Music Perception and Cognition

Introduction

Hearing: It's mostly for survival. It helps overcome some of our limitations of other senses such as vision.

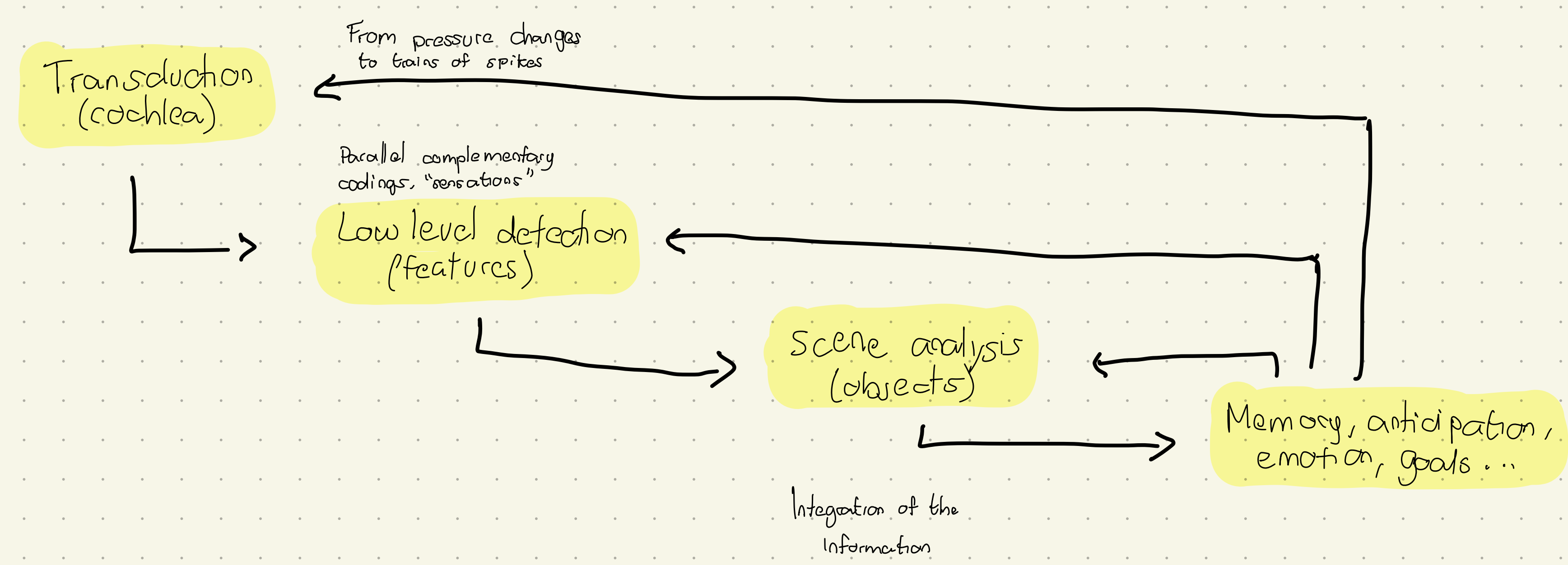
Listening: Putting attention. Gain information that becomes the basis for taking decisions, for example.

Understanding: Combination of both. You match what you have heard and what you listened to some pre-existing knowledge. This way you make sense of information.

THE PSYCHOLOGY OF AUDITION

Physical properties	Sensations	Musical properties
- Frequency	- Pitch	- Melodies
- Amplitude	- Loudness	- Dynamics
- Waveform	- Timbre	- Voices, instruments, texture, chords
- Duration	- Length	- Figures, rhythm

STAGES OF PERCEPTION



PERCEPTION VS. COGNITION

Bottom up processing: It's a data driven approach stating that the perception directs cognition. It influences in decisions and behavior

Top down processing: The behavior is influenced by conceptual data. The expectations influence perception and behavior

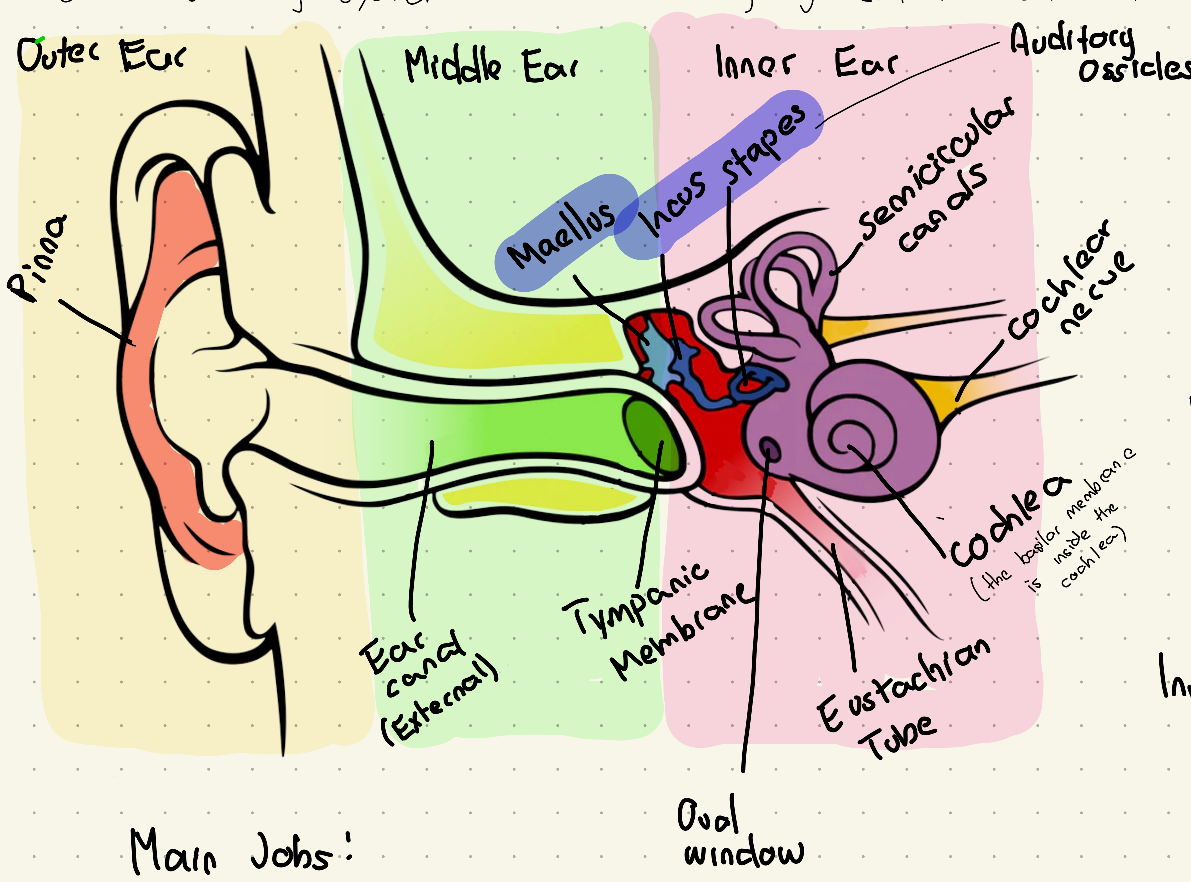
PERCEPTION	COGNITION
Pressure converted to electrochemical patterns	Electrochemical patterns becoming meaningful and adaptive
Transduction and basic sensations	Musicallity, music knowledge
Bottom-up	Top down
Automatic	Consciousness-mediated
Encapsulated	Distributed

SENSORY SYSTEMS

Interceptors: Keep the brain informed about the status of the internal organs
Propioceptors: keep the brain informed about the body position
Exteroceptors: keep the brain informed about the external world. The traditional 5 senses.

THE EA

The auditory system is the sensory system for the sense of hearing



Our hearing system is divided into three parts:

Outer Ear: Its composed mainly by the pinnea and the auditory canal. It's main job is locating, filtering sounds and protection.

Middle Ear: It's composed by the eardrum or tympanic membrane, the Malleus, the Incus and the stapes. It's main job is adequating sound waves and amplificating. It also protects the inner ear (with acoustic reflex)

Inner Ear: It's composed by the cochlea and the semicircular canals. It's main job is the transduction from mechanical to electrical. It also helps with equilibrium

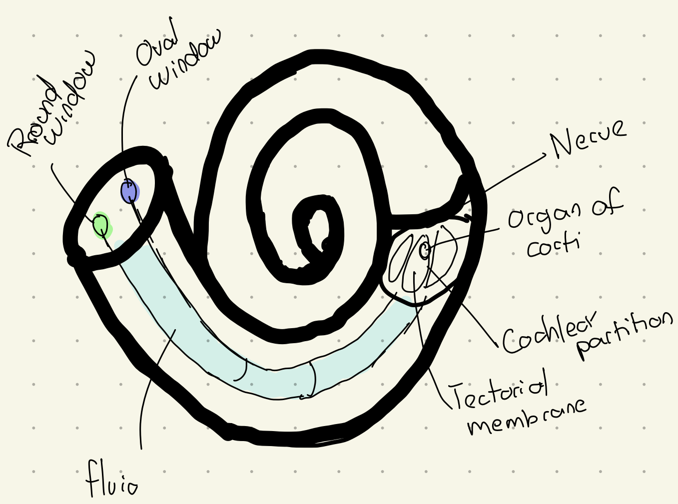
Main Jobs:

Amplification: auditory ossicles
Pressure compensation: eustachian tube
Equilibrium: semicircular canals
Transduction: Cochlea
Sound collection: Pinna macro-structure
Localization: Pinna micro-structure
Protection: wax, inner hair, inner canal, pinna

HOW IT WORKS

First, the soundwaves hit the ear. The preassure waves hit the tympanum and it vibrates. This vibrating causes mechanical energy and this energy transmits the vibrations to the 3 bones located in the middle ear. The stapes transmit the vibrations to the oval window, which is the outermost structure of the inner ear. The vibrations caused by the soundwaves in the oval window create preassure waves in in the fluid that's inside the cochlea (perilymph). Inside the cochlea, upon the basilar membrane, theres' the organ of corti, which contains inner hair cells, which are the ones responsible of the transduction of mechanical waves into electrical signals. Basically, the friction caused by the hair cells and the tectorial membrane (also located inside the cochlea) causes electro chemical impulses

THE COCHLEA:



The cochlea contains the basilar membrane, the tectorial membrane, the organ of corti and the inner hairs.

Depending on where the inner hairs are, they provide information about freqvencies.
The outer part of the cochlea contains the inner hairs that detect higher frequencies

EFFERENT AND AFFERENT NEURONS

The cochlea also contains efferent and afferent neurons. Afferent neurons carry information from sensory receptors of the skin and other organs to the central nervous system. Meanwhile, Efferent neurons carry motor information, away from the central nervous system, to the muscles and glands of the body

THE BASILAR MEMBRANE

The basilar membrane's properties change along its length, on the outside its thicker and narrower and, on the inside, it's thinner and floppier. Depending on the frequency, different different regions of the membrane vibrate and this is how the fluid-filled cochlea can detect the pitch at different regions of the membrane.

THE ORGAN OF CORTI

The organ of corti is the responsible of the transduction. The organ is composed of the hair cells placed above the basilar membrane.
The inner hairs are the primary auditory receptors and exist in a single row.
The intensity of a sound is determined by how many hair cells at a particular location are estimulated

THE NEURONS

Spontaneous stochastic activity
Binary devices
Electro chemical transmissions
Refractory period: time in which a nerve cell is unable to fire an action potential (nerve impulse).

CODING OF ACOUSTIC INFORMATION

Phase-locking: It's the tendency of a neuron to fire action potentials at particular phases of an ongoing periodic sound waveform, such as sinusoidal waveforms that are typically used in psychological studies of the auditory system. Phase-locking means of frequency coding.

Information about sound intensity is coded in 2 ways: rate (firing rate of the neurons) and place (the numer of neurons active). Neural correlates of percieved loudness

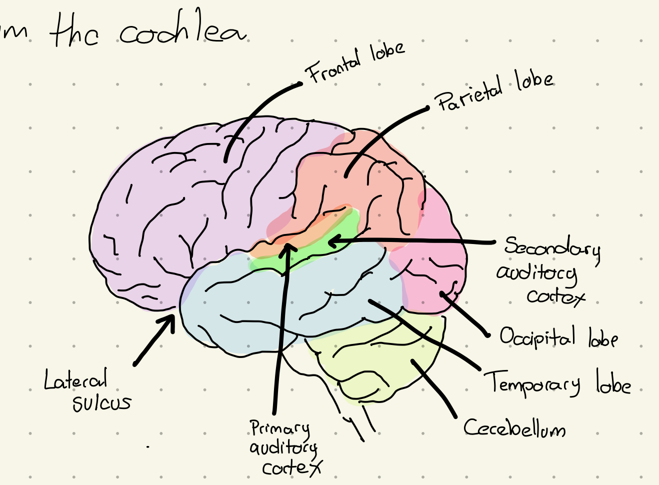
Statistical behaviour of neuron ensambles

AUDITORY PATHWAYS, THE BRAIN AND THE AUDITORY CORTEX

The auditory information travels through the auditory pathways until they reach the auditory cortex. Some processing is done during this process, such as localization and frequency identification.
This processing is distributed accross all neurons involved.
There are two pathways:

Primary auditory pathway: This pathway exclusively carries messages from the cochlea

Non-primary pathway: This pathway carries all types of sensory messages



The primary pathway is quite short and it ends in the primary auditory cortex, which is located in the temporal area, within the lateral sulcus.

The tonotopic organization of the auditory nerve is also preserved through the auditory pathway.
Another way to categorize it is:

What: It's monoaural and it recieves information from only one ear. Concerned with spectral (frequency) and temporal (time) features and hardly concerned spatial aspects. Focused on identifying and classifying different types of sounds.

Where: It's binaural and it recieves information from both ears. Involved in the localization of the stimulus.

The left hemisphere of the brain is specialized in processings such as the temporal structure, the speech and the fast changes, while the right hemisphere is specialized in the pitch, timbre, frequency and slow changes.

The auditory cortex is placed in both hemispheres in the upper side of the temporal lobes. It is divided into the primary auditory cortex, the secondary auditory cortex and the associative auditory cortex

When we listen to music, other cortical regions are activated. When the music is known we activate the motor area. For musical memory and experiences we activate the hypocampus. For emotional responses we activate the Nucleus accumbens amygdala. And for expectations we activate the prefrontal cortex

Each hemisphere of the brain contains four lobes, the frontal lobe (cognitive functions, control of voluntary movement), the parietal lobe (processing information about temperature, taste, touch and movement), the occipital lobe (vision) and the temporal lobe (memories, sound, sight and touch)

Lateralization: functions are performed by distinct regions of the brain

Neural plasticity: The ability of neural networks in the brain to change through growth and reorganization. These changes range from individual neuron pathways making new connections to systematic adjustments (circuit networks changes from learning something new or from practice)

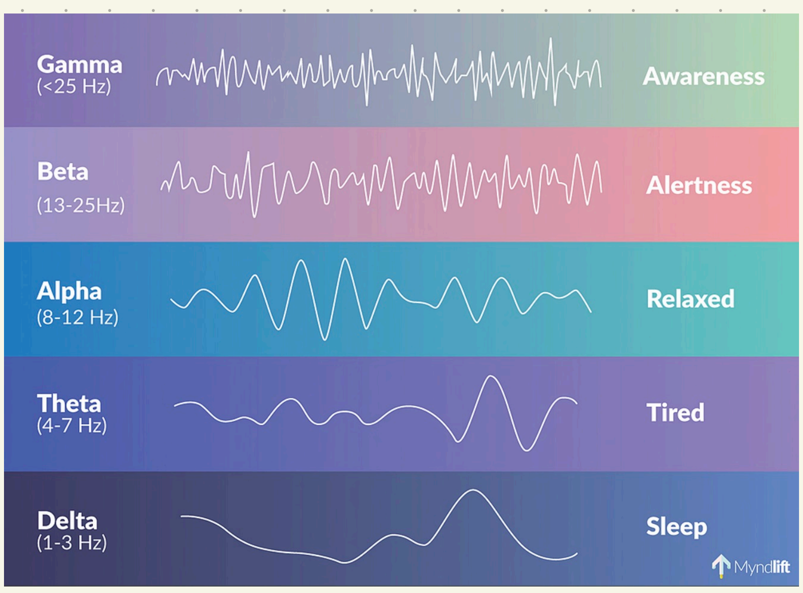
2.- RESEARCH METHODS AND TECHNIQUES

How do we measure music perception and cognition?

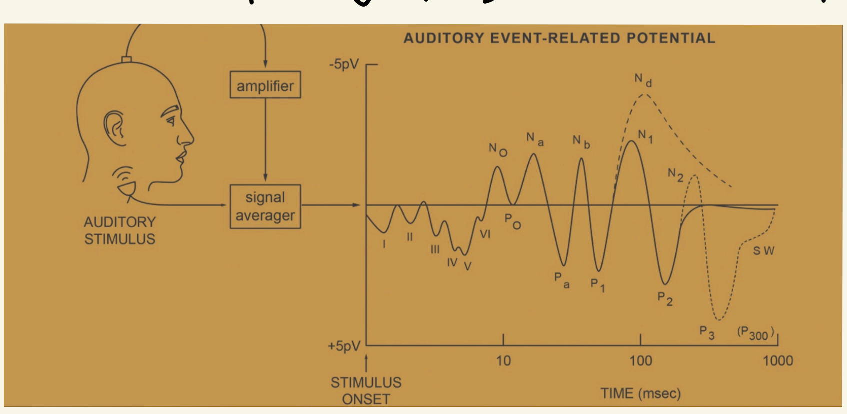
Psychological measures: heart rate, skin conductance, blood pressure, temperature, muscle tension

Behavioral measures: errors resolving tasks, reaction time, relatedness or similarity, choice (forced or not forced), eye-fixation, motion capture, quantified self.

Electro-encephalography (EEG)



Electro-encephalography: event related potentials



A sematic context is set either by speech or by music. Hearing unrelated words generate a negativity peak in centro-parietal electrodes around 400 ms. But when there is a sematic relationship between the prime and the word, the negativity is not observed. The same N400 is observed when the context is set using music and the words no semantic relationships.

ERP: Neural sign that reflects coordinated activity of an ensemble of neurons, observed after certain events or stimuli.

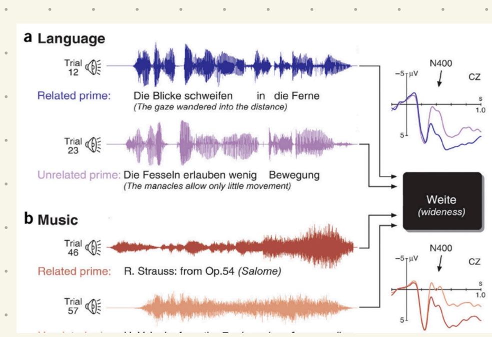
Event related potentials have a good temporal resolution but poor spatial resolution.

Pxxx: positive peaks Nxxx: Negative peaks

MMN: Mismatch Negativity: It reflects the changes in the content of our auditory short-term memory, even though we are not aware of that.

N400: Lack of sematic associations between terms or expressions

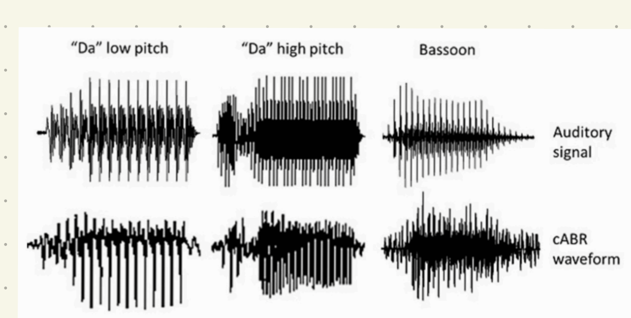
P600: syntactic violations in language and music



Frequency Followed Response (FFR or cABR)

The FFR is an evoked potential generated by periodic or nearly-periodic auditory stimuli. Part of the auditory brainstem response (ABR), the FFR reflects sustained neural activity integrated over a waveform and/or the envelope of the periodic stimuli.

cABR: auditory brainsystem response to complex sounds



Functional Magnetic Resonance Imaging (fMRI)

The fMRI mesures the brain activity by detecting changes associated with blood flow. This technique relies on the fact that cerebral flow and neuronal activation are coupled. When an area of the brain is in use, blood flow to that region also increases.

Changes in oxygen consumption in particular areas of the brain when performing a given task
oxygen nuclei have nuclear magnetic resonance.

Excellen spatial resolution but bad temporal resolution
Non-invasive and innocuous

Position Emission Tomography (PET)

Radioactive contrast substance injected.
Measurements of local blood flow, metabolism, neuro-receptor bindings
Good spatial resolution but poor temporal resolution
Health risk

Diffussion Spectrum Imaging

Mapping the gradient of water (instead of blood) molecule diffussion. Estimation of the trajectories of neuron firing (instead of activity) in the area.

3: PSYCHOPHYSICS OF BASIC SOUND DIMENSIONS : FREQUENCY RESOLUTION

To understand how frequency resolution works, it's important to have an insight into how frequencies are transducted from the cochlea to the auditory cortex.

The basilar membrane acts as a bank of overlapping pass-band filter, the so called auditory filters. Each location along its surface is responding to a limited amount of frequencies. These filters are non-linear, level-dependent, and it's bandwidth and sharpness increases with frequencies. The bandwidth of the auditory filters is called critical bandwidth

The critical bandwidth plays a crucial role in frequency resolution. In fact, when two tones fall into the same auditory filter, the auditory system cannot discriminate between them, and thus code the information as if would be only one. This effect is called unresolved frequencies, and is the base of different effects, such as beating and roughness.

The frequency resolution ability can be interpreted as the filtering capability of the auditory system. It can be studied by the shape of the auditory filter, that is centrated at the frequency component of interest.

The ability to hear frequencies separately is known as frequency resolution or frequency selectivity. When signals are perceived as a combination tone, they are said to reside in the same critical bandwidth.
A complex sound is split into different frequency components and these components cause a peak in the pattern of vibration at a specific place on the cilia inside the basilar membrane. These components are then coded independently on the auditory nerve which transmits sound information to the brain. This individual coding only occurs if the frequency components are different enough in frequency, otherwise they are in the same critical band and are coded at the same place, therefore they are percieved as one sound instead of many.

Frequency resolution occurs on the basilar membrane due to the listener choosing a filter which is centered over the frequency they expect to hear, the signal frequency.

Theres different methods to estimulate the shape of the auditory filter. Normally they use pure tones and are based on the assumption that the subject can detect the tone if the signal-to-noise ratio (SNR) within the auditory filter exceeds a critical threshold. The reduction of the SNR is the reason that problems arise in a noisy environment.

Noise can be interpretated as all sounds different from the signal pure tone. Most methods use the principle of changing the level and/or frequency content of the noise in order to obtain the shape of the filter. The noise added on purpose in the method is called "the masker". and the difference between methods are mainly related to the maskers applied. The auditory filter is nonlinear and it's shape depends on the method and signal level used.

The sensitivity of human hearing depends of the frequency range where the sound is present. We are more sensitive in the low regions and less sensitive at high frequencies.

In the cochlea, sounds are processed in spectral bands, which are independent, such that sounds in separate bands do not interfere with each other, but sounds within the same bands do. The width of these bands is frequency dependent and increases with increased frequencies. This has been approximated by several models such as bark and ERB scales.

The distance between pitches are perceived differently depending on their frequencies. A perceptualy small step in pitch is much larger at higher frequencies than at low frequencies. This can be approximated by the Mel Scale.

RANGE OF HEARING

Hearing threshold: human's hearing is limitated by the hearing threshold. Faintest sounds that are just audible without any other sounds being present. Measured by playing tones and someone detecting the presence of the tone. Threshold of hearing as a function of frequency. Tones beyond that threshold are audible.
Risk of damage threshold
Threshold of pain
Music: from 10Hz to 10kHz   Speech: from 80Hz to 7kHz.

MASKING

Masking is another effect closely related. When two signals are presented on the same CB, the one with more level will increase the perceptible threshold of the other one. Masking will have more effect in higher frequencies other than the lower ones, due to the asymmetry of the basilar membrane vibration response. This effect increases, and will have a proportional broader effect in frequency in low freq than in high freqs.

Masking is measured in $dB_{mask} = 10 * \log(I_{mask}/I)$, where I_{mask} is the hearing threshold with masking, and I the level without masking.

Masking effects can be classified according to different criteria. Regardid the sound sources location:

- Ipsilateral masking: when the sources are perceived by the same ear
- Contralateral masking: when sources come from differen ears.

If the masker and the masked signals are not presented in a synchronous way:
- forward masking: the masker is presented before in time (betwen 80ms and 300ms)
- backward masking: the master is presented after in time (up to 5ms)

Masking effects are more strong when there is no time gap between the master and the masked signals, If it's not the case, usually forward masking will produce more effect than backward masking.

Masking can alter also pitch sensations. When masking a tone with narrow-band noise, it can happen that the tone will be perceived as higher in frequency if the noise is lower, and vice-versa. This is due to the off-frequency effect, since the CB's are overlapping, it can happen that the tone is perceived also in another CB, thus biasing its pitch

We can see masking as a relative timing issue of the tone relative to the master. There is also a situation where the probe tone is present before a masking sound and there is even masking. You have to raise the level of the probe, about 20ms or can be 50ms. with hearing impairement these durations can be more.