Digital Audio Processing
5. An Introduction to Human Hearing

Nicolas Obin nicolas.obin@sorbonne-universite.fr

2021

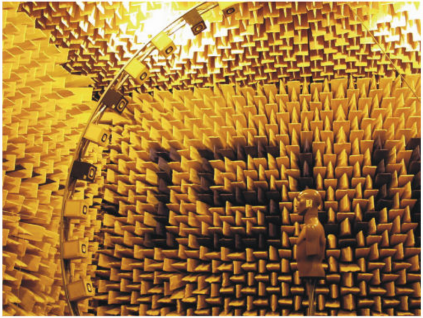
Sorbonne Université

Introduction

Sound Perception

Sound perception is the study of sound perception by humans, which covers:

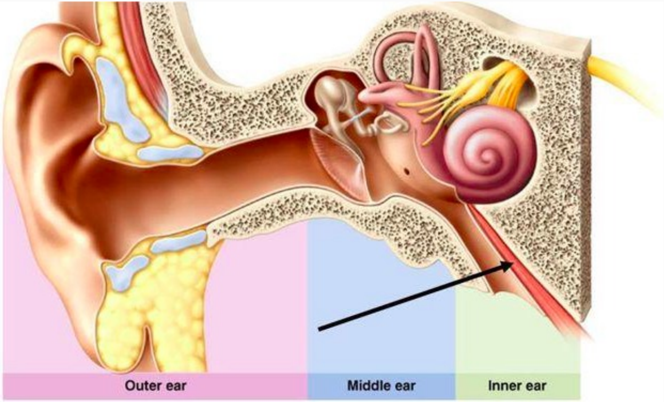
- Psycho-acoustics: hearing as a phycological/physiological system (physics).
- Cognition: hearing as a cognitive process (experience, memory).



- Further readings:
  Brian Moore, An Introduction to the Psychology of Hearing. 1977.

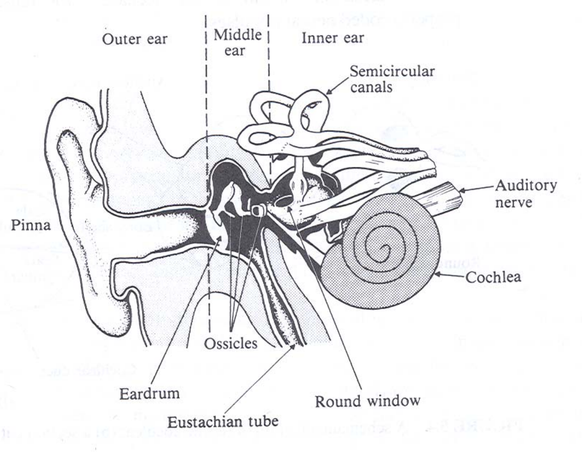
The Ear

The Ear



The human ear.

The Ear



The human ear in details.

The Ear

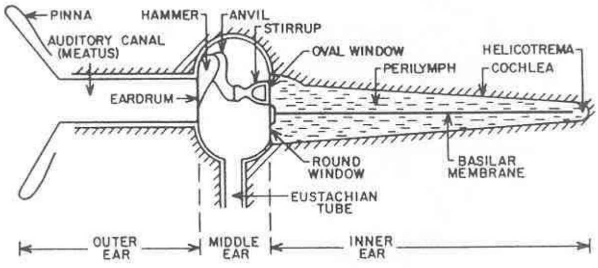
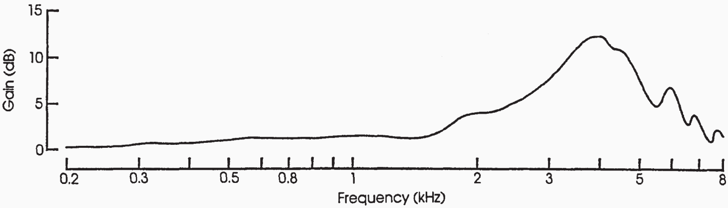


FIG. 1. Schematic diagram of the human ear, with the cochlea uncoiled.

Schematic illustration of the human ear.

Outer Ear

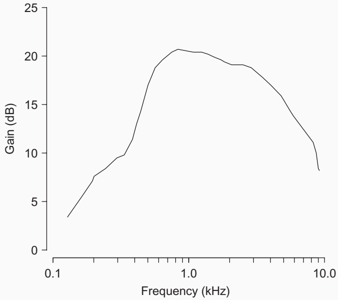
- Anatomy: pinna and auditory canal.
- Function: transmission of the external air pressure vibrations through the ear.
- Signal: band-pass resonant filter which a resonance around 4 kHz.



Transfer function of the outer ear.

Middle Ear

- Anatomy: ear drum (aka tympani) and ear bones.
- Function: transformation of the sound pressure vibrations into mechanical vibrations (inverse of a loudspeaker) that are trasmitted to the inner ear.
- Signal: band-pass filter which a broad resonance around 1 kHz.



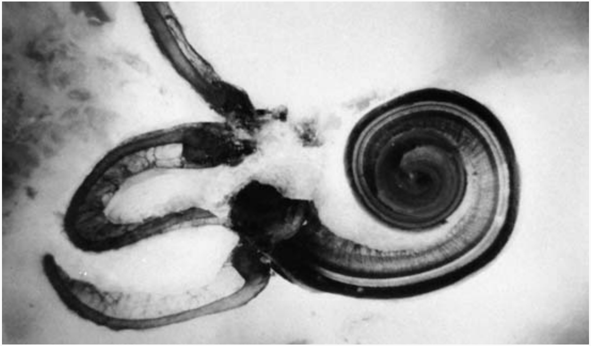
Frequency response of the middle ear.

Inner Ear

The inner ear is the main part of the ear.

- Anatomy: cochlea, basilar membrane, hair cells, auditory nerves.
- Input: mechanical vibrations (middle ear).
- Intermediate: acoustic vibrations in a fluid (cochlea, hair cells).
- Output: electric signal (auditory nerve)
- Signal: a filter-bank of resonating oscillators at different frequencies.

Cochlea



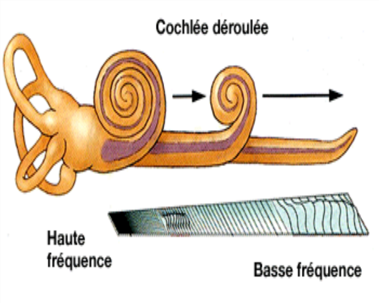
Picture of a cochlea.

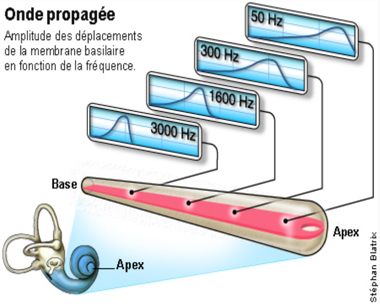
Basilar Membrane

The basilar membrane is basicaly equivalent to a bank of oscillators, each responding to a specific frequency range.

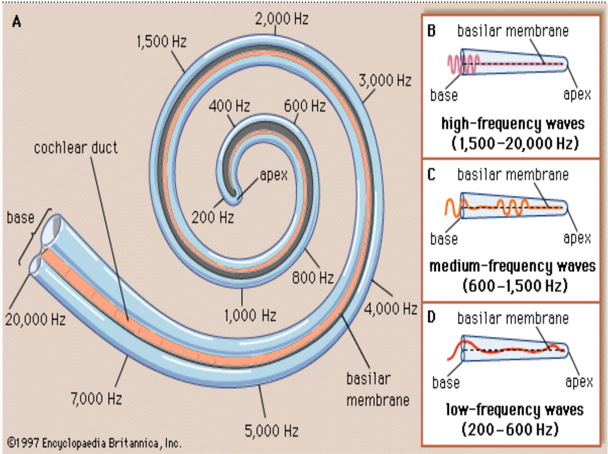
Physically, the membrane can be represented as a succession of rectangular plates with increasing stiffness.

- Base: high stiffness, high-frequency resonance.
- Apex: low stiffness, low-frequency resonance.





Basilar Membrane

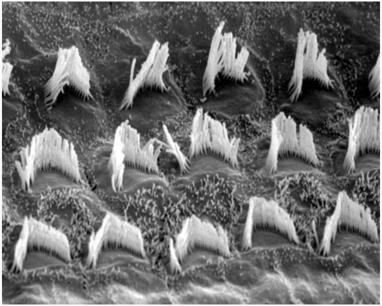


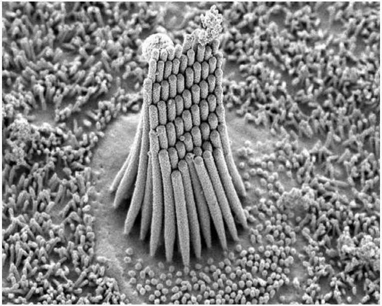
Frequency response of the basilar membrane.

Hair cells

The vibration of the basilar membrane is transmitted to the hair cells located around the spatial region being excited.

The hair cells are displaced depending on the amplitude and the frequency of the excitation.

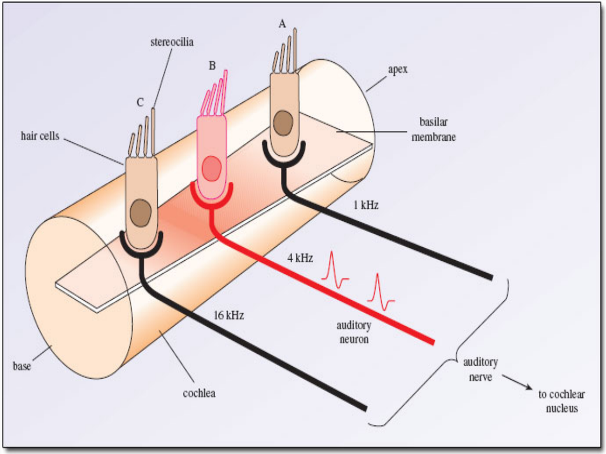




Picture of hair cells in the Corti organ.

Auditory Nerves

Hair cells movements are tranduced into electrical signal, and trasmitted to the auditory nerve.



Auditory Cortex

The signal travels along the auditory nerve to the auditory cortex in the brain.

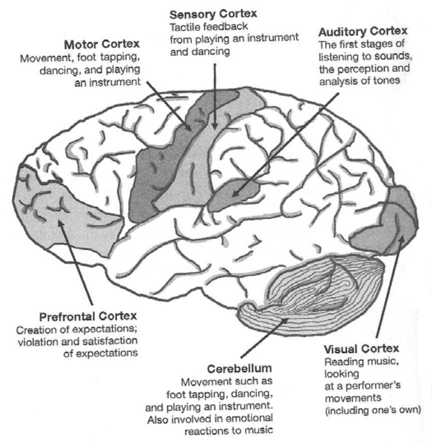


Illustration of the brain Cortex.

The Perception of Sounds

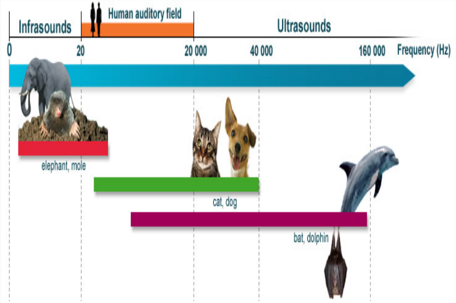
The Audible World of Sounds

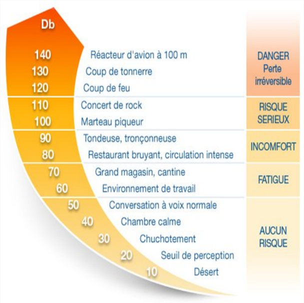
The human hearing is limited:

- frequency range: 20 Hz - 20 kHz
- intensity range : 0 dB (silence) - 120 dB (pain/destruction)

Some audio coding facts:

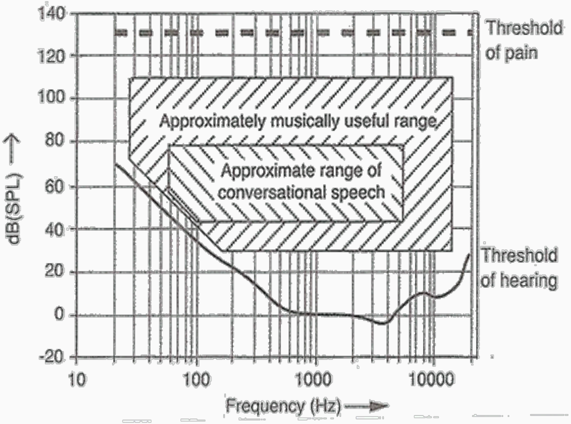
- sampling frequency: 44.1 kHz (0 Hz - 22.05 kHz)
- dynamic range : 16/24 bits (96/144 dB)





Auditory Sensitivity

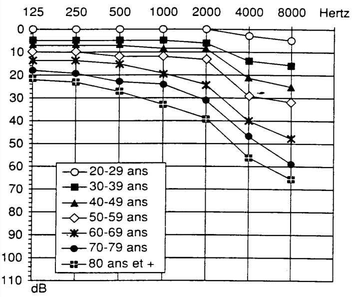
The auditory sensitivity depends on the frequency of a sound.



Hearing threshold, hearing ranges, and pain threshold as a function of the frequency.

Hearing Damages

The auditory sensitivity naturally decreases with aging.



Auditory sensitivity with respect to the age.

Some external causes of hearing damages, and partial/total loss of auditory sensitivity:

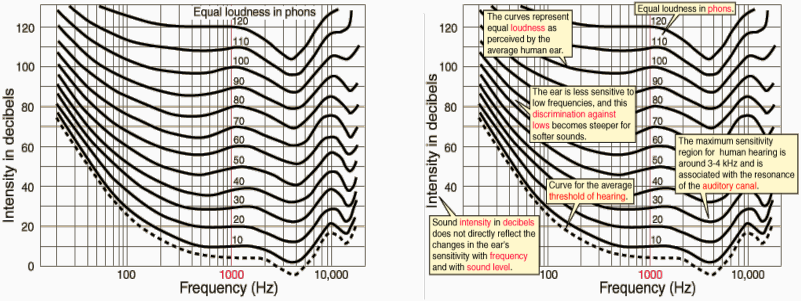
- ponctual exposure to sound levels which exceed the threshold of destruction (140 dB).
- long and repeated exposure to high sound levels (e.g., headphones, concerts, night-clubs)

Loudness

The loudness is the perceived intensity of a sound, expressed in phons/sones.

The loudness of a sound depends on the frequency.

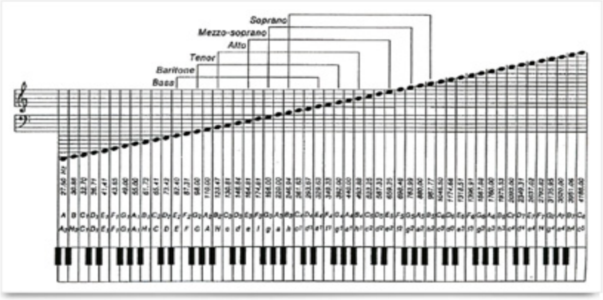
The equal-loudness contours represent the correspondance between measured (dB SPL) and perceived (phons/sones) intensity of a sound.



Equal-loudness contours.

Pitch

The pitch is the perceived frequency of a sound, expressed in Tones.



Pitch scale.

Most people can perceive relative pitch, i.e. pitch intervals between two frequencies.

Some people can perceive absolute pitch, i.e. the frequency of a sound without any apparent pitch reference.

Auditory Masking

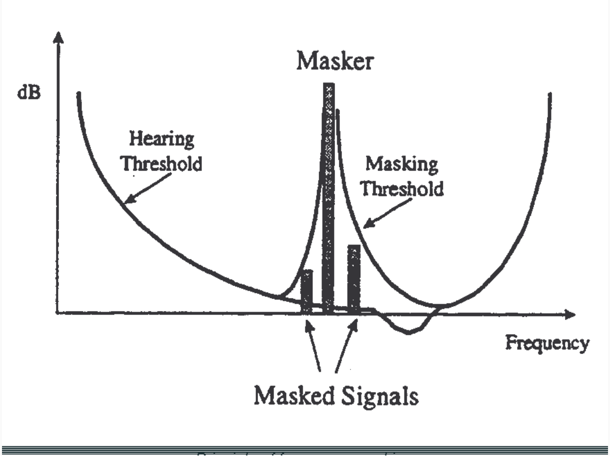
Auditory Masking

Masking is the process by which the detection threshold of a sound (called "the masked signal") is increased by the presence of another sound (called "the masker signal").

Auditory Masking.

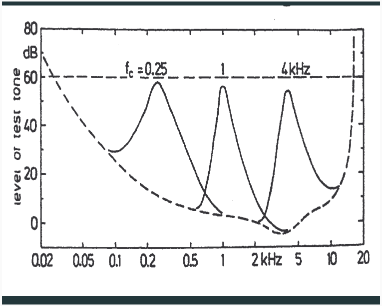
Frequency Masking

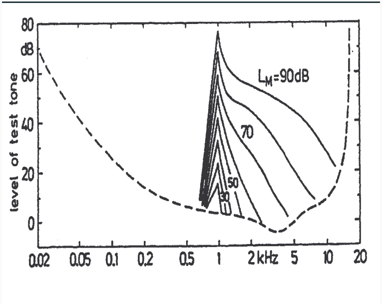
Auditory masking acts in the frequency domain: in a complex sound, a frequency masks surrounding frequencies.



Frequency Masking

The frequency mask depends on the frequency and the intensity of the masker.

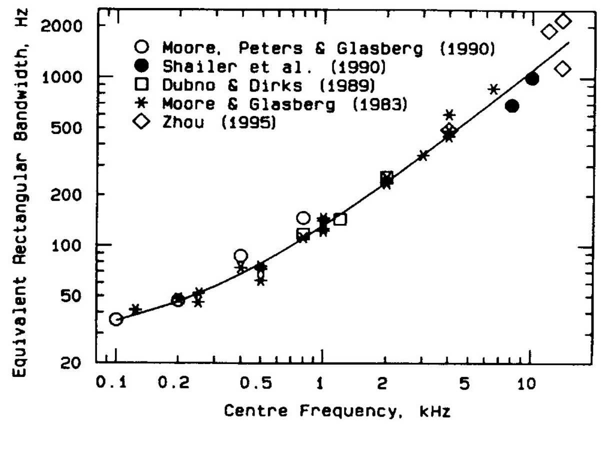




Frequency masking as a function of frequency and intensity.

Frequency Selectivity

The frequency masking is directly relate to the frequency selectivity, which is the ability to resolve two frequencies of a complex sound.



Frequency selectivity as a function of the frequency.

Frequency Scales

Frequency masking can be interpreted as fuzzy filters whose bandwidth depends on the frequency.

- high frequency selectivity for low frequencies
- low frequency selectivity for high frequencies

Several frequency scales exist: Bark scale, ERB scale, and by a extension the Mel scale.

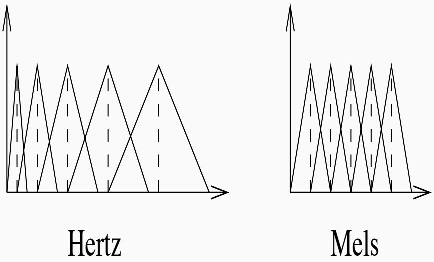
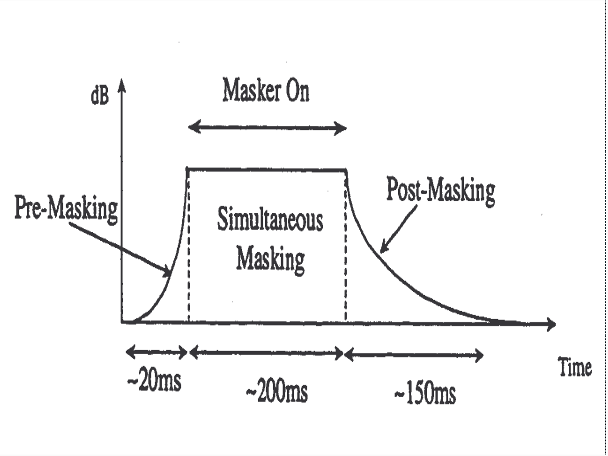


Illustration of the Mel frequency scale.

Temporal Masking

Auditory masking also acts in the time domain: a signal masks the past and the future of a sound.



Principle of temporal masking.