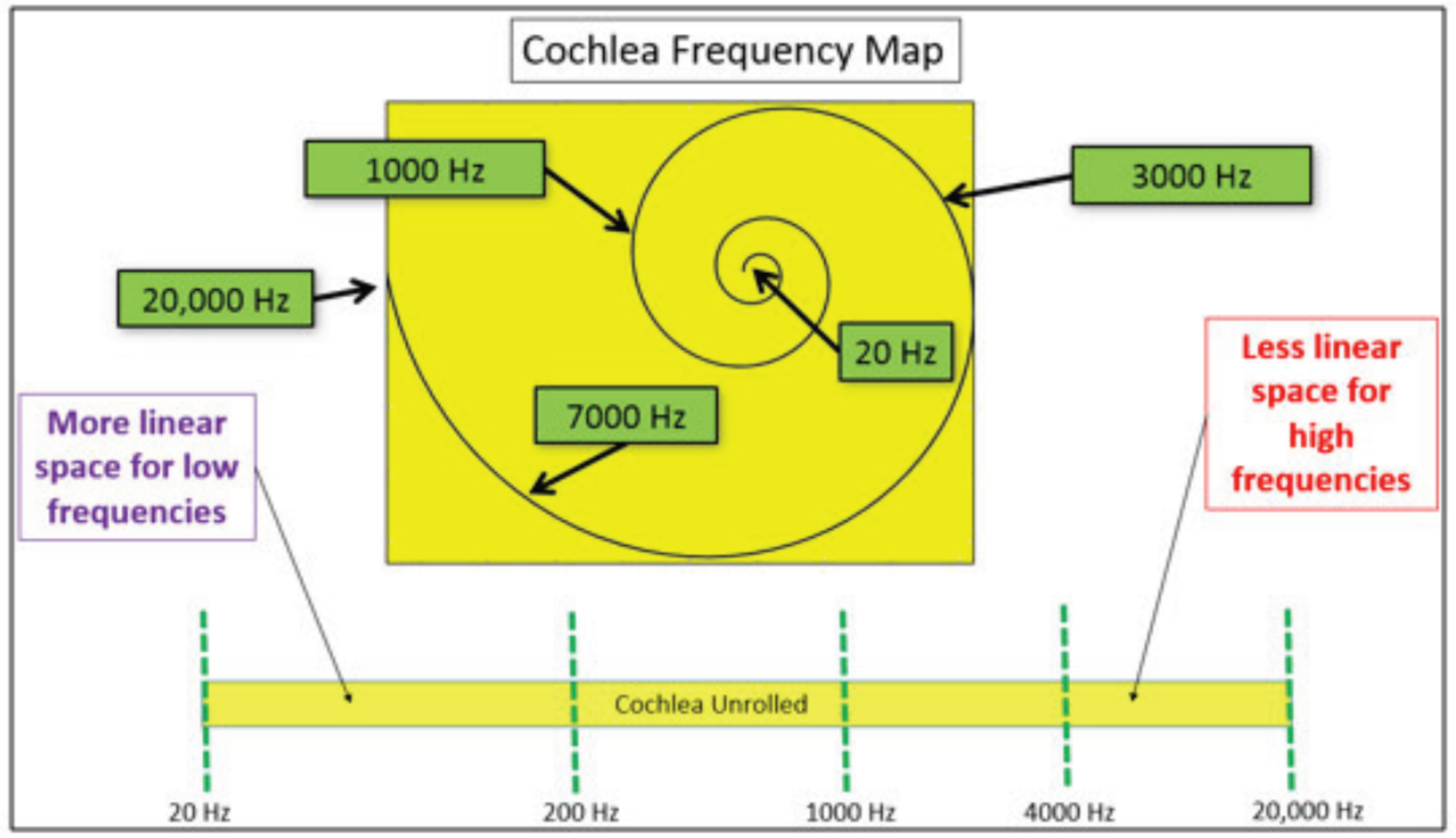


Figure 1 is a conceptual drawing of auditory filters. In reality there are more auditory filters and they are closer together. On the horizontal axis is frequency and sound intensity is shown on the vertical axis. (Plack, 2013).

Auditory Specifications

Critical Bands (II)





from: http://hyperphysics.phy-astr.gsu.edu/hbase/Sound

Critical Band (Bark)	Center Frequency (Hz)	Bandwidth (Hz)
1	50	100
2	150	100
3	250	100
4	350	100
5	450	110
6	570	120
7	700	140
8	840	150
9	1000	160
10	1170	190
11	1370	210
12	1600	240
13	1850	280
14	2150	320
15	2500	380
16	2900	450
17	3400	550
18	4000	700
19	4800	900
20	5800	1100
21	7000	1300
22	8500	1800
23	10500	2500
24	13500	3500

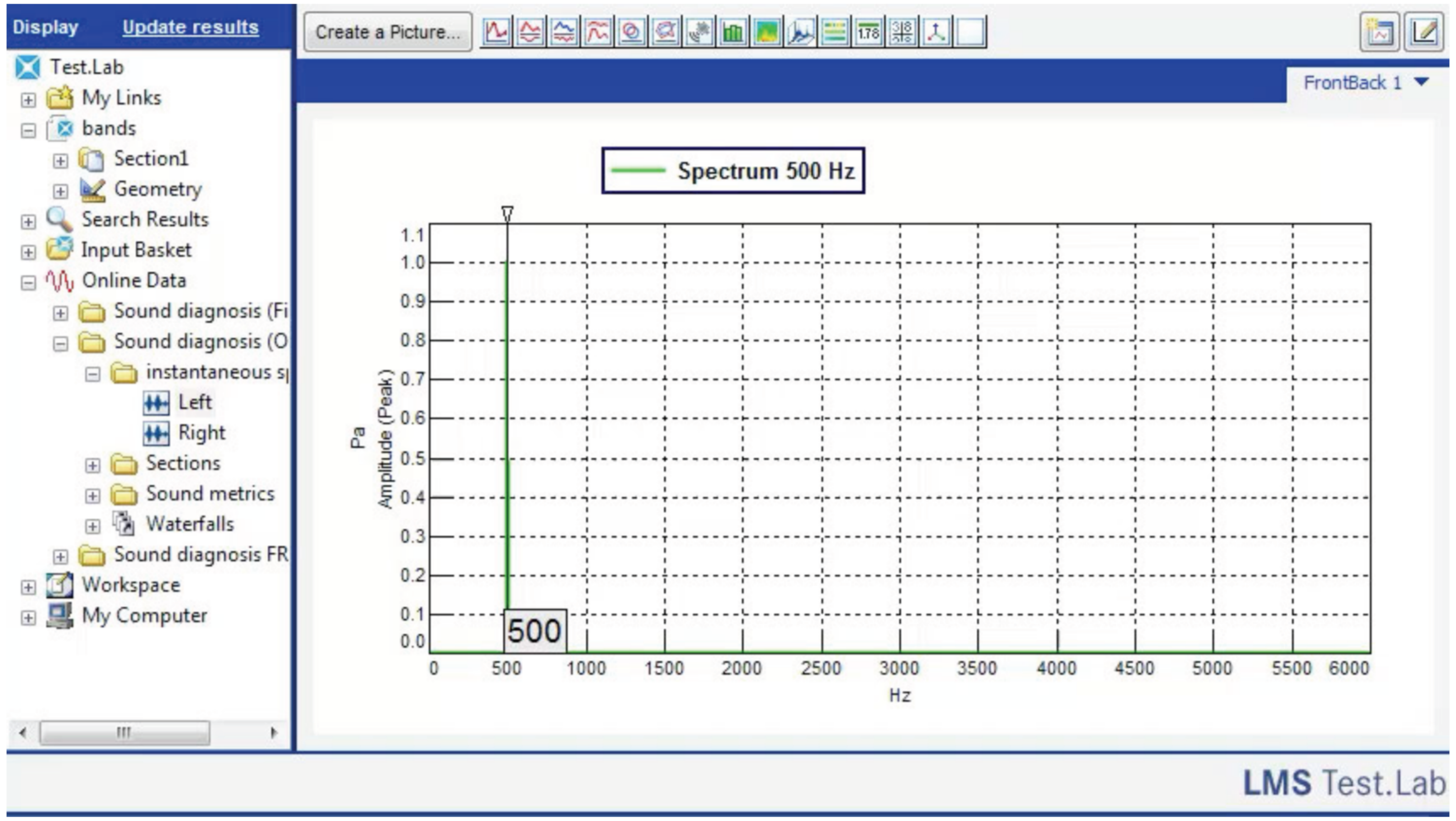
Critical Bands and the Bark Scale

from: https://community.sw.siemens.com/s/article/critical-bands-in-human-hearing

[See the Hudspeth Cochlear Animation]

Auditory Specifications

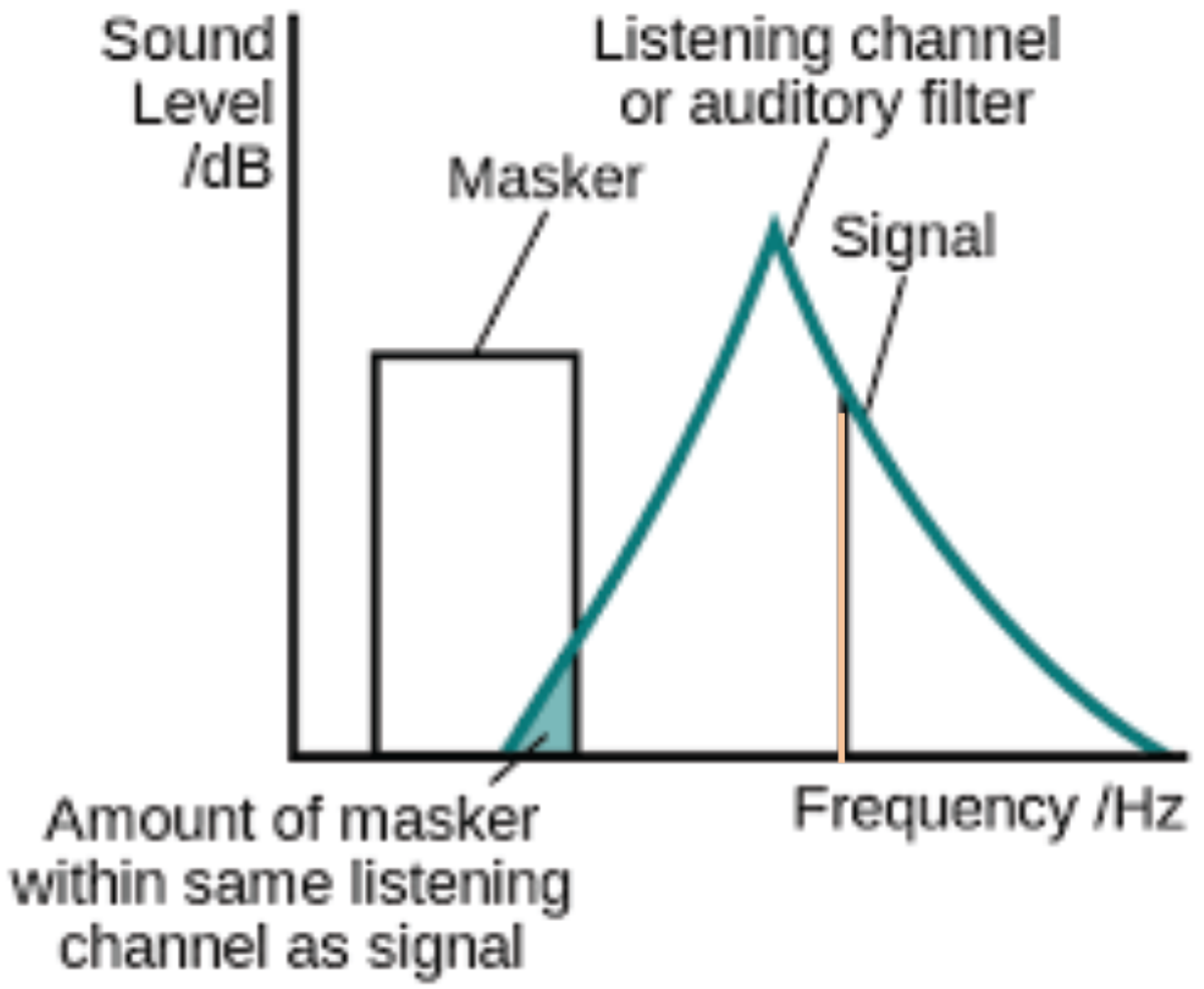
Frequency Resolution



From: https://community.sw.siemens.com/

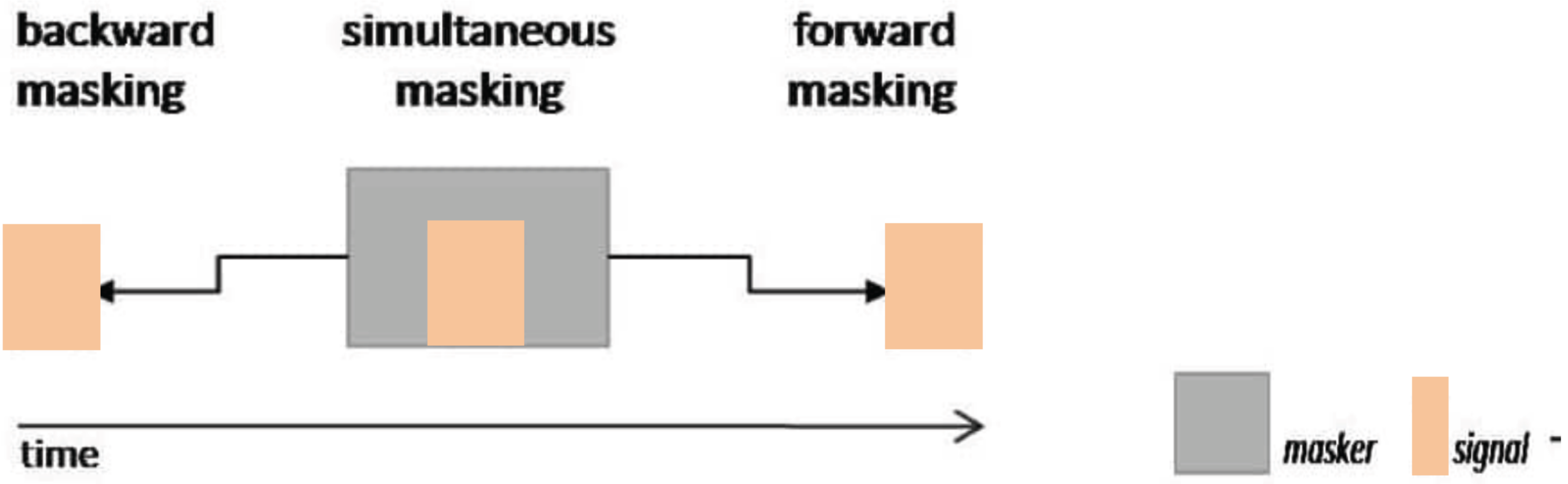
Auditory Masking

Spectral

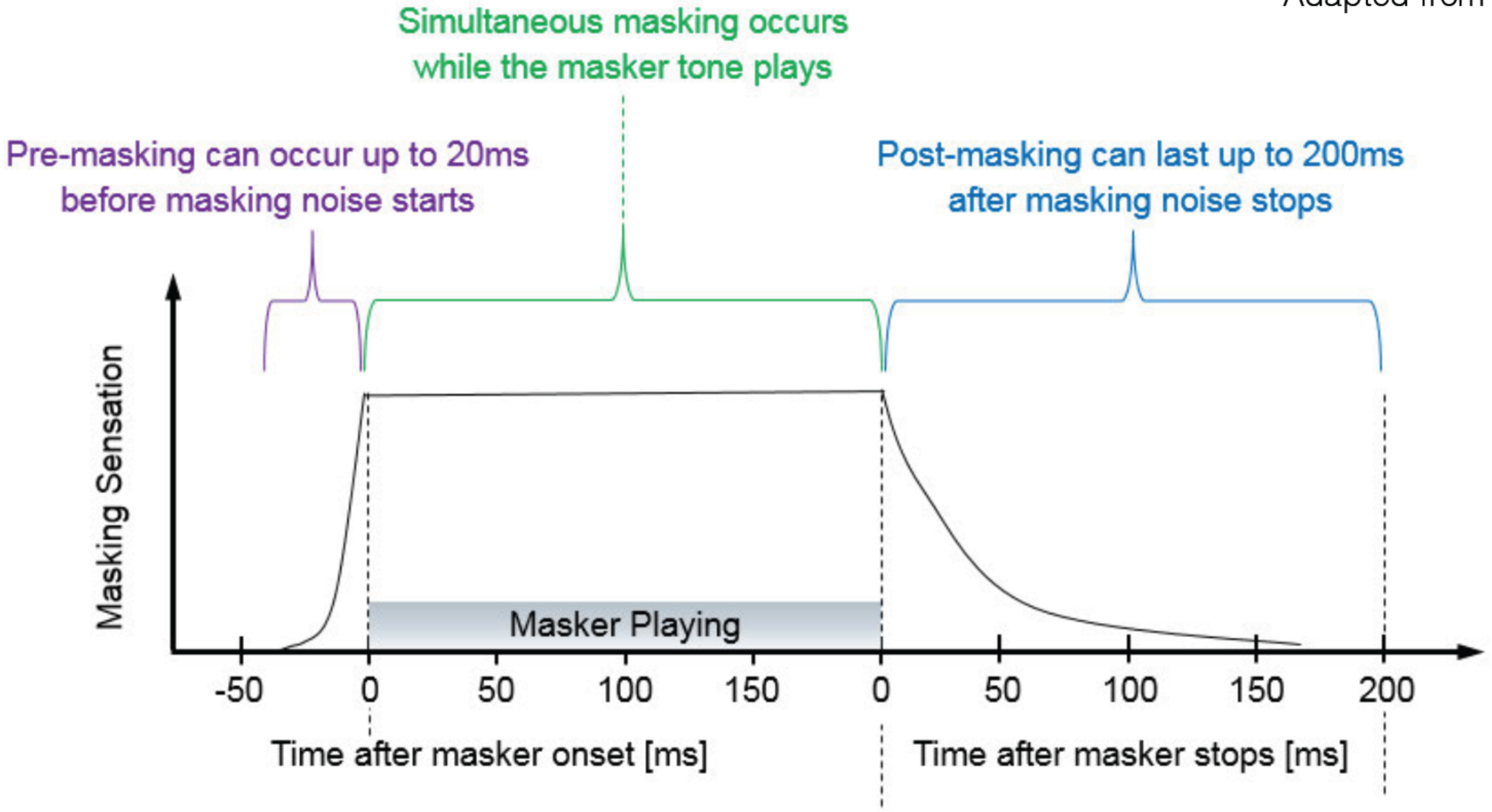


Adapted from Wikipedia

Temporal



Adapted from Wikipedia

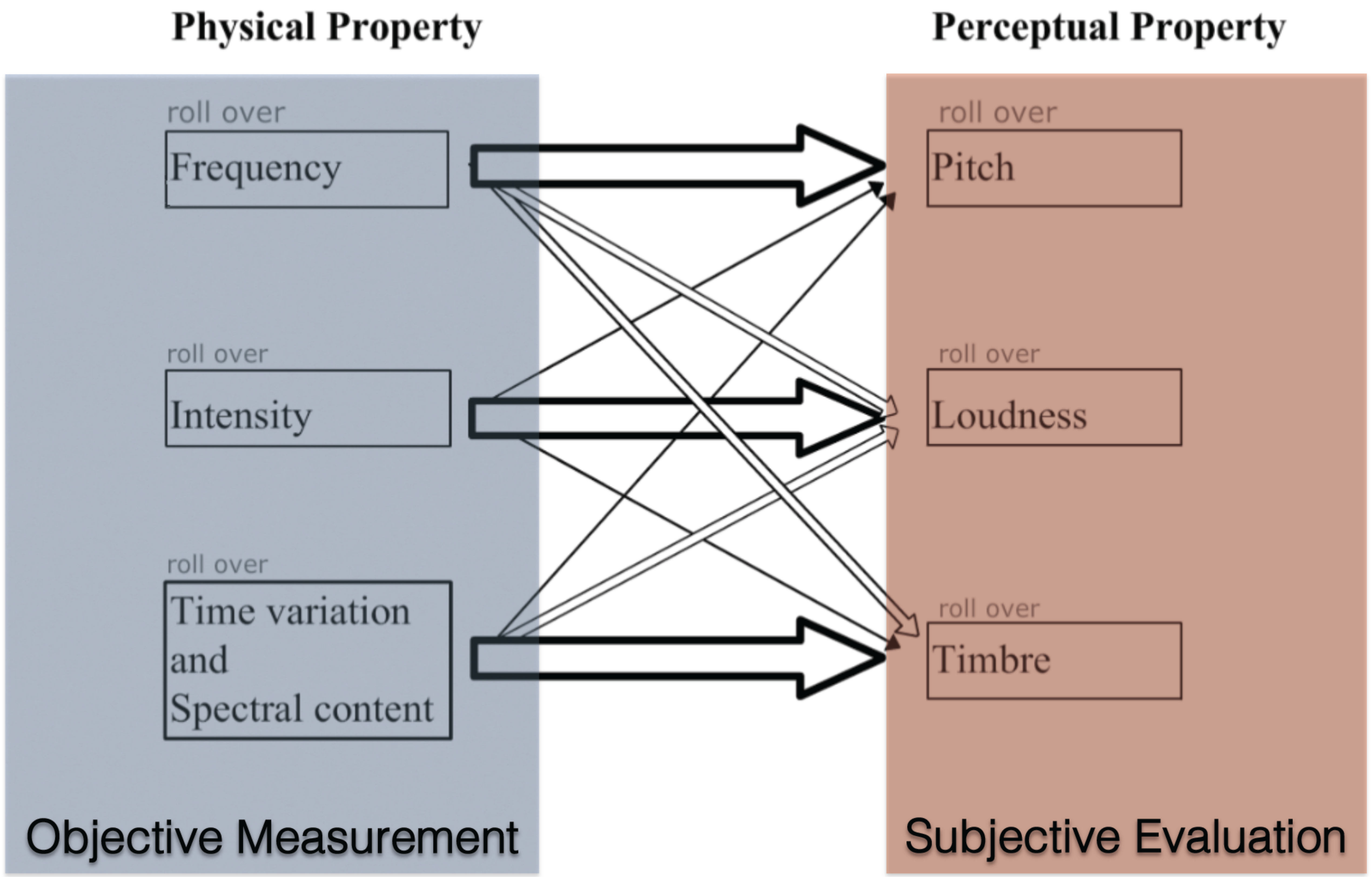


Auditory Masking

Masking a pure tone
with white noise

From: https://community.sw.siemens.com/

Subjective Evaluation

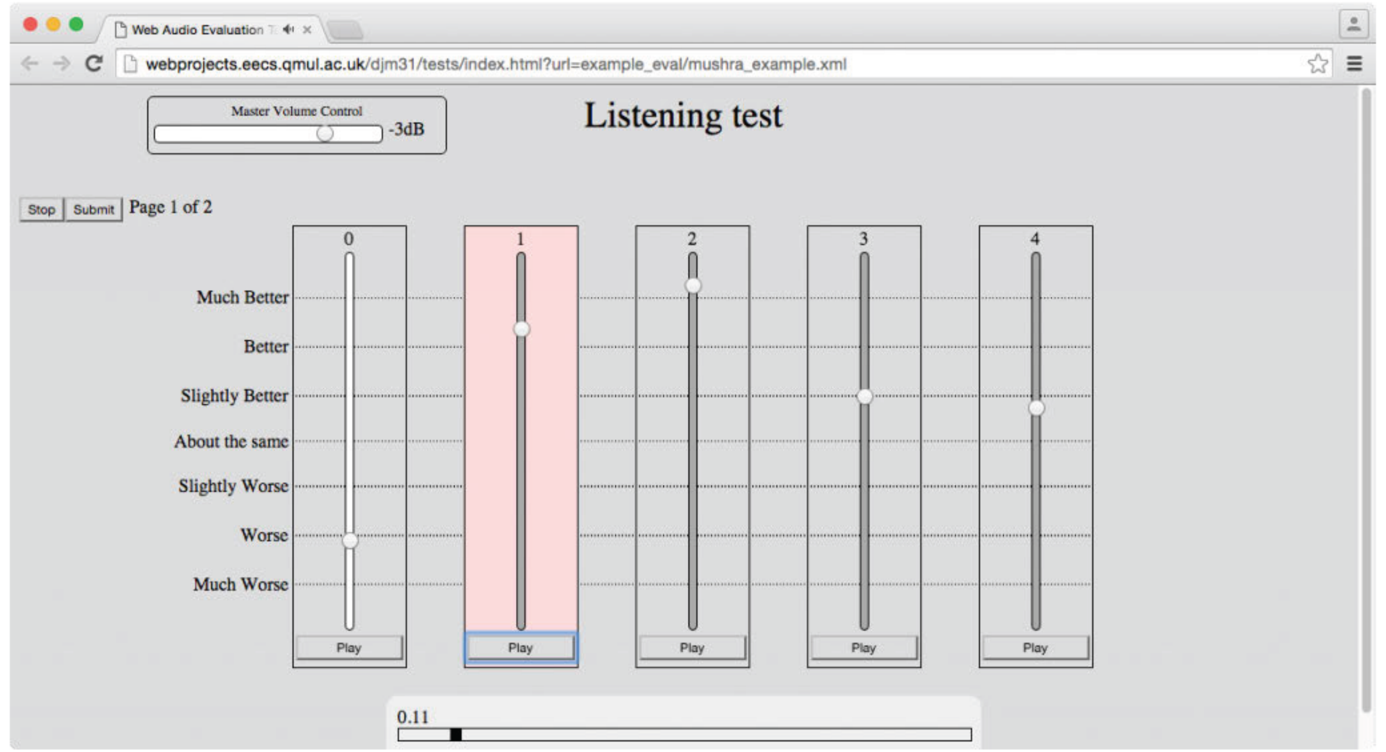


Subjective Evaluation (I)

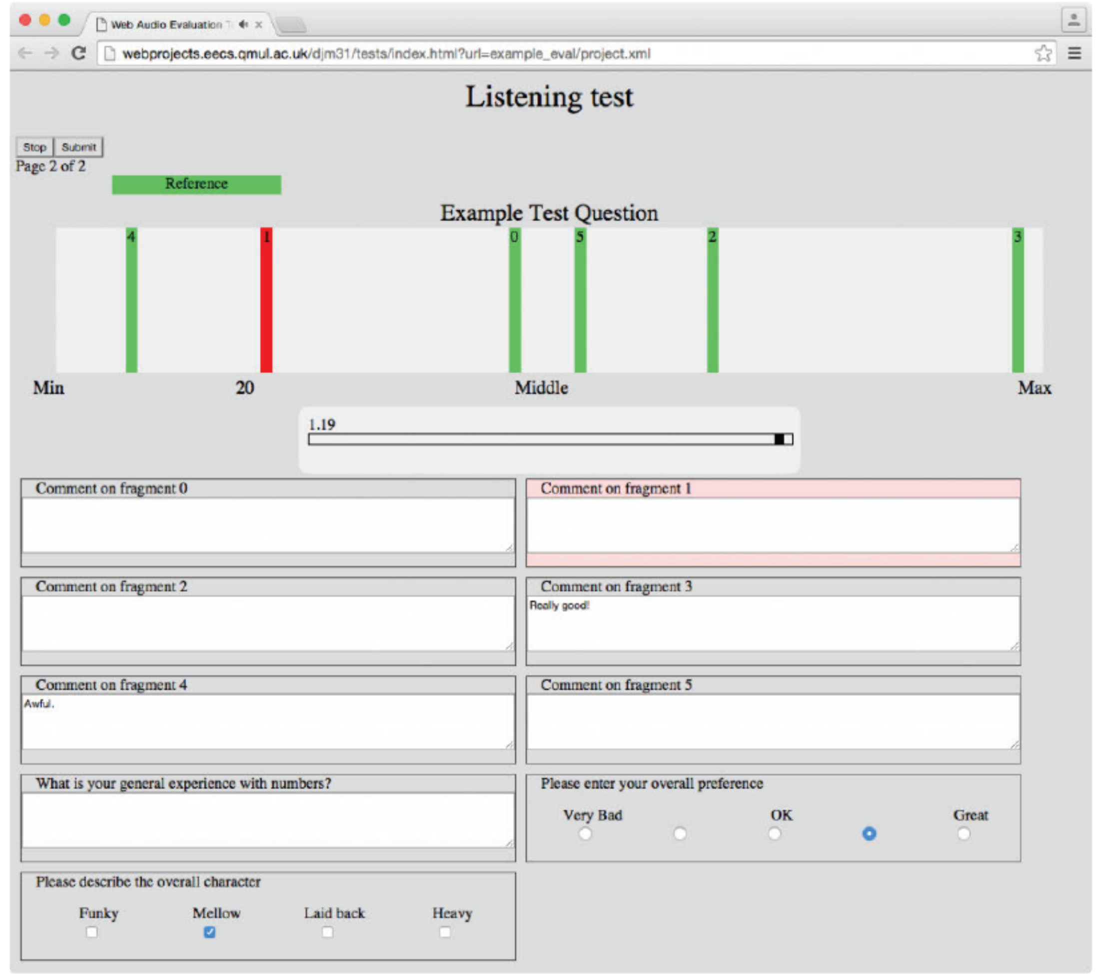
Perceptual listening tests

- fundamental in audio and music research evaluation

- Audio Perceptual Evaluation – APE
- ITU-R BS.1534-1 (MUSHRA)
- ITU-R BS.1116-1 (ABC / Hidden Reference)
- ITU-R BS. 562-3 (Continuous Impairment Scale)
- ITU-T P. 800 Absolute Category Rating (ACR)
- A/B Testing and Pairwise
- -50 to 50 / Bipolar
- ABX Likert Scale [Rank Scale]
- 9 Point Hedonic Scale



MUSHRA Test



APE Test

From Web Audio Evaluation Tool http://www.semanticaudio.co.uk/projects/waet/

Subjective Evaluation (II)

Loudness – *Sone*

The **sone** is derived from psychophysical measurements which involved volunteers adjusting sounds until they judge them to be twice as loud. This allows one to relate perceived loudness to phons:

1 sone is defined to be equal to 40 phons

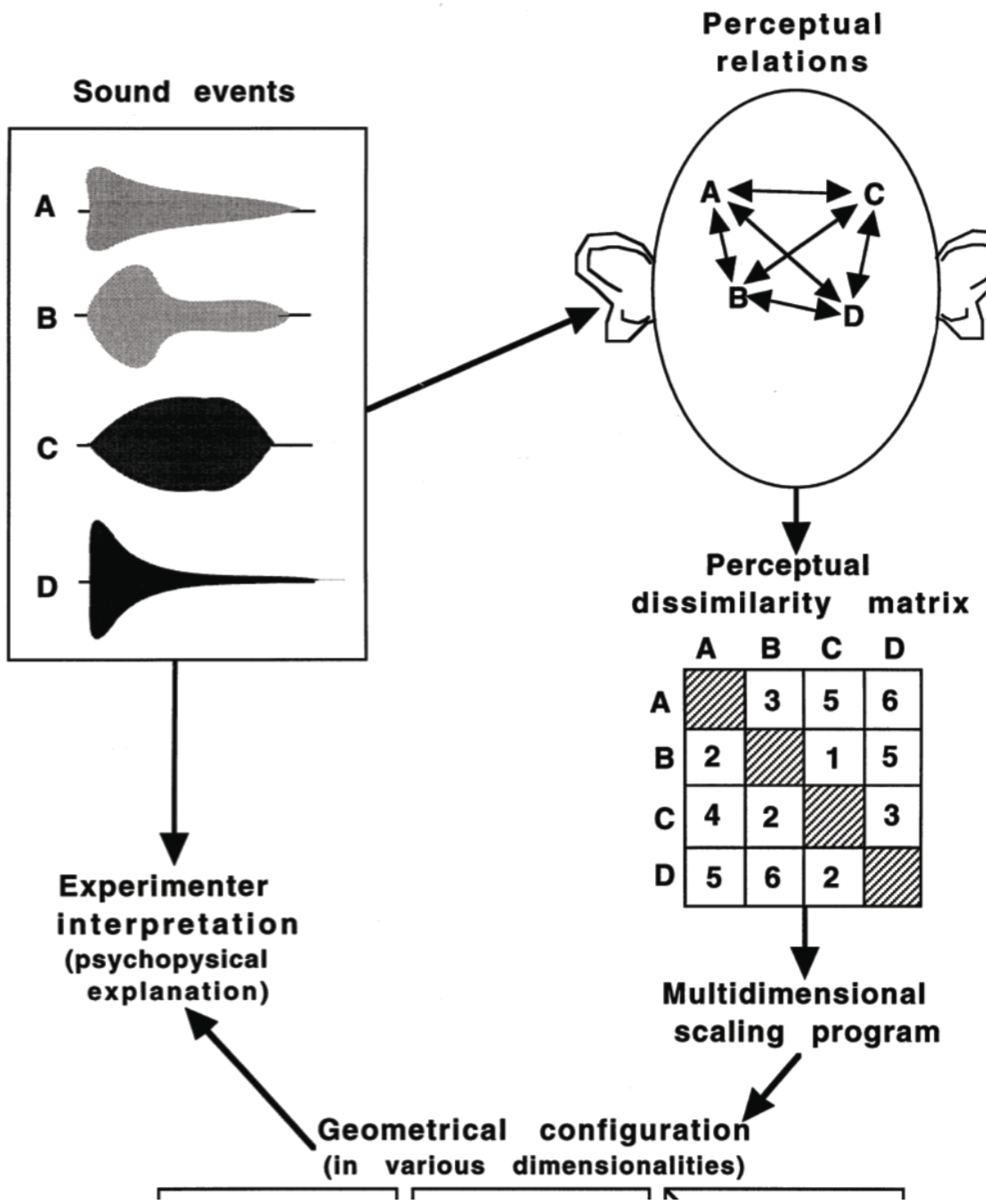
Experimentally it was found that, above 40 phons, a 10 dB increase in sound level corresponds approximately to a perceived doubling of loudness. So that approximation is used in the definition of the sone: 1 sone = 40 phon, 2 sone = 50 phon, 4 sone = 60 phon, etc.

sone	1	2	4	8	16	32	64	128	256	512	1024
phon	40	50	60	70	80	90	100	110	120	130	140

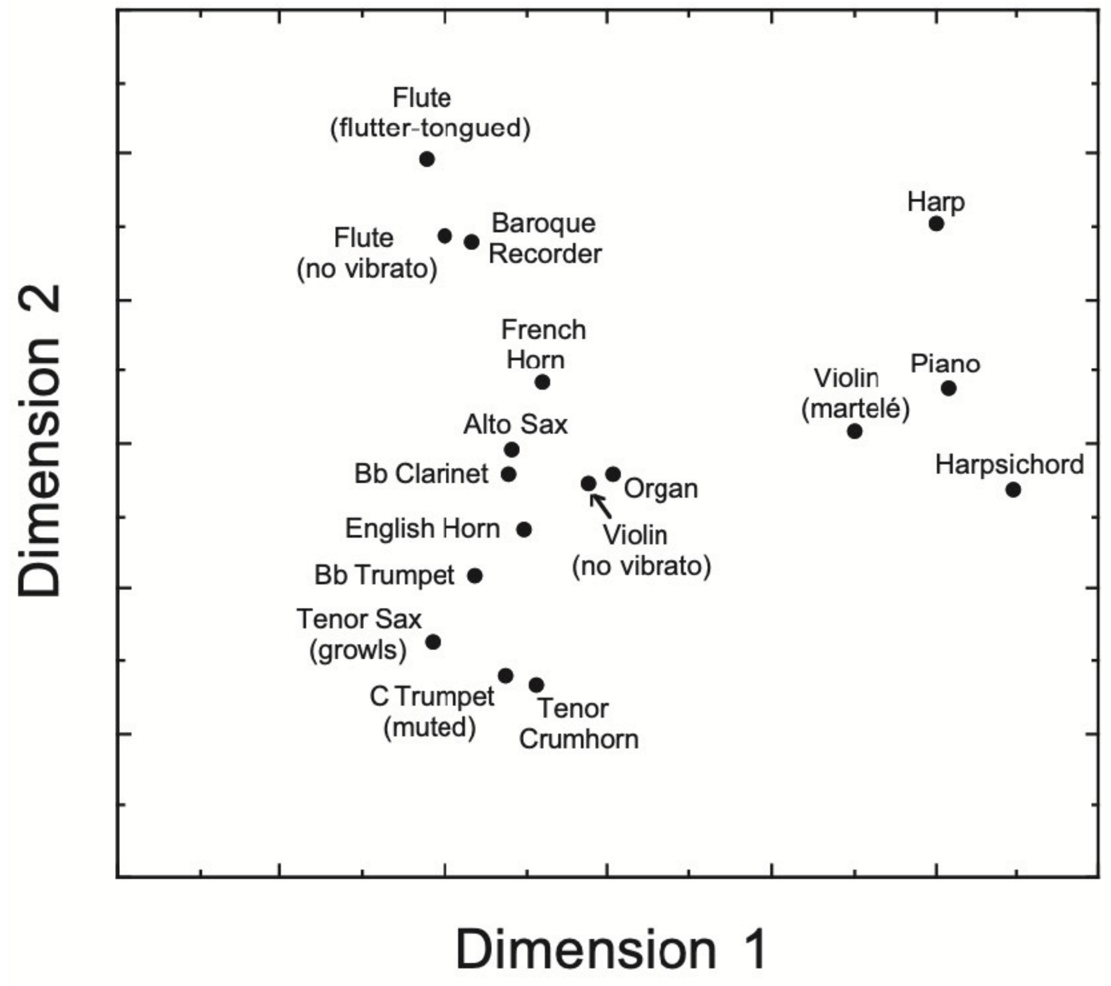
from: Wikipedia

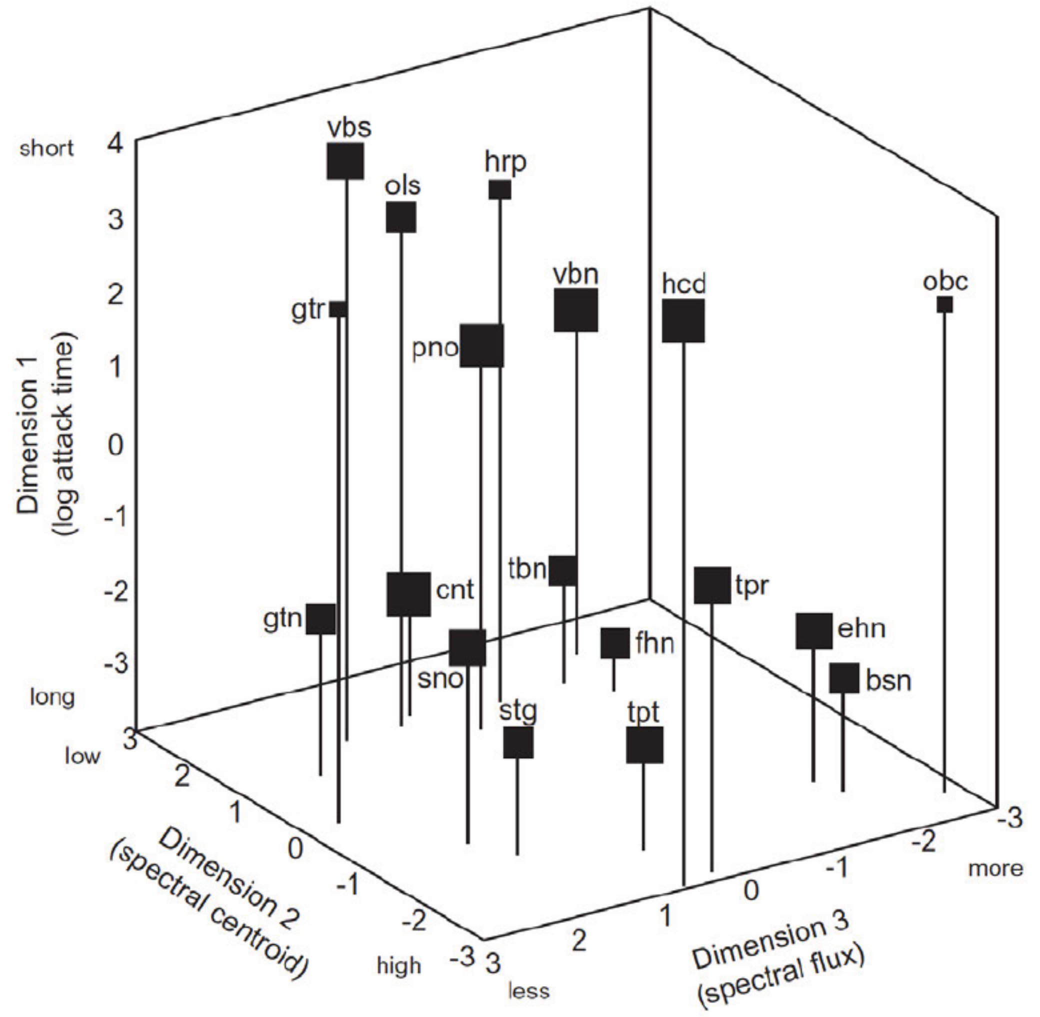
Subjective Evaluation (III)

Timbre Spaces



from: (Deutsch) The Psychology of Music. *Academic Press, 2013*





McAdams et al. Timbre Space (1995);

from: (Deutsch) The Psychology of Music. *Academic Press, 2013*

Subjective Evaluation (III)

Timbre Spaces

Demonstration

Resources

General
- Introduction to Computer Music – Vol I
- Hyperphysics - Sound and Hearing

(Do your) Hearing Test Online
https://newt.phys.unsw.edu.au/jw/hearing.html

Perceptual Audio Demonstrations
https://www.ece.uvic.ca/~elec499/2003a/group09/p/demos.htm
https://pure.tue.nl/ws/files/79033010/402660.pdf

2.2. The Auditory System