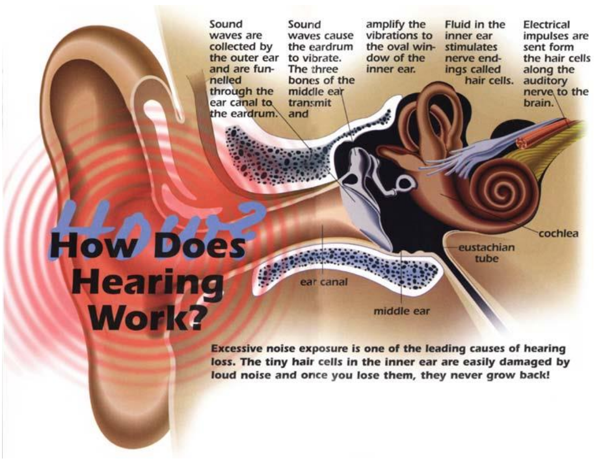
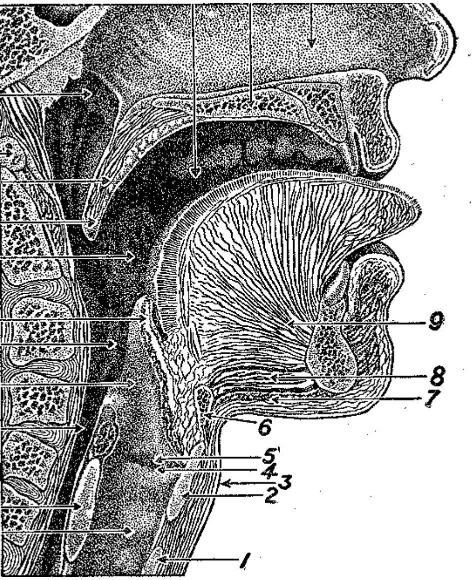


Infocommunication
Sound and hearing



Dr. Mohammed Salah Al-Radhi

Dr. Tamás Gábor Csapó

malradhi@tmit.bme.hu

Copyright

Topics

- Basic signal processing
- Sampling and quantization
- Analog modulation
- Digital baseband modulation
- Digital carrier modulation
- Error Detection Coding
- Error Correction Coding
- **Sound, hearing and speech**
- Light and vision
- Radio Communication
- Video Broadcasting
- Mobile communication (1G, 2G, 3G, 4G, 5G)

http://smartlab.tmit.bme.hu/education-infocommunications

SOUND AND HEARING

Speech

- the most natural form of human-human communications
- related to language; linguistics is a branch of social science
- related to human physiological capability; physiology is a branch of medical science
- also related to sound and acoustics, a branch of physical science
- one of the most interesting signals that humans work with every day

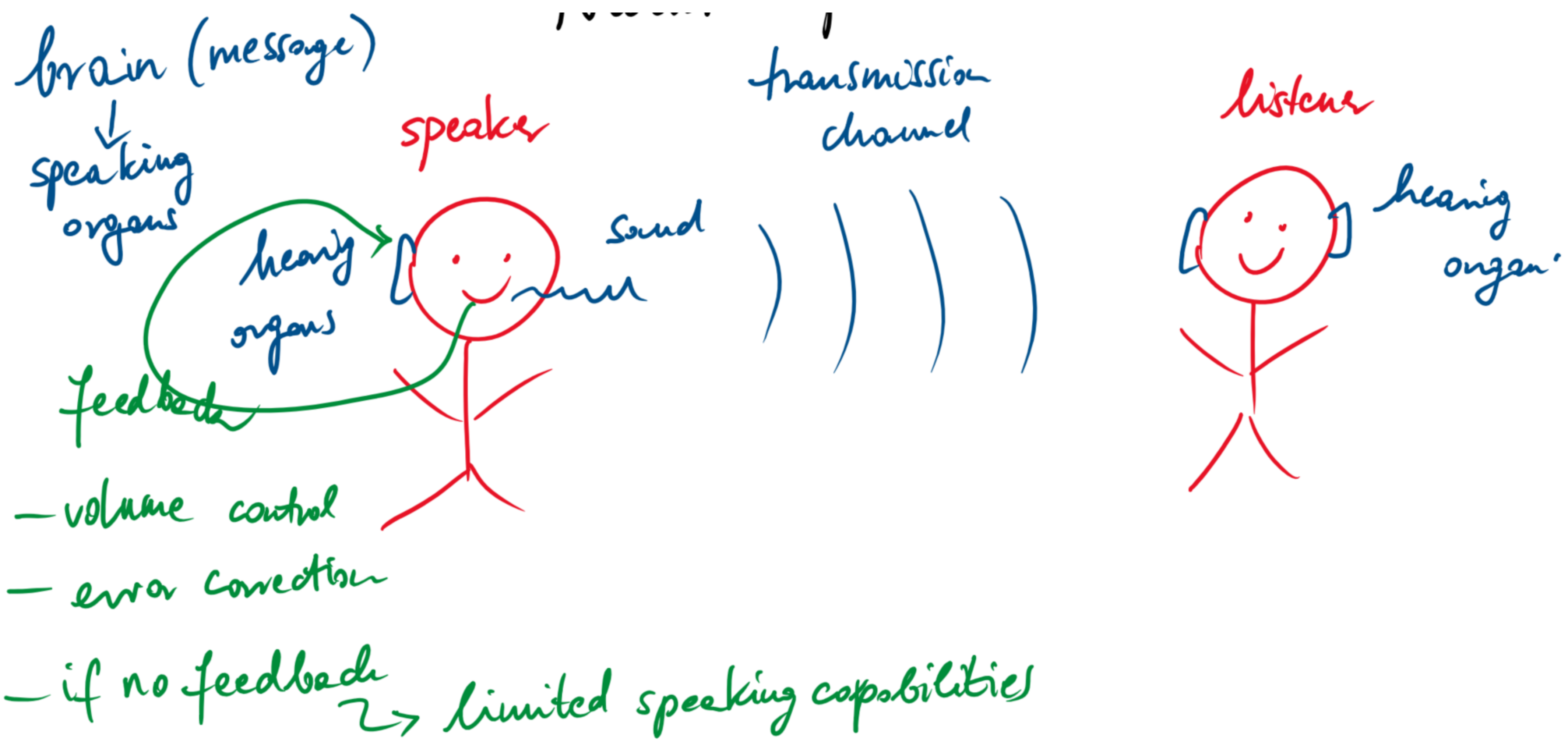
Source: Rabiner (2015) http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/speech%20course.html

Speech processing

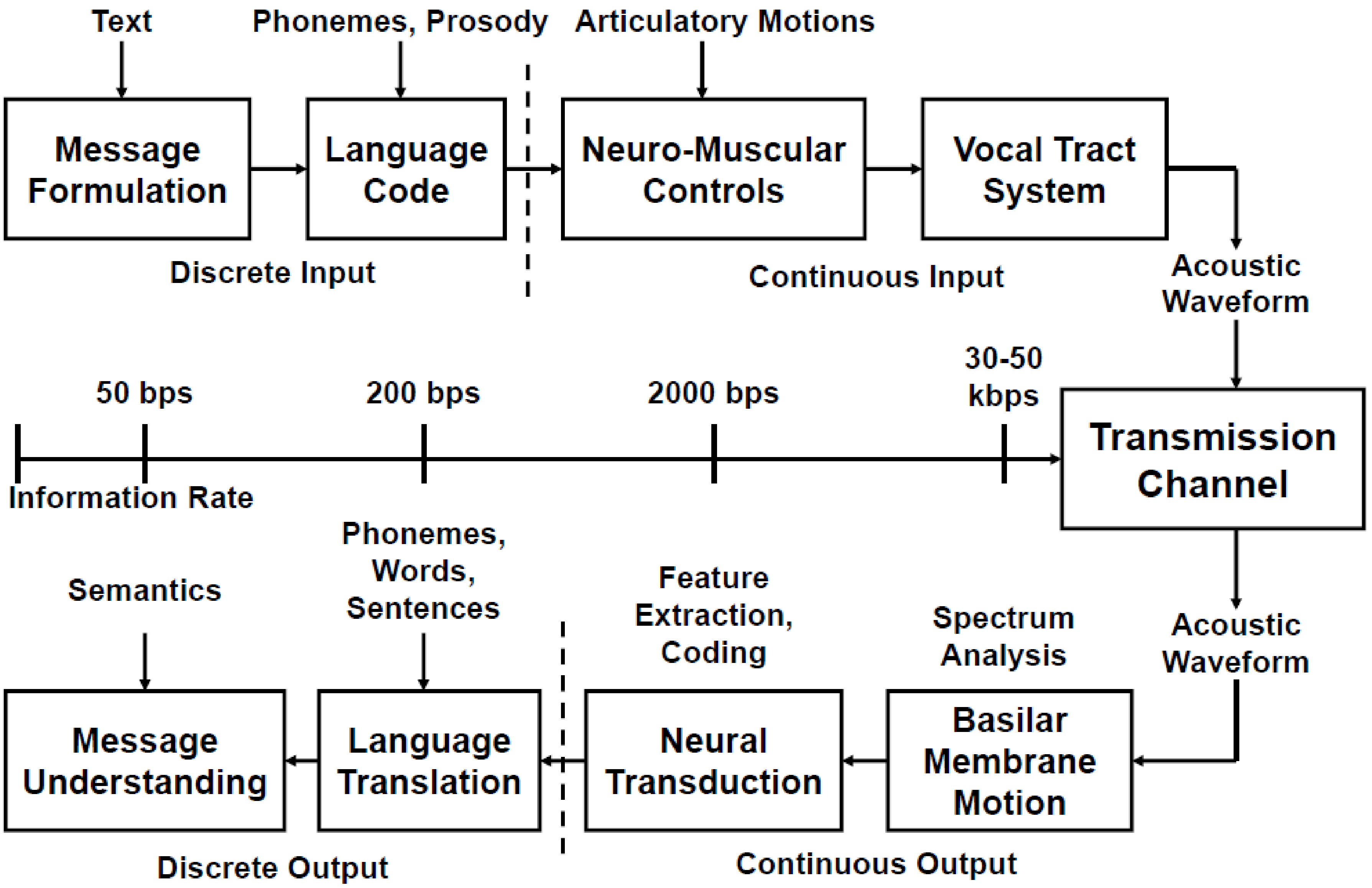
- Purposes:
  - to understand speech as a means of communication
  - to represent speech for transmission and reproduction
  - to analyze speech for automatic recognition and extraction of information
  - to discover some physiological characteristics of the talker

Source: Rabiner (2015) http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/speech%20course.html

Natural speech communication chain

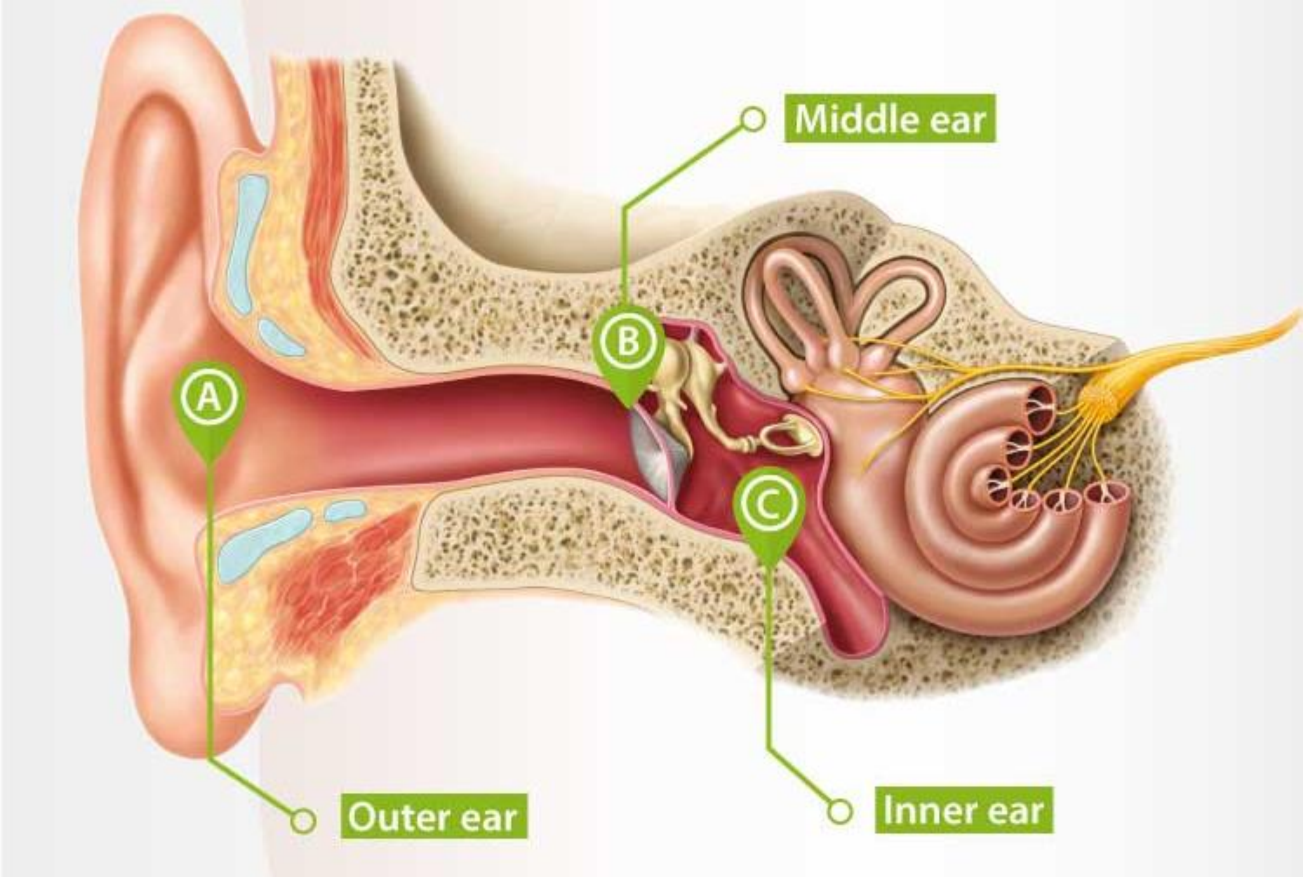


The Speech Chain



Source: Rabiner (2015) http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/speech%20course.html

Structure of the human ear

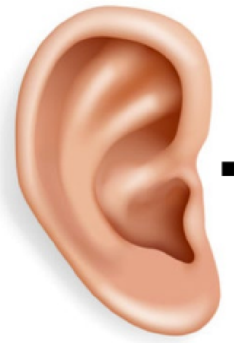


1.Outer Ear: collects sound waves and directs them into the ear canal, which then vibrates as the sound waves hit the eardrum.

2.Middle Ear: responsible for transmitting sound waves from the outer ear to the inner ear. The three smallest bones in the body, the malleus, incus, and stapes, work together to amplify the sound waves and transmit them to the inner ear via the oval window.

3.Inner Ear: Converts sound waves to electrical signals for the brain to interpret and helps with balance and spatial orientation. The cochlea contains tiny hair cells that are stimulated by the vibrations of fluid inside and send electrical signals to the brain via the auditory nerve, responsible for interpreting sound.

How Do We Hear?



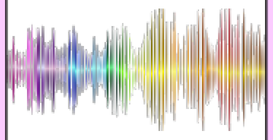
Sound waves are funnelled into our ear canal by the pinna.

The sound waves make our ear drum vibrate.

Vibrations of the ear drum cause the 3 small bones in the ear, collectively called the ossicles, to move.

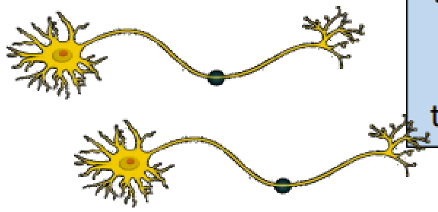
As the last bone in the chain, the stapes, vibrates this causes wave like movements to be generated in the fluid inside the cochlear.

These waves stimulate microscopic hairs inside the cochlear.

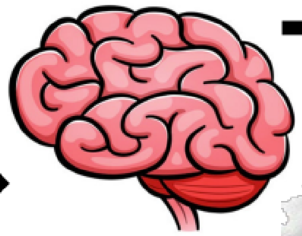


The hairs in the cochlear are tuned to respond to differences in sound frequency and pitch.

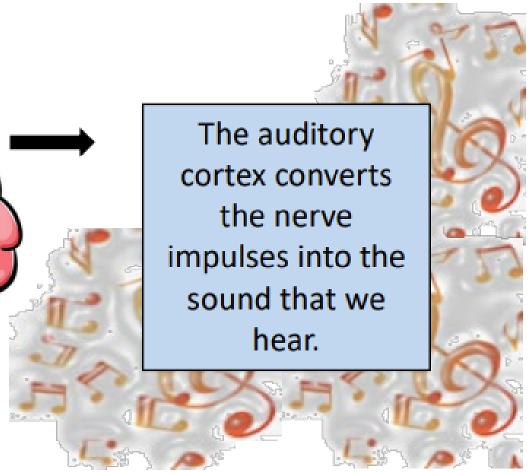
When stimulated these hair cells generate nerve impulses that are transferred to the auditory nerve.



The nerve impulses travel along the auditory nerve into the hearing centre of the brain, called the auditory cortex.



The auditory cortex converts the nerve impulses into the sound that we hear.



https://youtu.be/eQEaiZ2j9oc

Békésy György / Georg von Békésy

Nobel prize in 1961 (function of the cochlea)

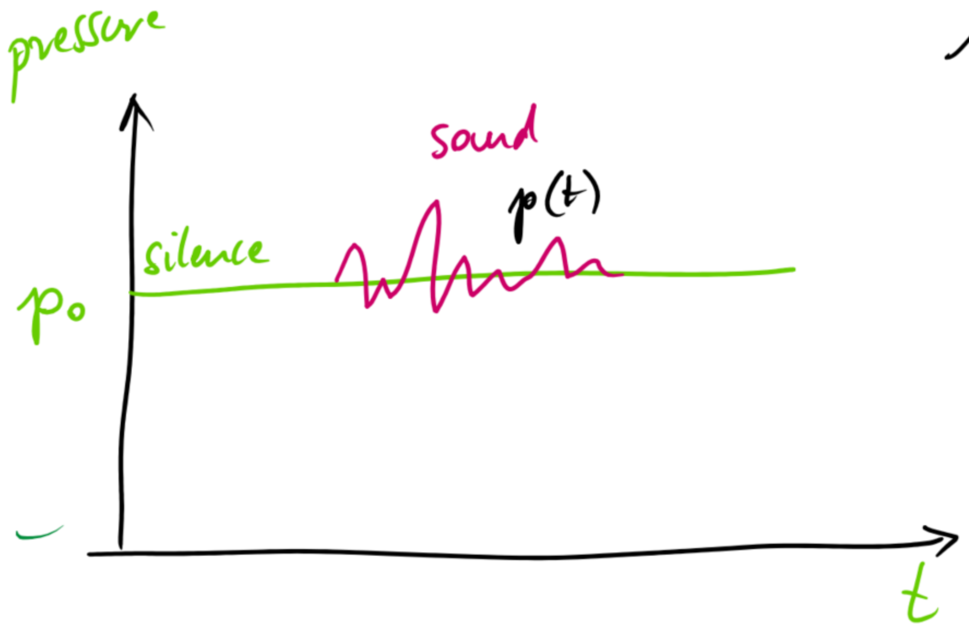
Source: http://braintour.harvard.edu/archives/portfolio-items/von-bekesy-experiments-hearing

Physical modelling of sounds

Physical modelling of sounds

Sound: mechanical vibration of an elastic medium

human ear perceives airborne sounds



$$P_0 = 100\,\mathrm{kPa}$$

Sound pressure change

constant atmospheric pressure
+
alternating component

decreasing: $\frac{1}{\text{distance}}$

$$P(t) = P_0 + p(t)$$

$$p = \frac{F}{A} \quad \left[\frac{N}{m^2}\right] \quad [Pa]$$

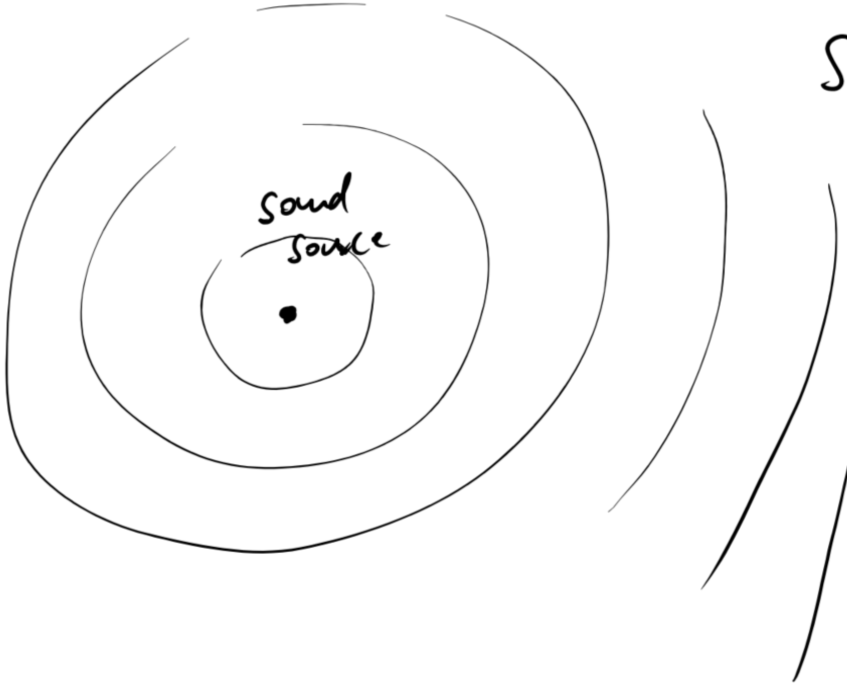
Propagation in short distance / long distance

effective value of pressure change

$$P_{eff} = \sqrt{\frac{1}{t_2-t_1}\int_{t_1}^{t_2}(P(t)-P_0)^2\,dt}$$

decreasing:

$$\frac{1}{\text{distance}}$$

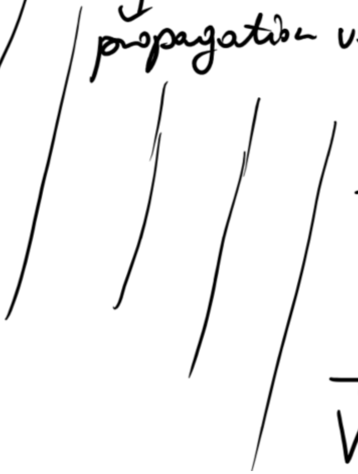


Short distance:

spherical sand waves

$$c_{sand} = f \cdot \lambda = 340\,\mathrm{m/s}$$

propagation velocity of sand



Large distance:

plain waves

air density

$$\frac{p}{v} = \text{constant} = 410\,\frac{\mathrm{kg}}{\mathrm{m^2 s}} = \rho_0 \cdot c$$

- Intensity:
- Volume:
- SPL:

Intensity

$$I = \frac{P \text{ (power)}}{A \text{ (area)}} \left[\frac{W}{m^2}\right]$$

$I_0 = 10^{-12}\, W/m^2$ (reference I)

acoustic power which comes through a unit of area

decreasing $\frac{1}{\text{distance}^2}$

Volume (sound pressure level)

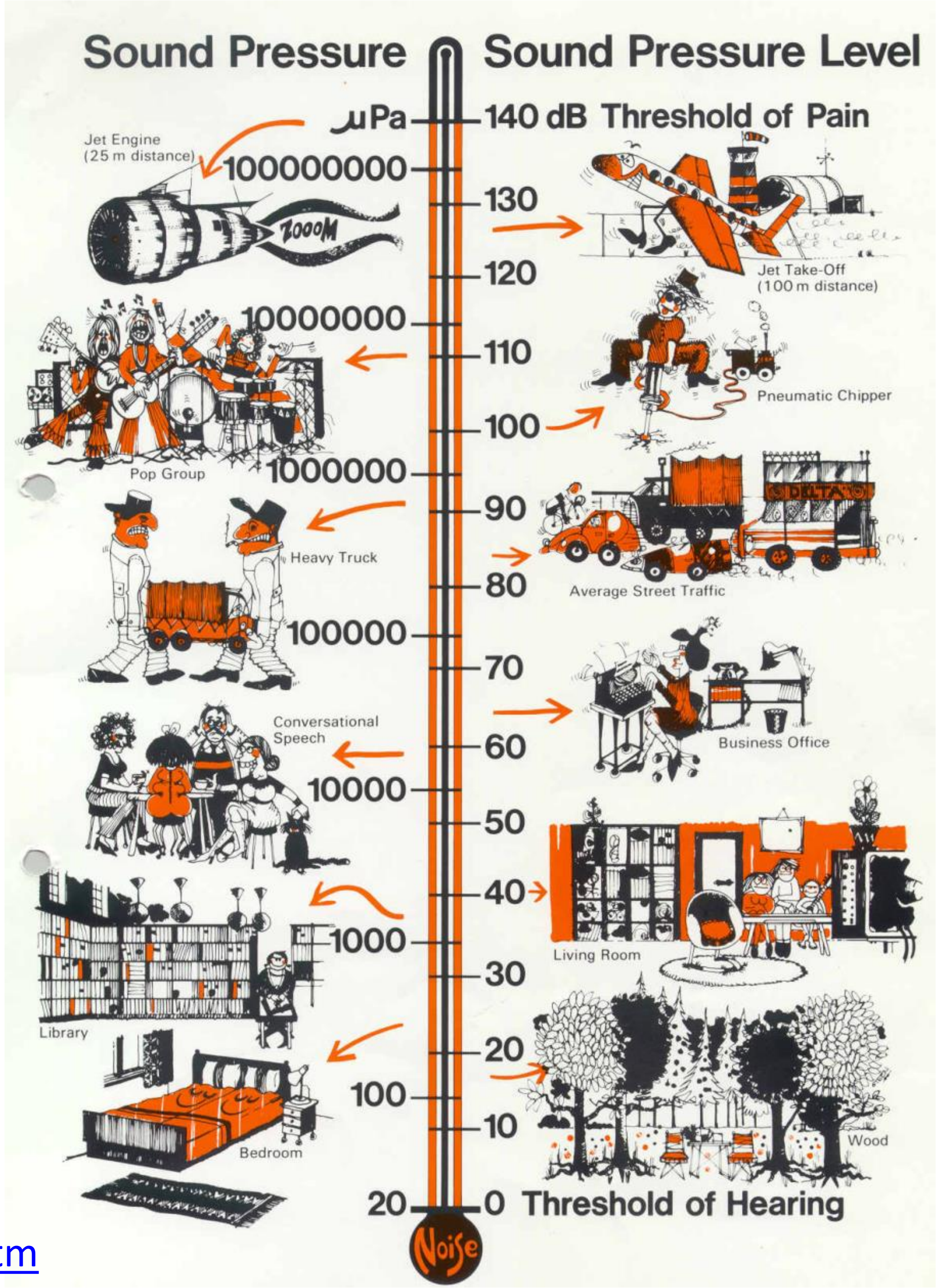
$$L = 20 \cdot \lg \frac{p}{p_0} \; [dB] = 10 \cdot \lg \frac{I}{I_0}$$

(reference pressure)

$p_0 = 20 \mu Pa \rightarrow$ just yet audible sound

SPL : sound pressure relative to the reference p. in acoustic decibel

Sound pressure level



Source:
http://personal.cityu.edu.hk/~bsapplec/sound.htm

Sound pressure level

- 440 Hz tone (A4 on musical scale)
  - reduced in 1 dB steps
  - reduced in 3 dB steps
  - reduced in 5 dB steps

Source: http://www.ece.uvic.ca/~elec499/2003a/group09/p/demos.htm

Physiological & psychoacoustical properties of hearing

- Objective parameters

Subjective parameters

Physiological, psychoacoustical properties of hearing

Objective parameters
[can be measured]

Subjective parameters
[perceptual correlates]
good expressions

– pressure / intensity level ⟷ loudness

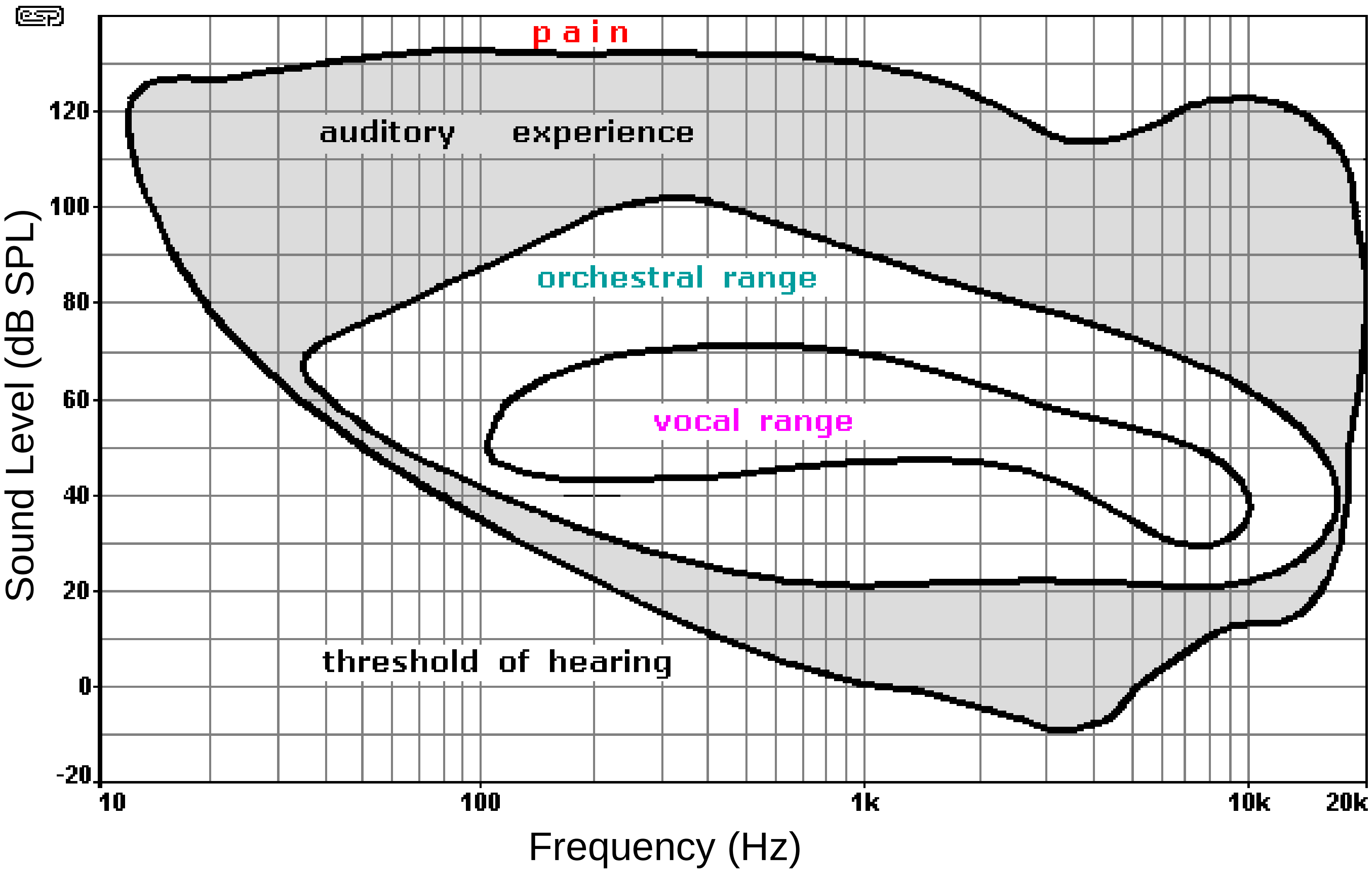
– F_0, fundamental frequency ⟷ pitch
($1/F_0 = T$, period)

"a" sound

– spectrum of sound ⟷ tone, timbre
(frequency components)

all 3 interact with each other
louder → higher pitch (?)

Limits of human hearing



Source: https://people.ece.cornell.edu/land/courses/ece5030/FinalProjects/s2014/kkp37_rjs483/kkp37_rjs483/AudioGram.html

Equal loudness level contours

- Def: loudness level

Equal loudness level contours L_N

↓

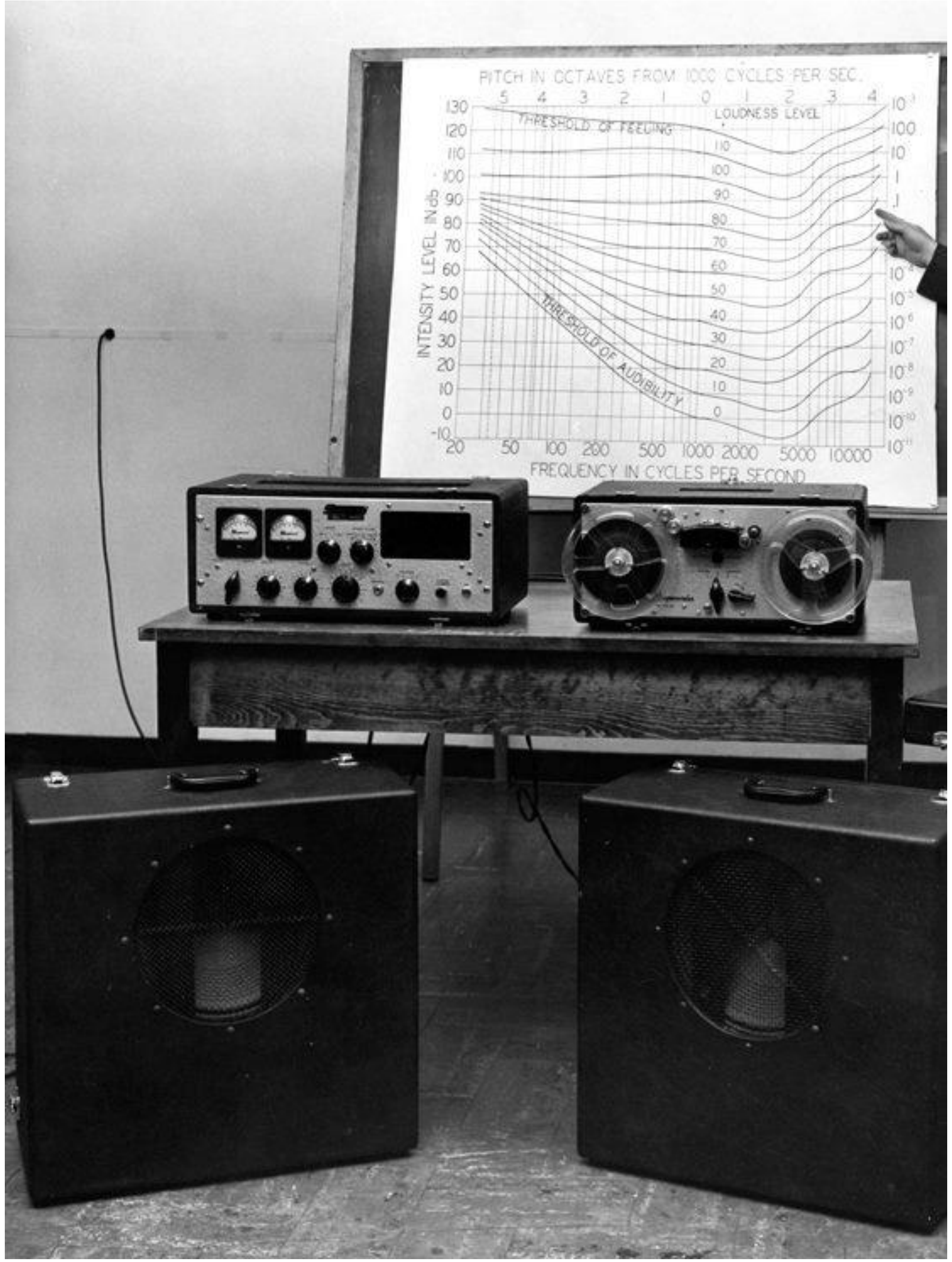
comparison of sounds with different freqs

1933, Fletcher & Munson

Loudness level : [phon]

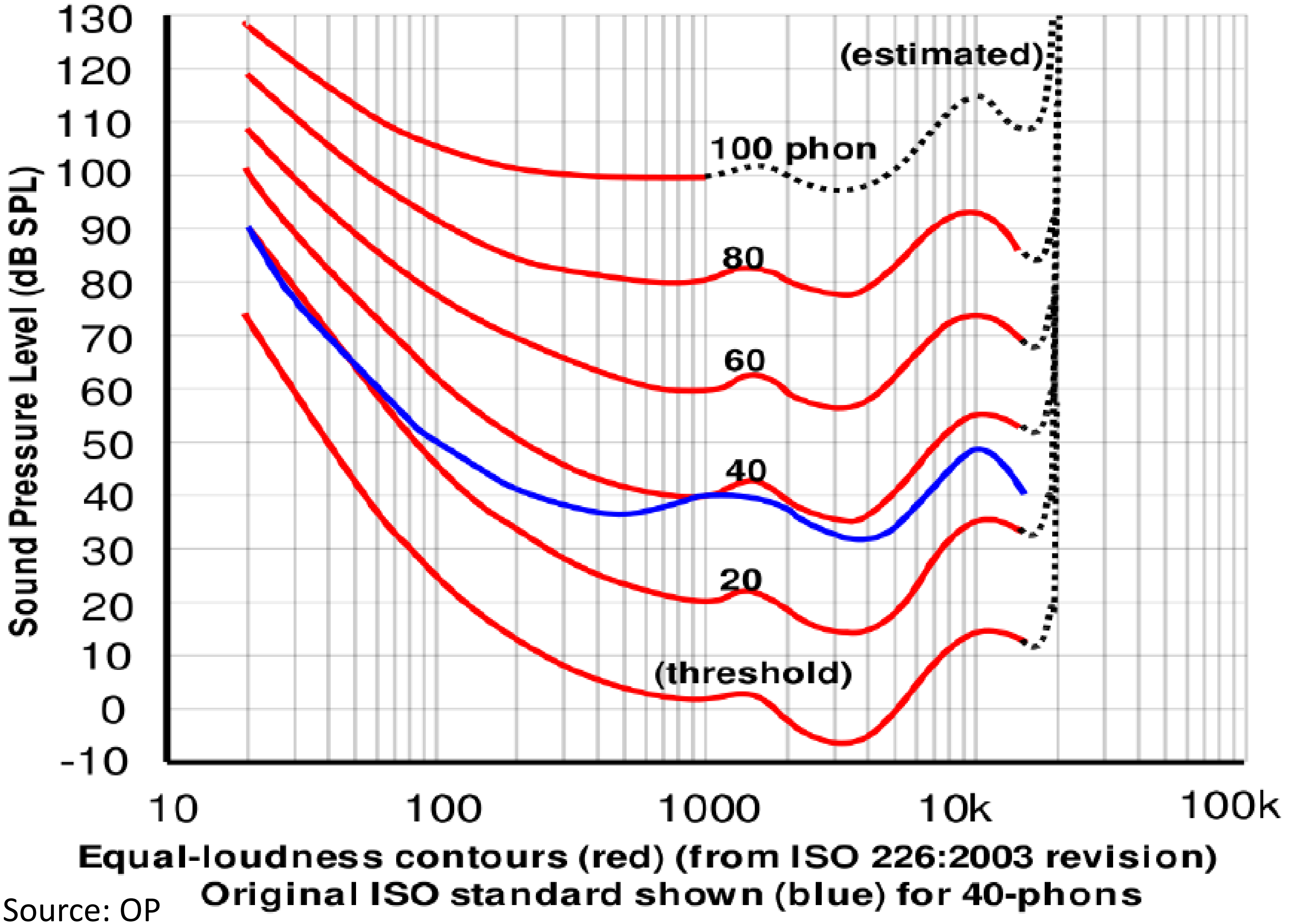
the [phon] number of a sound is the dB SPL of another sound at a frequency of 1kHz that sounds just as loud.

Fletcher & Munson
1933
experiment

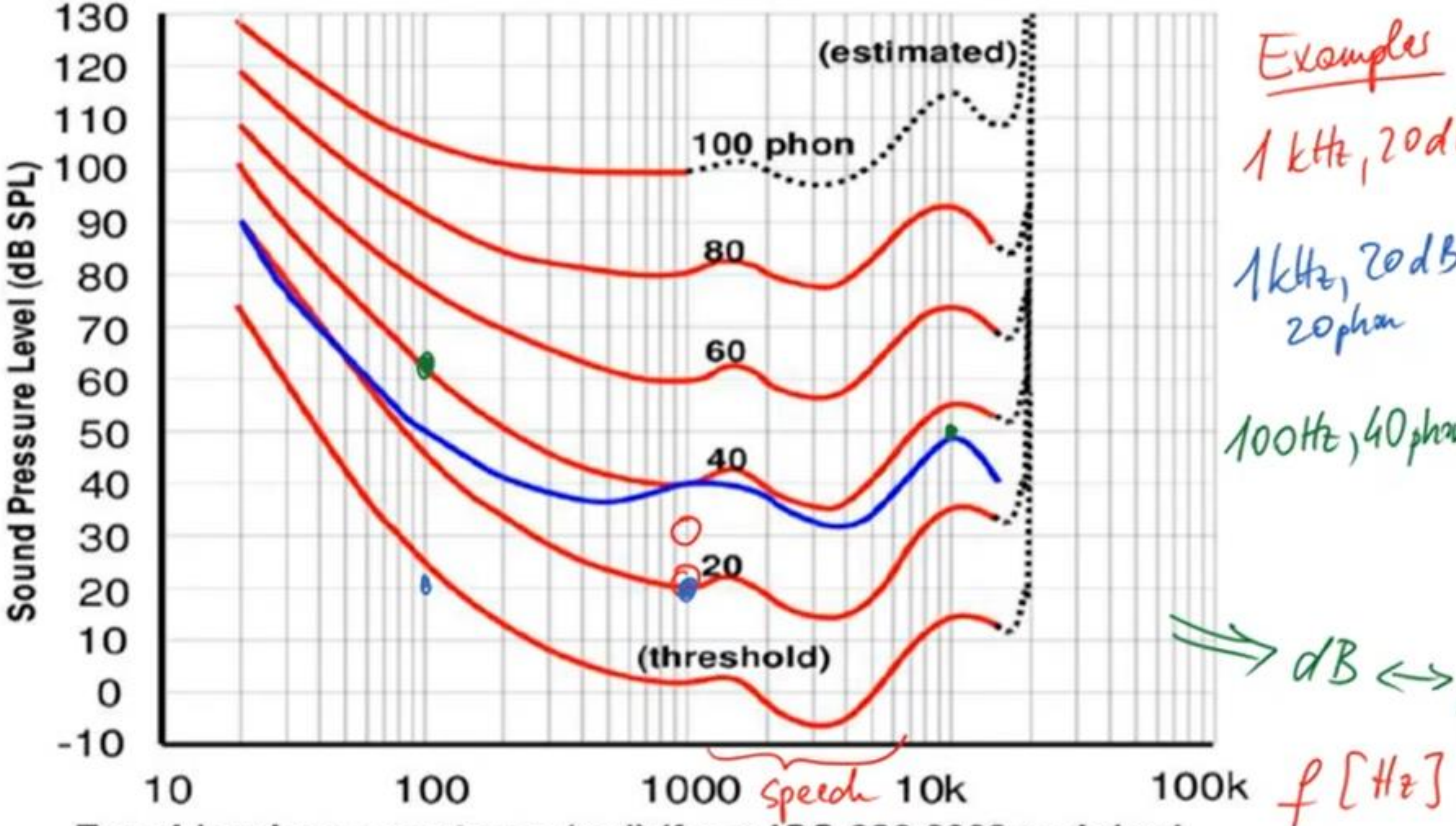


Source: http://www.effectrode.com/wp-content/uploads/fletcher_munson_chart.jpg

Equal loudness contours



Source: OP



Equal-loudness contours (red) (from ISO 226:2003 revision)
Original ISO standard shown (blue) for 40-phons

Source: OP

17

Examples which is louder

1 kHz, 20dB < 1kHz, 30dB

1kHz, 20dB > 100Hz, 20dB
20phon 0phon

100Hz, 40phon > 10kHz
50dB
37phon

$\Rightarrow$ dB $\Leftrightarrow$ phon conversion

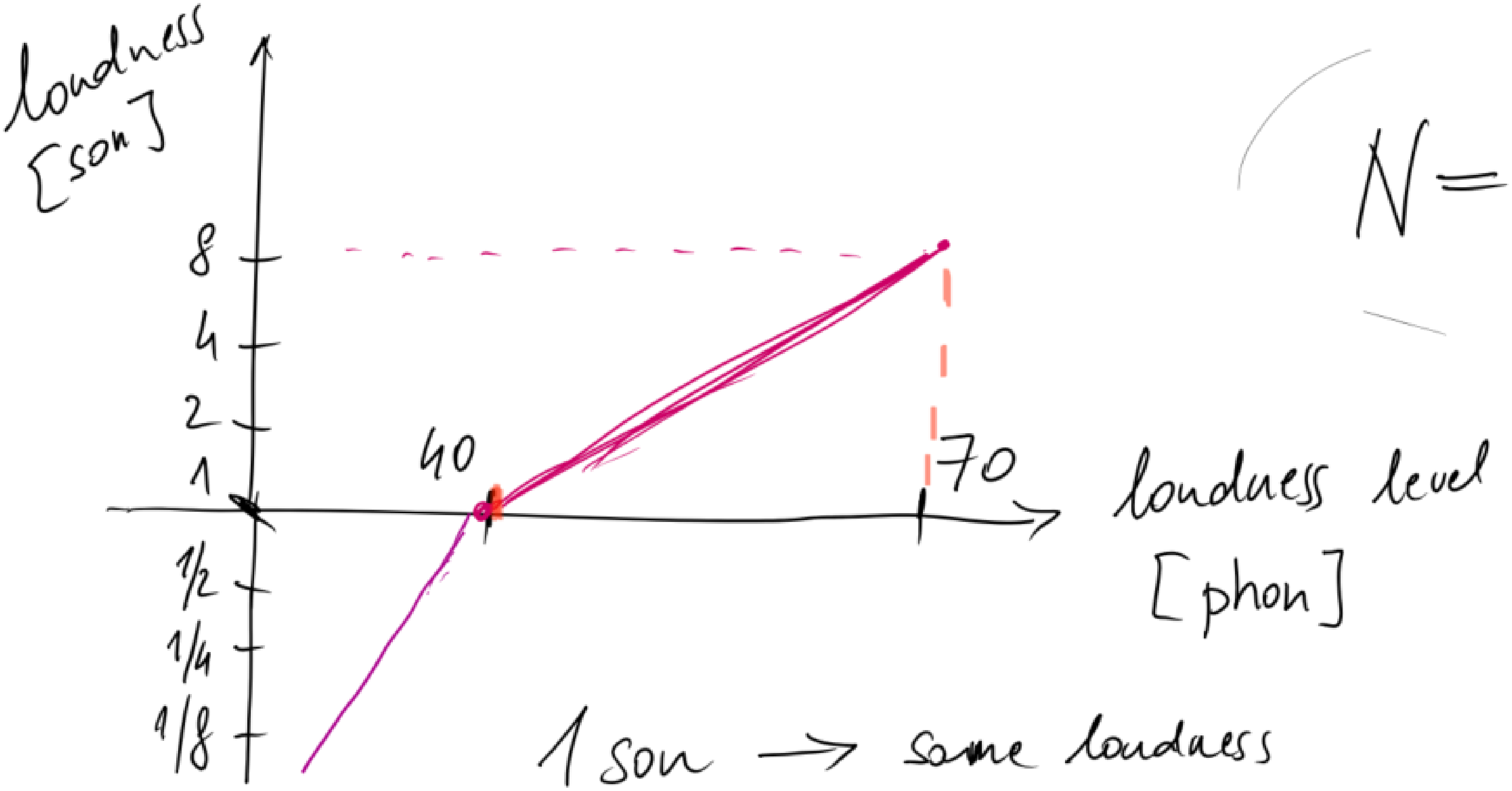
Source: OP

Loudness

Loudness (relative loudness)

perceptual correlate of intensity

how loud a sound is perceived compared to another?



$$N = 2^{\frac{L_N - 40}{10}}$$ (between 40-70 phones)

40 phon → 1 son

70 phon → 8 son

1 son → same loudness

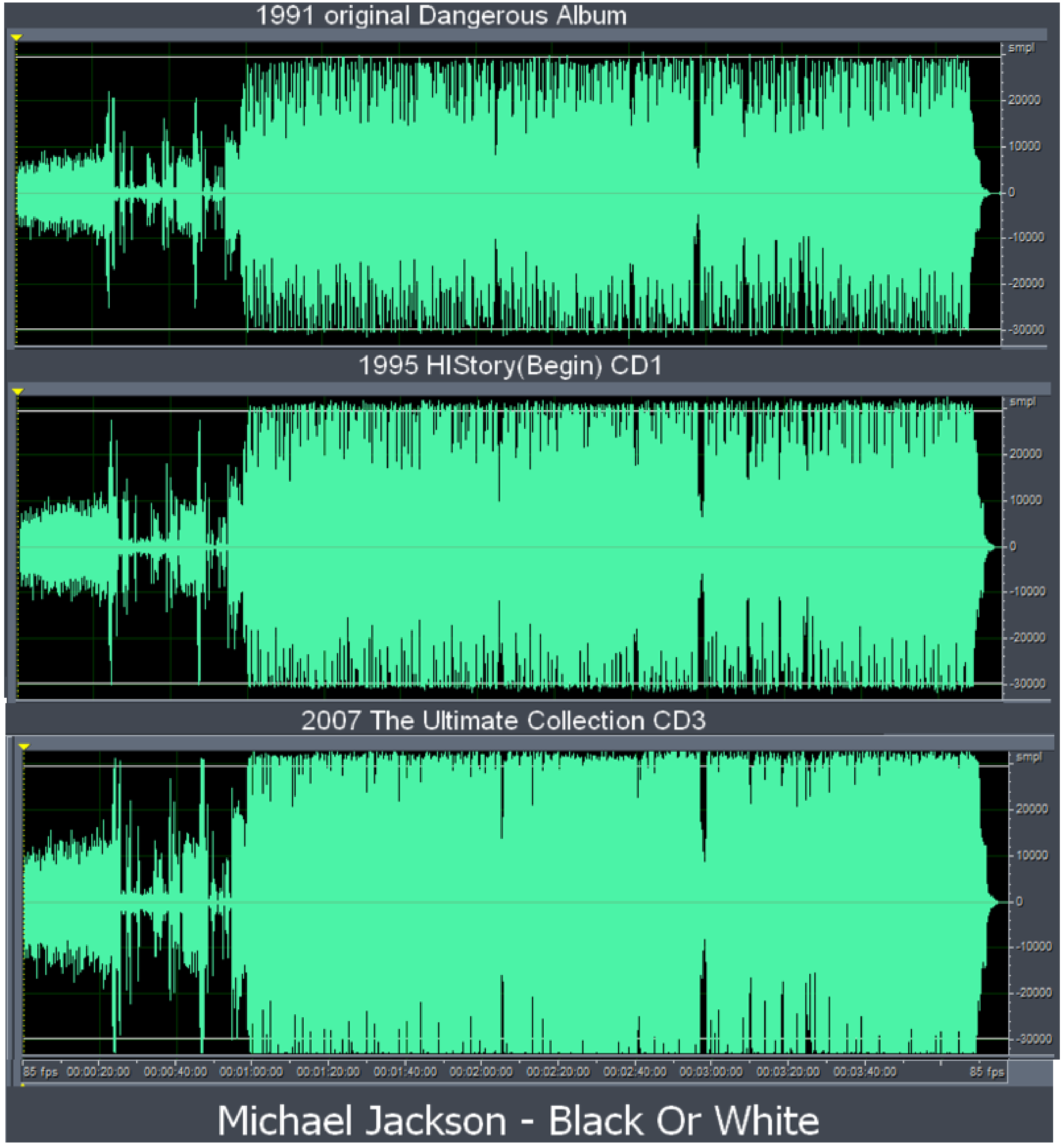
Loudness

- various frequencies at a constant SPL
  (the perceived loudness of tones varies at equal sound intensity)

- which tone sounds twice as loud as the reference tone?
  - reference tone + same tone 5 dB higher
  - reference tone + same tone 8 dB higher
  - reference tone + same tone 10 dB higher

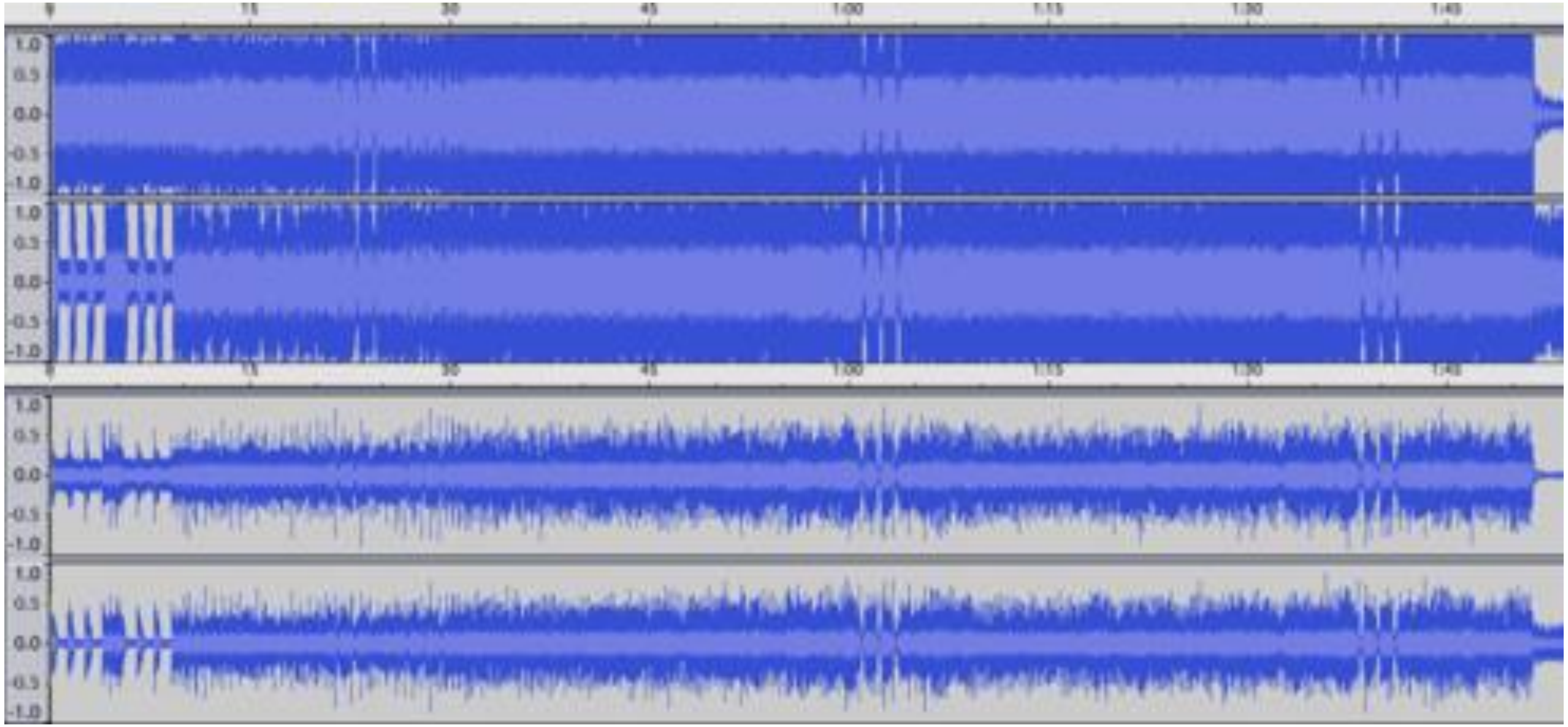
Source: http://www.ece.uvic.ca/~elec499/2003a/group09/p/demos.htm

Loudness war



Source:
https://en.wikipedia.org/wiki/Loudness_war

Loudness war
Metallica: Death Magnetic



Source:
https://en.wikipedia.org/wiki/Loudness_war

Loudness war

Source: https://www.youtube.com/watch?v=3Gmex_4hreQ

Loudness war

Some of the albums that have been criticized for their sound quality include the following:

Artist	Album
Arctic Monkeys	*Whatever People Say I Am, That's What I'm Not*[6]
Black Sabbath	*13*[57]
Bob Dylan	*Modern Times*[40]
	Together Through Life[40]
Christina Aguilera	*Back to Basics*[3]
The Cure	*4:13 Dream*[58]
Depeche Mode	*Playing the Angel*[59]
The Flaming Lips	*At War with the Mystics*[6][note 3]
Led Zeppelin	*Mothership*[60]
Lily Allen	*Alright, Still*[61]
Los Lonely Boys	*Sacred*[3]
Nine Inch Nails	*Pretty Hate Machine (2010 Remaster)*[62]
Metallica	*Death Magnetic*[63][note 4]
Miranda Lambert	*Revolution*[64]
Oasis	*(What's the Story) Morning Glory?*[6]
Paul McCartney	*Memory Almost Full*[65]
Paul Simon	*Surprise*[66]
Pearl Jam	*Ten* (2009 remaster)[67][68][69]
Queens of the Stone Age	*Songs for the Deaf*[6]
Red Hot Chili Peppers	*Californication*[3][6]
Ghost	*Infestissumam*[70]
Rush	*Vapor Trails*[71]
The Stooges	*Raw Power* (1997 remaster)[66]

Source:
https://en.wikipedia.org/wiki/Loudness_war

Dynamic range compression artistic effect

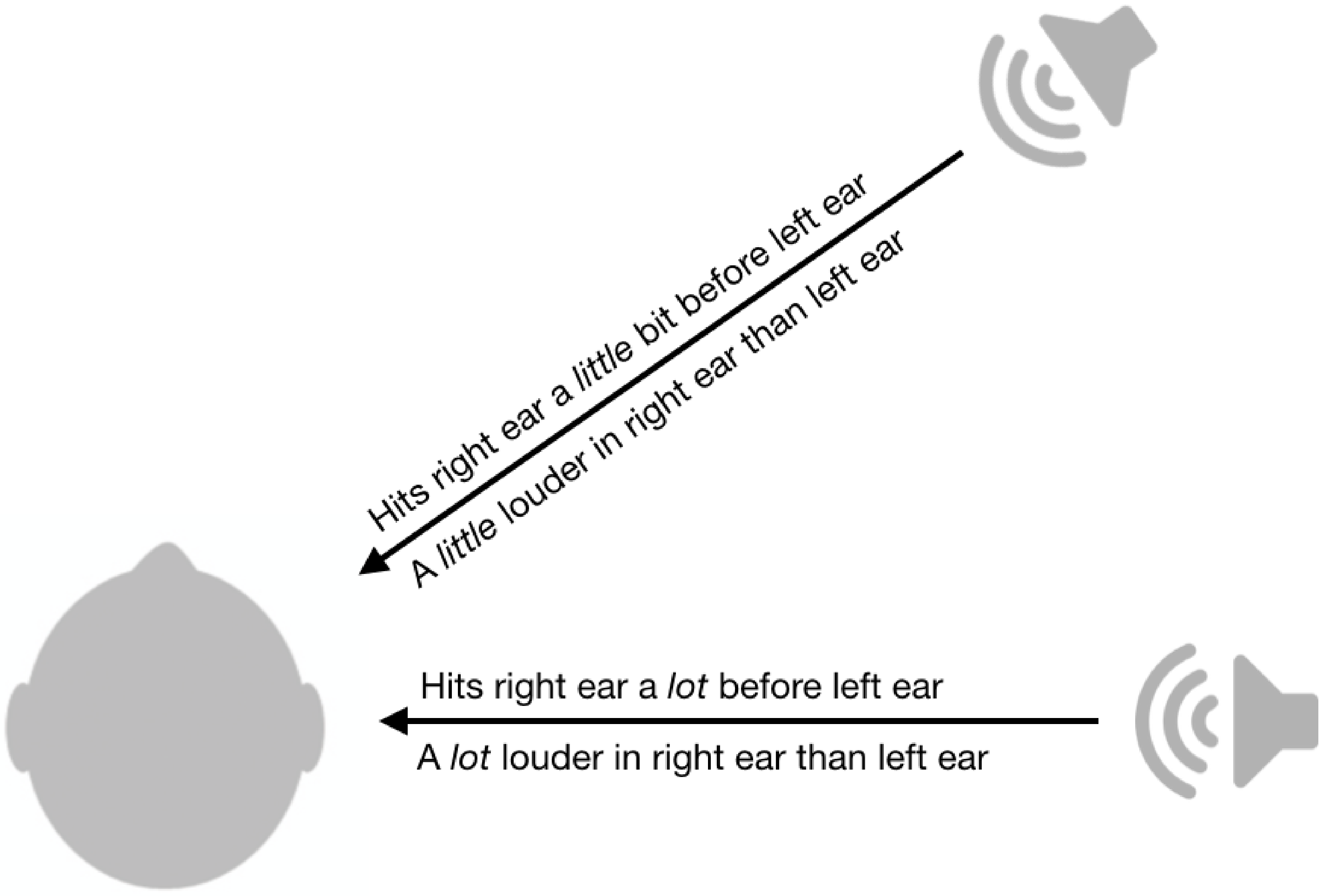
Listen at around 0:43 for the bass drum; you'll heard the rest of the track's volume drop.

Source: http://www.howtogeek.com/57903/htg-explains-how-does-dynamic-range-compression-work/
https://www.youtube.com/watch?v=RIZdjT1472Y

Spatial parameters, direction of sound

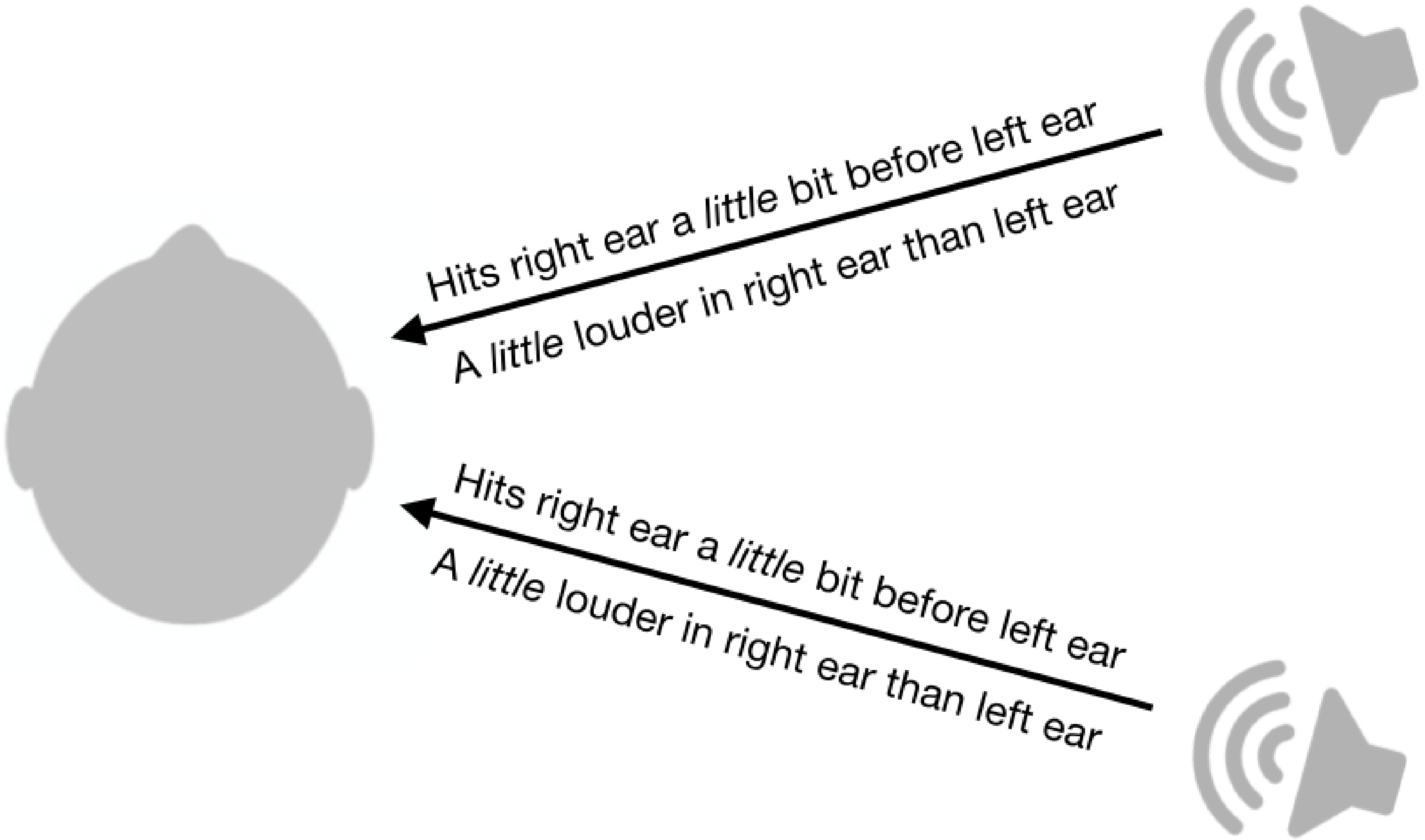
- horizontal plane
  - low freq. (wavelength < size of head, 50-1600Hz)
    phase difference
  - high freq.
    (w.l. > size of head, 1600Hz-...)
    intensity difference
- vertical plane.

Hearing / directions…



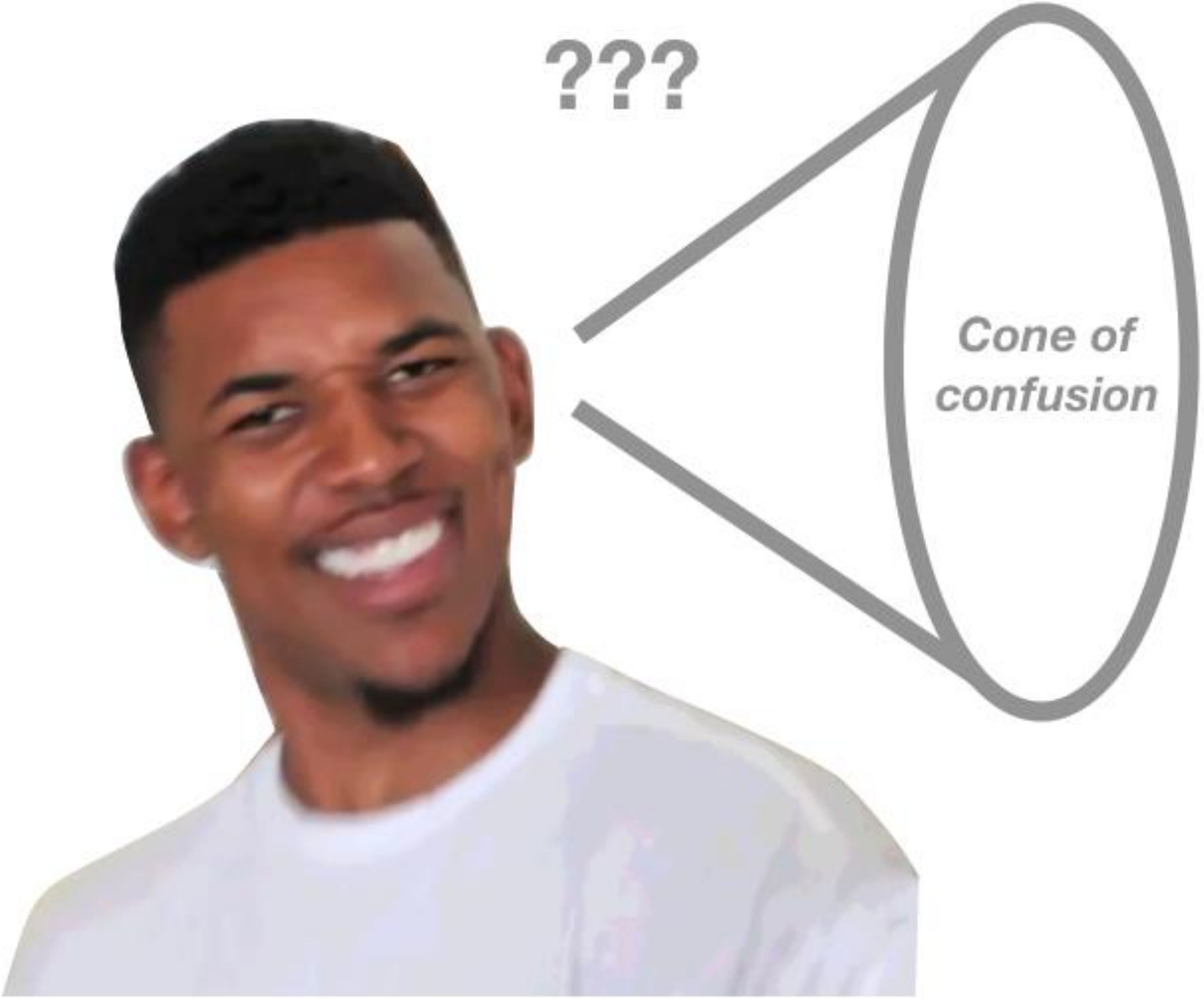
Source: https://chris-said.io/2018/08/06/cone-of-confusion/

Hearing / directions…



Source: https://chris-said.io/2018/08/06/cone-of-confusion/

Hearing / directions / ambiguities



Source: https://chris-said.io/2018/08/06/cone-of-confusion/

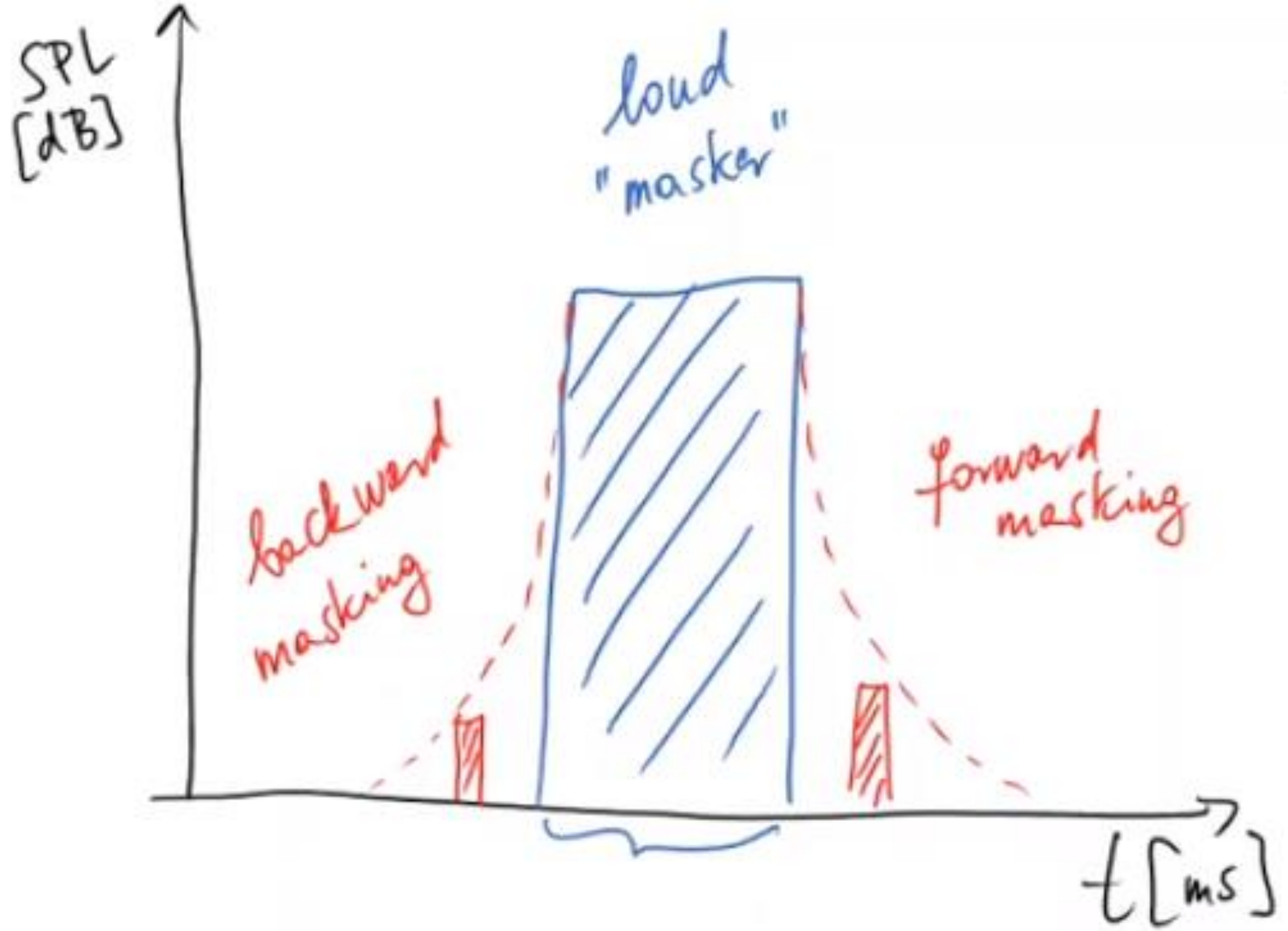
Spatial hearing
„hearing throne”

Source: https://auditoryneuroscience.com/book/export/html/15

Masking

- Time domain masking
- Frequency domain masking
- Directional masking

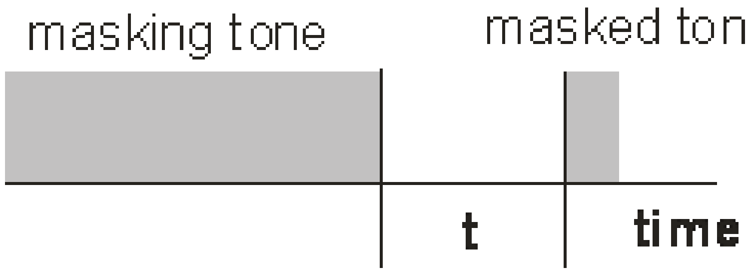
Time domain masking



a) forward : several 10ms

b) backward : several ms

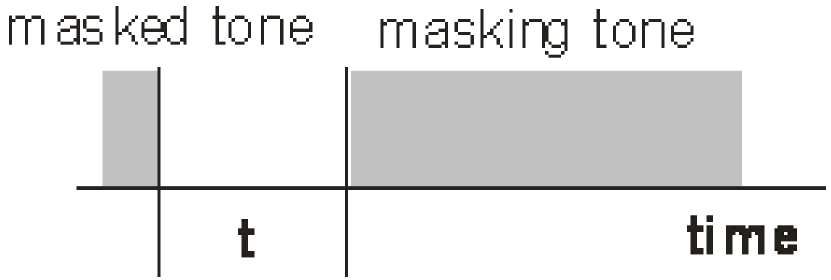
Time domain masking - Forward



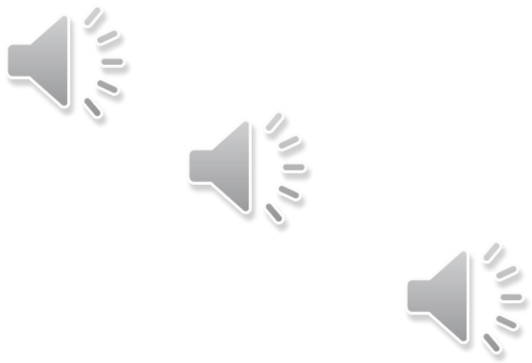
- masking tone + tone that is semitone down
  - with a 100 ms delay in between
  - with a 10 ms delay in between

Source: http://www.ece.uvic.ca/~elec499/2003a/group09/p/demos.htm

Time domain masking - Backward

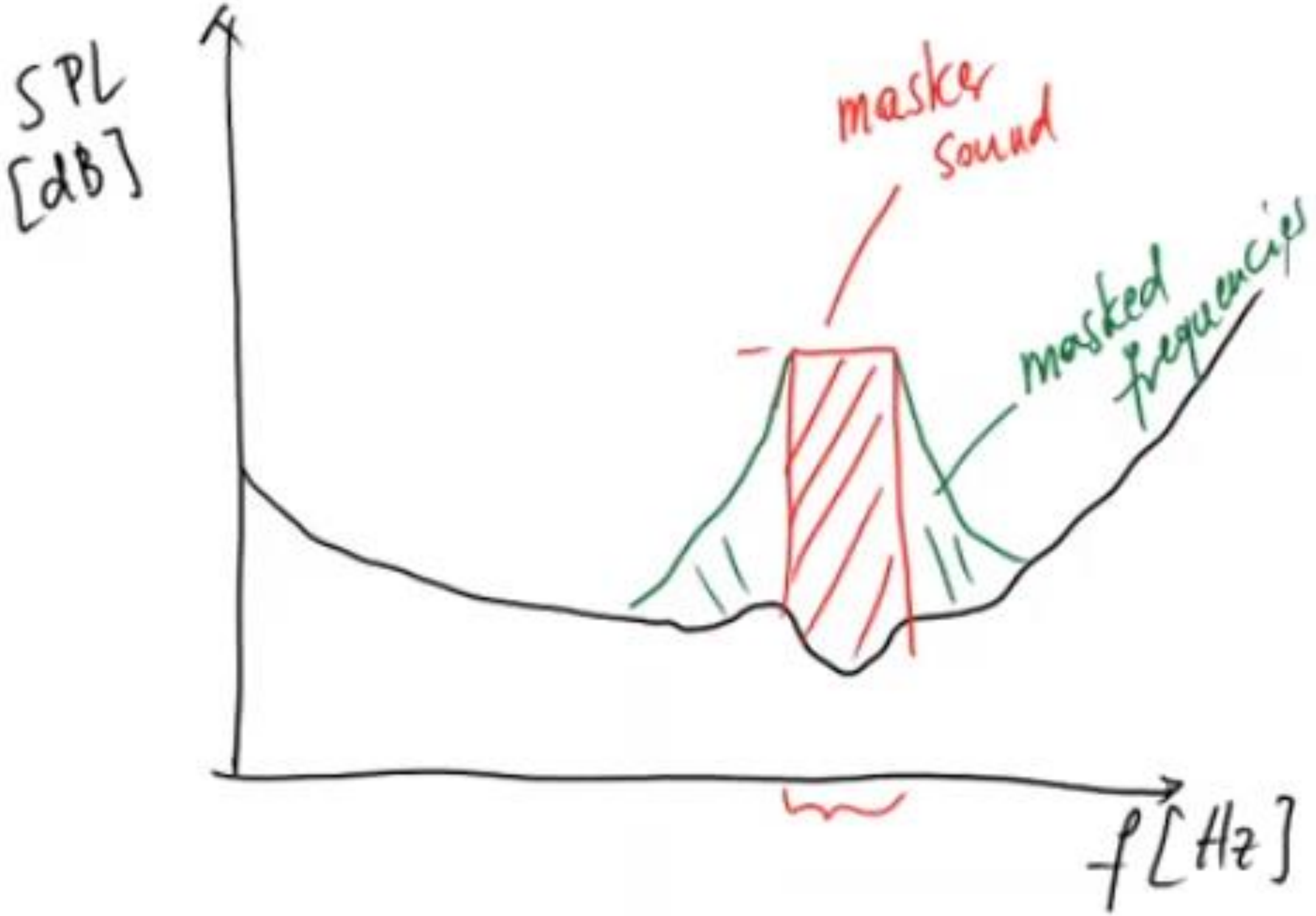


- initial tone is going to be masked by the tone that follows
  - delay: 100 ms
  - delay: above 10 ms
  - delay: below 10 ms



Source: http://www.ece.uvic.ca/~elec499/2003a/group09/p/demos.htm

Frequency domain masking

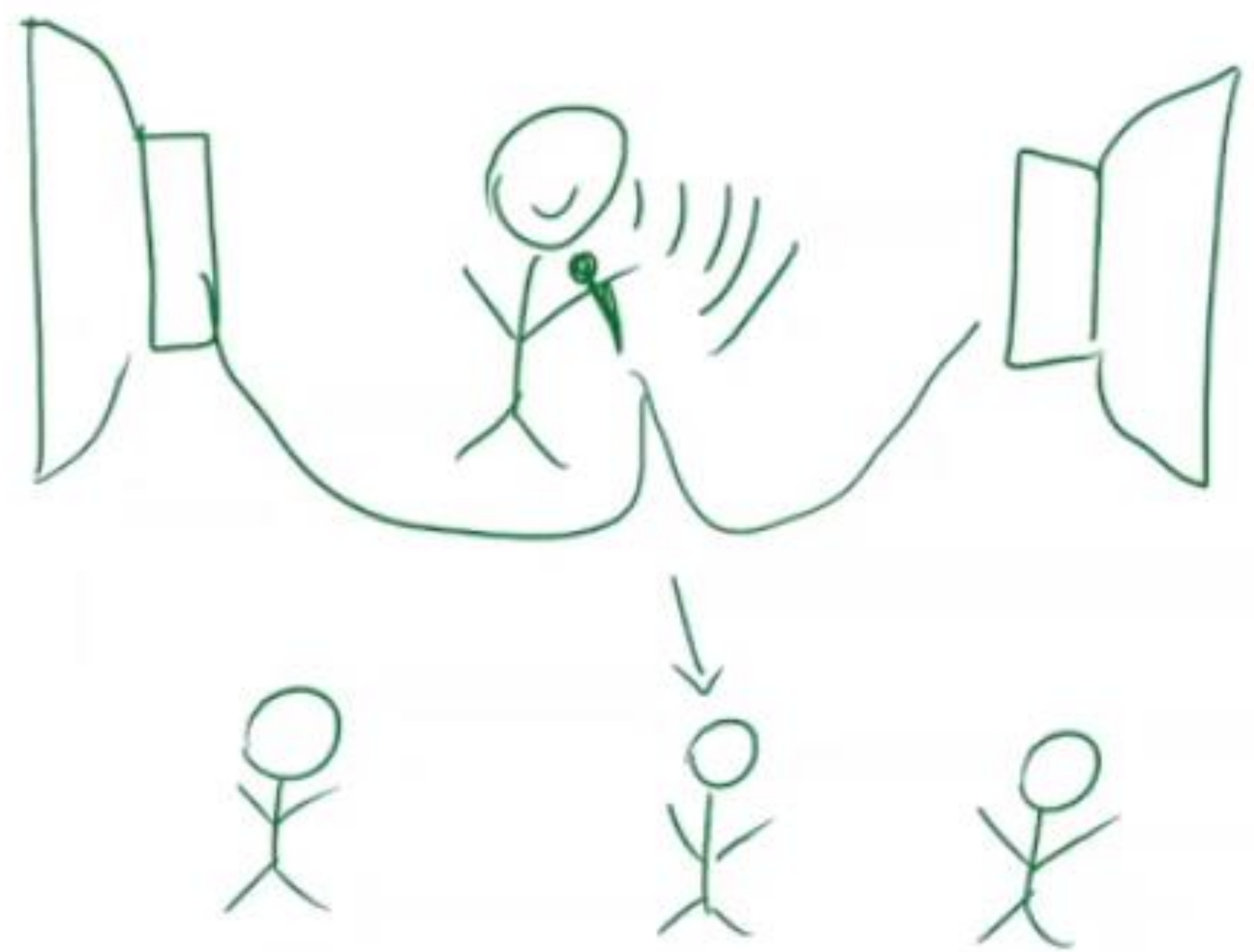


Frequency domain masking

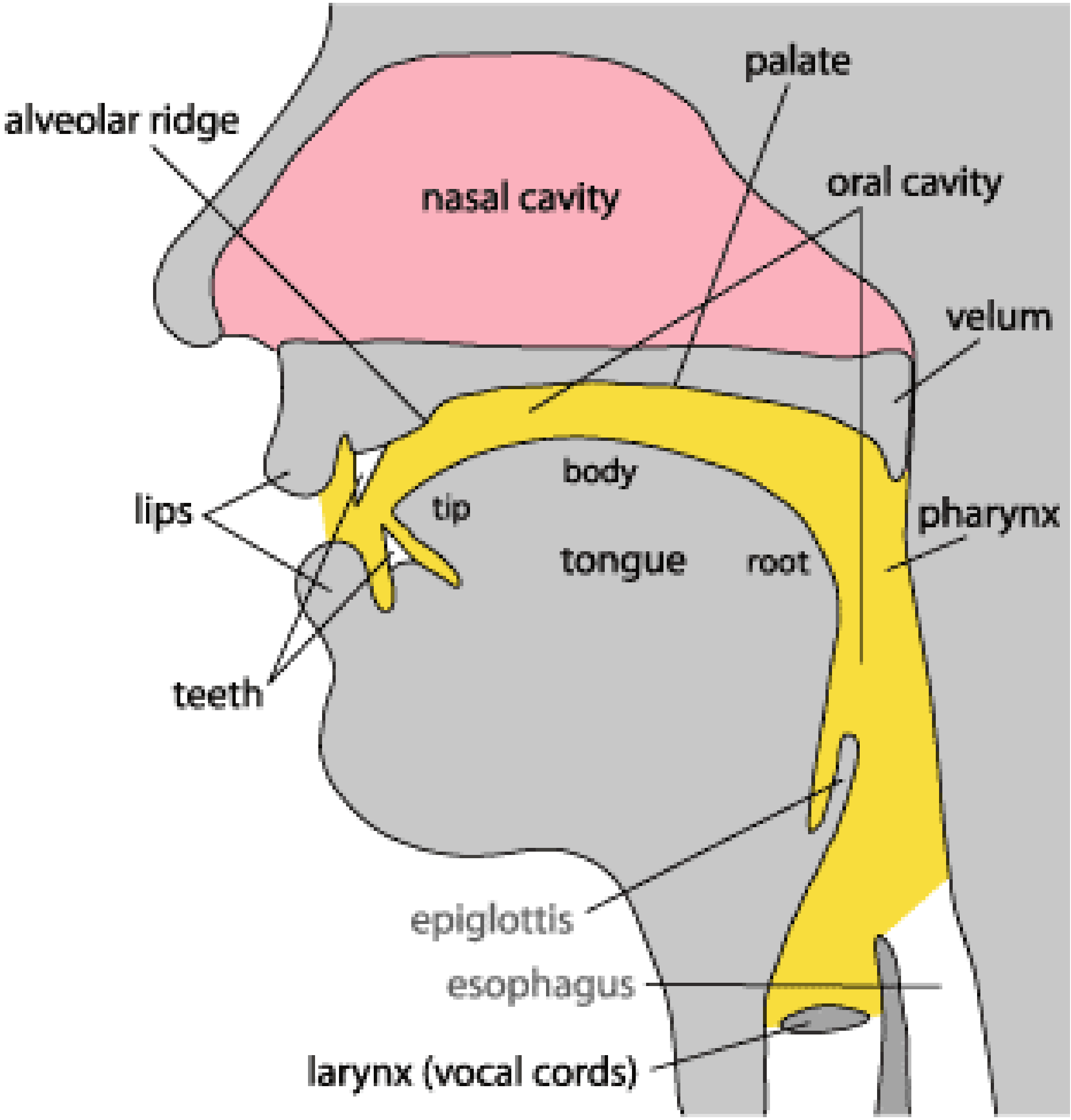
- Pure tones mask higher frequencies better than lower frequencies
  - Mask high freqs
  - Mask low freqs

Source: http://www.ece.uvic.ca/~elec499/2003a/group09/p/demos.htm

Directional masking

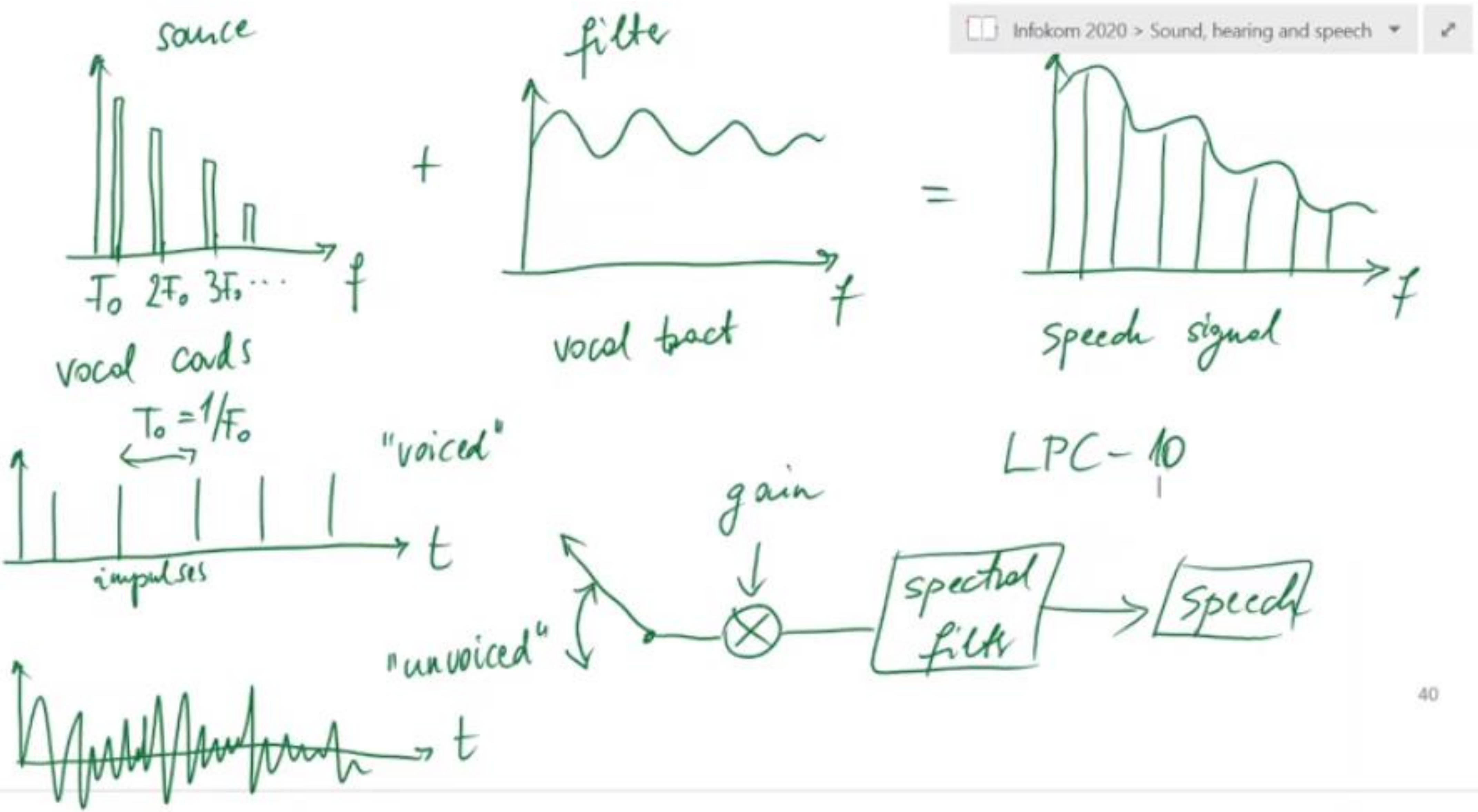


The organs of speech



Source: http://www.englishbaby.com/lessons/3201/member_submitted/vocal_organs_of_speech

Source-filter model of speech production



Linear Predictive Coding

- Bishnu Atal
  - born in India in 1933
  - Bell Labs, AT&T Labs Research
  - 1960's: ‚pulses'

Source: https://www.youtube.com/watch?v=bs5_cuYfoRI

Waveform coding vs Speech coding

- sampling & quant.
- any kind of signal
- store every sample
- e.g. $f_s = \underbrace{8\,kHz,\ 8\ bit}_{64\ kbit/s}$ quantization

LPC-10

GSM

- analyse speech to spectral parameters
- use knowledge about speech production
- high compression

as low as 2 kbit/s

20–30 × compression

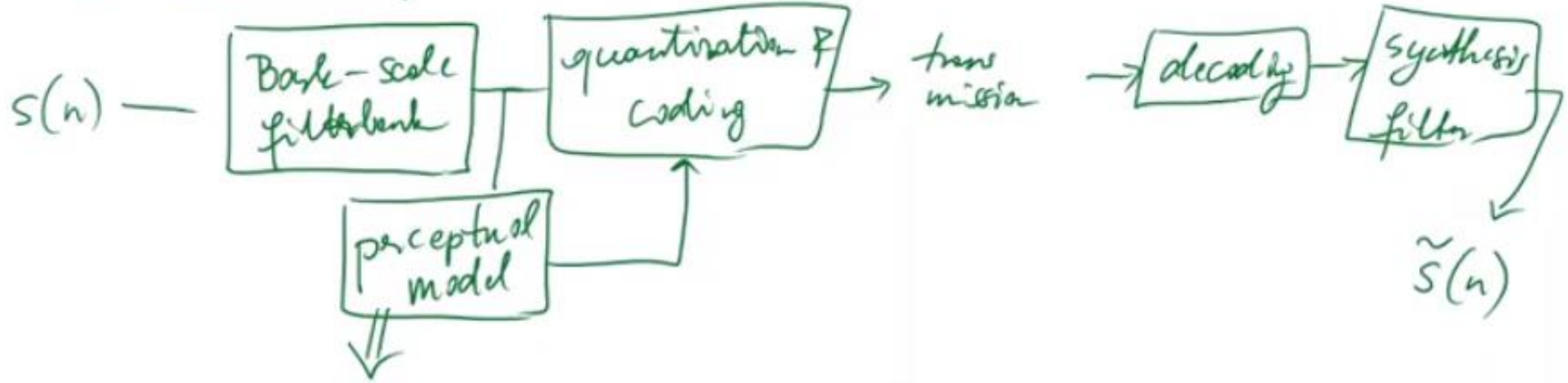
- Waveform coding
  - Original (64 kbps)
  - ADPCM (32 kbps)

- Linear Predictive Coding
  - CELP (4800 bps)
  - LPC-10 (2400 bps)

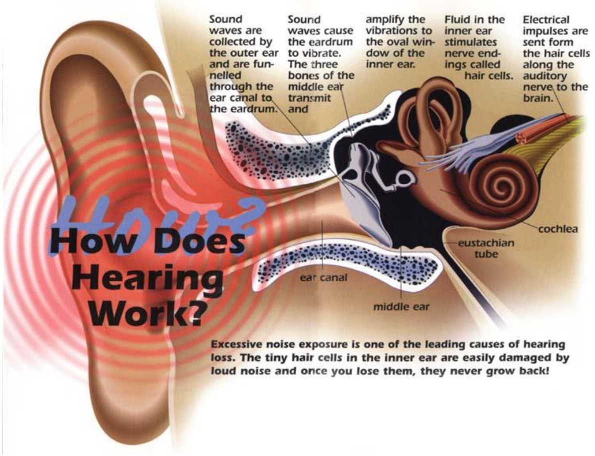
Source: http://www.data-compression.com/speech.html

Perceptual / subband coding

– masked components can be left out



e.g. MP3, MPEG1-Layer 3



The END

Infocommunication
Sound and hearing

Dr. Mohammed Salah Al-Radhi

Dr. Tamás Gábor Csapó

malradhi@tmit.bme.hu