MUSIC PERCEPTION AND COGNITION

ANDRÉS PÉREZ LÓPEZ

based on PERFECTO HERRERA slides

DECEMBER 2013

MUSIC TECHNOLOGY GROUP – UNIVERSITAT POMPEU FABRA

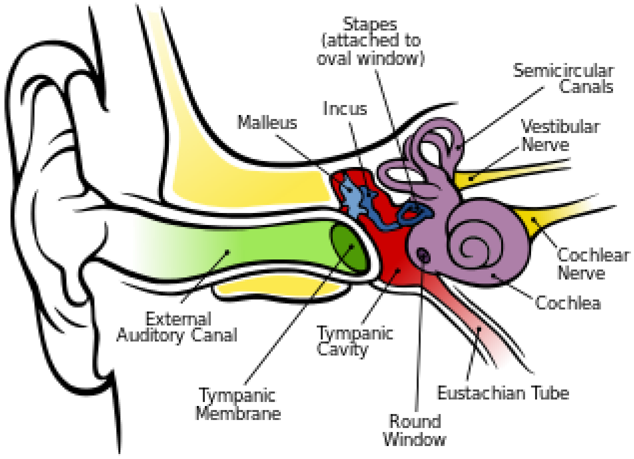
Contents:

1 Music Perception
   1.1 Physiology of music perception and cognition
   1.2 Psychophysics of the basic sound dimensions
      1.2.1 Frequency resolution
      1.2.2 Loudness
      1.2.3 Pitch
      1.2.4 Timbre
2 Music Cognition
   2.1 Perceptual organization
   2.2 Music and Memory
   2.3 Rhythm
   2.4 Music and Emotion

1 – MUSIC PERCEPTION

1.1 - Physiology of music perception and cognition

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



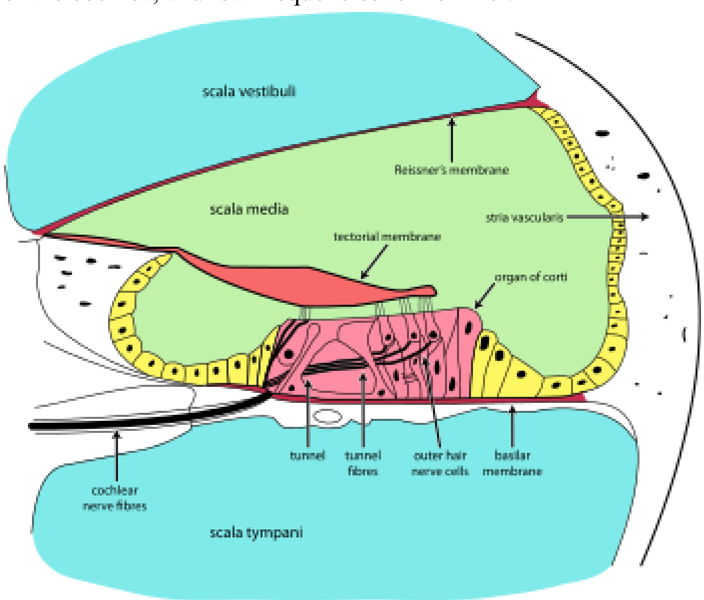
Anatomy of the Human Hear (Chittka L, Brockmann A)

The human auditory system is divided into three parts:

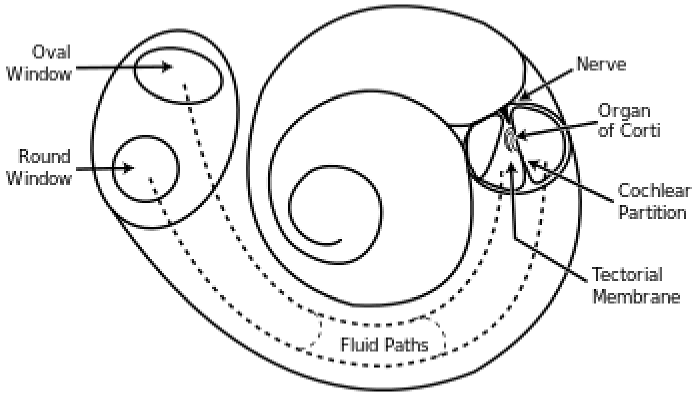
- Outer Ear (Pinnea and Auditory Canal): locating and filtering sounds
- Middle Ear (Eardrum/tympanic membrane, Malleus, Incus and Stapes): wave adequation and amplification
- Inner Ear (Cochlea): mechanical-electrical transduction

Mechanical waves propagate from Stapes to Oval window, thus causing a perturbation in the fluid inside the Cochlea, and moving the Basilar Membrane accordingly. The Organ of Corti, which is placed upon the Basilar Membrane, contains the Hair Cells, responsible of the transduction. Its movement causes the friction of the Hair Cells with the tectorial membrane, liberating electrochemical impulses.

Depending on its position, each Hair Cell will provide information about one specific frequency. High frequencies correspond to the outer part of the cochlea, and low frequencies to the inner.



Cochlea cross-section (Fred the Oyster)



Cochlea (Fred the Oyster)

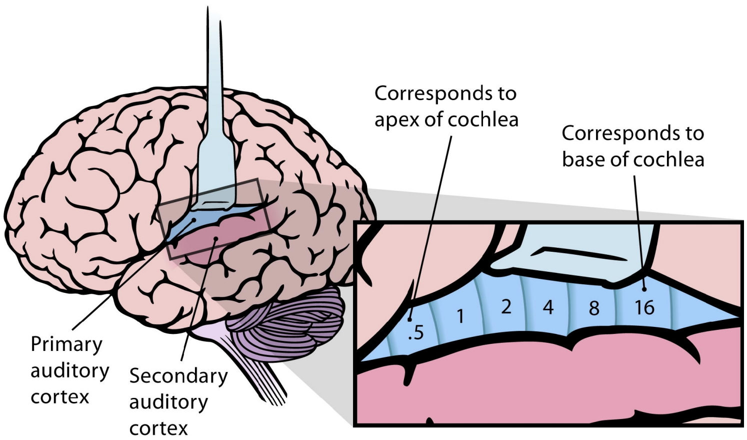
Auditory information travels then through the auditory pathway, until it reaches the Auditory Cortex. In this way, some processing is done, such as spatial localization and frequency identification. This processing is distributed across all neurons involved.

Auditory cortex is placed in both hemispheres, in the upper side of temporal lobes. It is divided into the Primary Auditory Cortex (A1), the Secondary Auditory Cortex (A2), and the Associative Auditory Cortex.

Different hemispheres are specialized in different auditory processing:

- Left hemisphere: temporal structure, speech, fast changes
- Right hemisphere: pitch, frequency, timbre, slow changes

As we can see in the image below, there is a *tonotopic organization*. That means that, as in the case of the Cochlea, information from different frequencies is placed separately also in A1. Due to *cortical plasticity*, it is possible to rearrange the organization after training.



Auditory cortex frequency mapping (Chittka L, Brockmann)

When listening to music, other cortical regions are activated:

- Motor area: active when the music is known
- Nucleus Accumbens, Amygdala: emotional responses
- Hypocampus: musical memory, experiences and contexts
- Prefrontal cortex: expectations

1.2 – Psychophysics of the basic sound dimensions

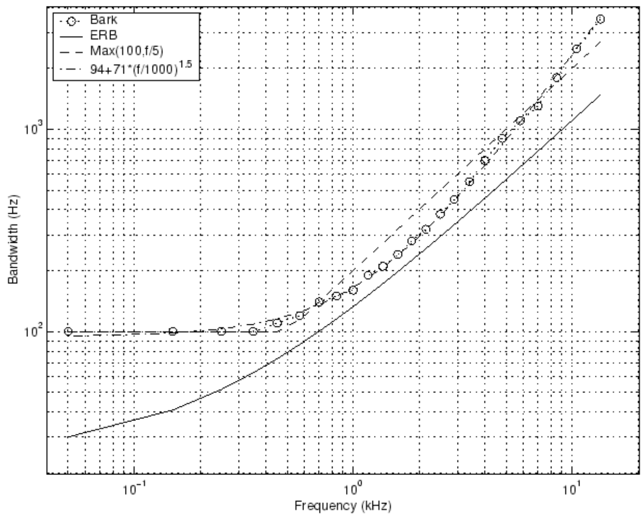
1.2.1 – Frequency resolution

In order to understand how the frequency resolution works, we must first have an insight into how frequencies are transduced from the Cochlea to the Auditory Cortex.

The Basilar Membrane acts as a bank of overlapping pass-band filters, the so called Auditory Filters (proposed by Helmholtz). Each location along its surface is responding to a limited amount of frequencies. These filters are non-linear, level-dependent, and its bandwidth and sharpness increases with frequency (similarly to the *constant Q*). The bandwidth of the Auditory Filters is called the Critical Bandwidth (CB, introduced by Fletcher).

Different measures have been proposed to emulate CB response:

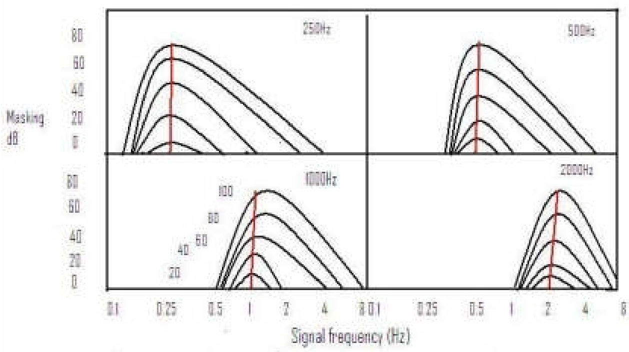
- Equivalent Rectangular Bands (ERB) are equivalent rectangular filters, which pass the same amount of energy that the equivalent auditory filters. They provide a useful computational simplification. ERB can be calculated by the formula $ERB = 24.7 * (0.00437 * f) + 1$
- Bark Scale is a frequency scale, in which each unit (bark) corresponds to one critical band. It have a shape with equal bandwidth up to 700 Hz, and 1/3 octave above. $Z_{bark} = 8.7 + 14.2 * 2\log(f/1000)$



Bark and ERB scales

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*.

Masking is another effect closely related. When two signals are presented in 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 than in lower ones, due to the asymmetry of the basilar membrane vibration response. This effect increases with level increases, and will have a proportional broader effect in frequency in low frequencies than in high frequencies.



Masking patterns

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

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

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

If the masker and the masked signals are not presented in a synchronous way, we have:

- Forward masking: the masker is presented before in time (between 30 ms and 300 ms)
- Backwards masking: the masker is presented after in time (up to 5 ms)

Masking effects are more strong when there is no time gap between masker and masked signals. If is not the case, usually forward masking will produce more effect than backwards 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 are overlapping, it can happen that the tone is perceived also in another CB, thus biasing its pitch.

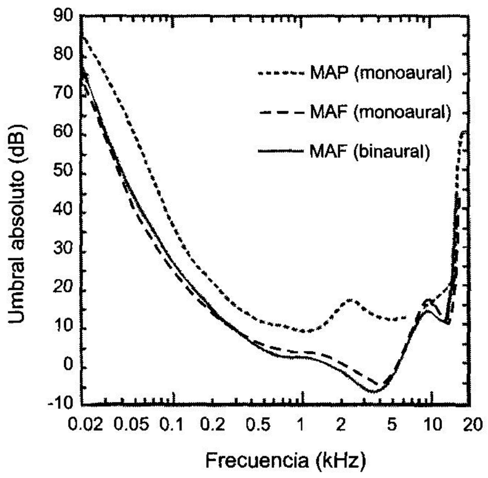
1.2.2 – Loudness

Loudness is the perceptual dimension related with the air pressure level.

Sound pressure level can be measured in different ways, but typically dB_{SPL} are used: $dB_{SPL}=20*\log(P/P_0)$,with P_0 usually taken as 20 µPa (the minimum audible pressure measured for a young human at 1 kHz).

There are different ways of measuring absolute pressure thresholds:

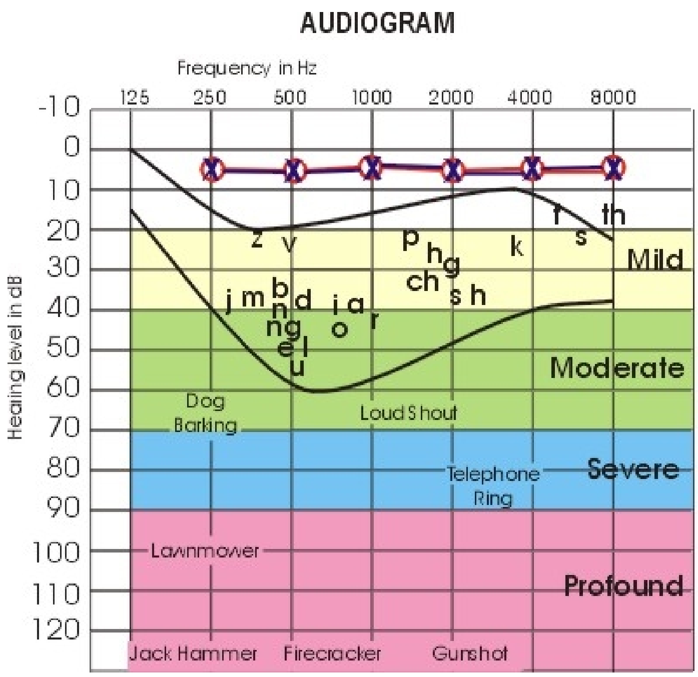
- MAP (Minimum Audible Pressure): measured in the ear canal
- MAF (Minimum Audible Field): measured in the free field, in front of the listener



Absolute sound pressure level thresholds

Another way to measure the pressure thresholds is by means of the Hearing Level, specially regarding hear impairment measures. Hearing Level measures provide information about how much the low threshold has raised, due to auditory system deterioration. Typically this information is displayed by means of an Audiogram.

Hearing Level is a magnitude relative to the population average – we consider 0 dB_{HL} as the average level threshold at 1 kHz, which is around 4 dB_{SPL}.

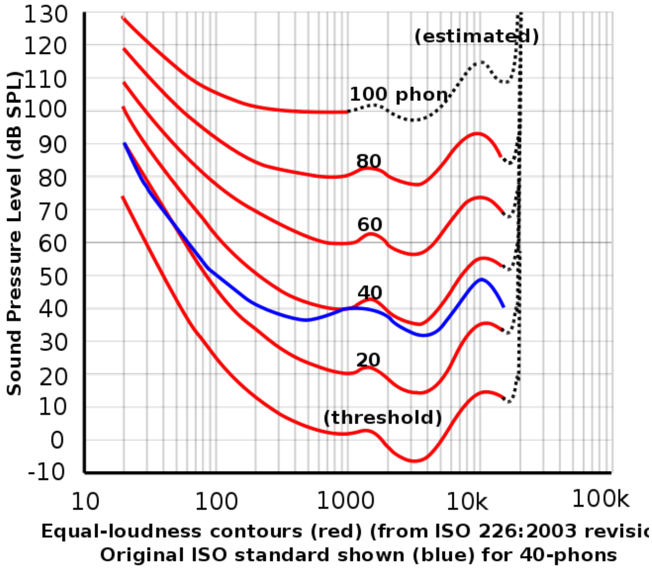


Hearing Level information in an Audiogram

According to Steven's Law ($R = kL^p$), it is possible to make a quantitative valoration of a perceptual parameter according to its related physical magnitude – for instance, to have an ordered scale of the loudness. Based on experimental results, it is accepted that an augment of 10 dB produces a doubling in loudness sensation (around 1 kHz).

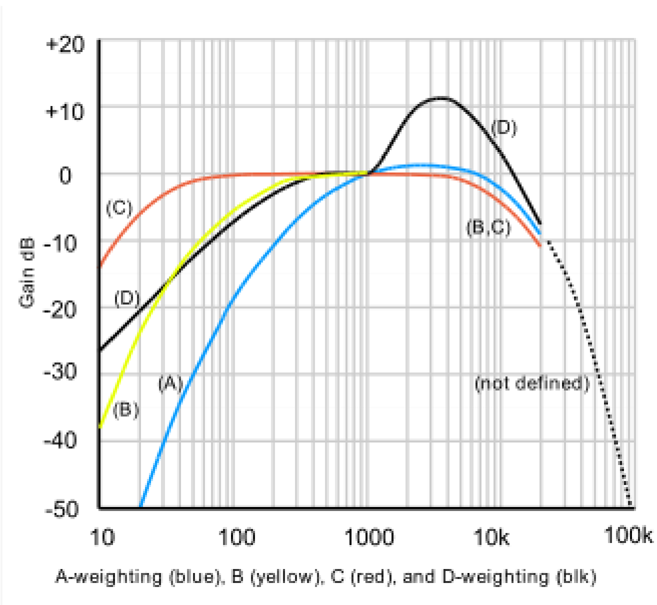
The *sone* is then defined as a unit for loudness measurement. 1 sone is defined as the loudness of a sinusoid at 40 dB_{SPL} and 1 kHz. So, according to the aforementioned, the same tone at 50 dB, which causes a doubling in loudness, will have a loudness of 2 sone, and so on.

The *phon* is another loudness unit, used for loudness comparison between sounds. By definition, the number of phon of a sound is the amount of dB_{SPL} of a tone at a frequency of 1 kHz that sounds just as loud. It is possible then to build graphs, called *Equal Loudness Contours* or *isophonic curves*, which provides information of loudness across frequencies.



ISO Equal Loudness Contours

There also exist some curves or filters which tend to take into account and compensate the loudness variation across frequencies. Most common weighting curve is the *A-weighting*, which emulates the isophonic contour at 40dB. Measures with this weighting use the dB(A) unit, which reflects that this concrete weighting was used.



Different standard weighting curves

Loudness is a perceptual feature which depends on the duration of the sound. In fact, the auditory system *averages* the input energy with a window of 200 ms. Consequently, a stimulus with length shorter than 150 will appear as with decreased loudness; its loudness will decrease 10 dB_{SPL} with a decreased duration of factor 10. It is shown that broadband noise is more affected by duration than pure tones.

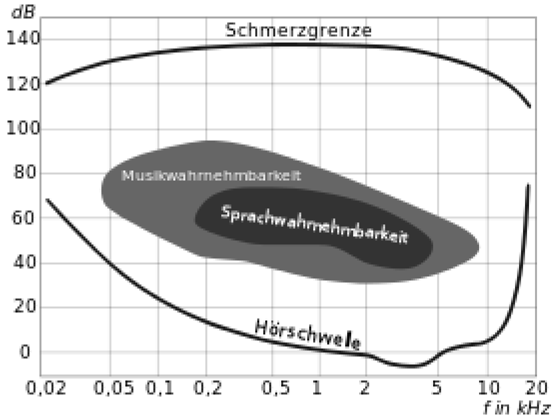
For simultaneous stimulus, loudness is also influenced by the stimulus' frequencies. As we saw before, CB play a crucial role: loudness sensation will only increase if the different partials are placed in different CB. If they are too close, they will *exceed* the CB *capacity*, so to say, and thus will not be perceived accordingly loud.

It is commonly accepted that loudness is encoded in the auditory system through changes in the BM neuron's firing rate: the louder the sound, the bigger the amount of spikes. However, this is not entirely true, since neurones with different spontaneous firing rates have different reactions to the sound pressure level; furthermore, they can easily be saturated after a certain amount of sound pressure, and anyway its dynamic range is much smaller than the perceivable human dynamic range for sounds. Other hypothesis, such as the spikes time pattern, and the number of neurons spiking, are also accepted.

1.2.3 – Pitch

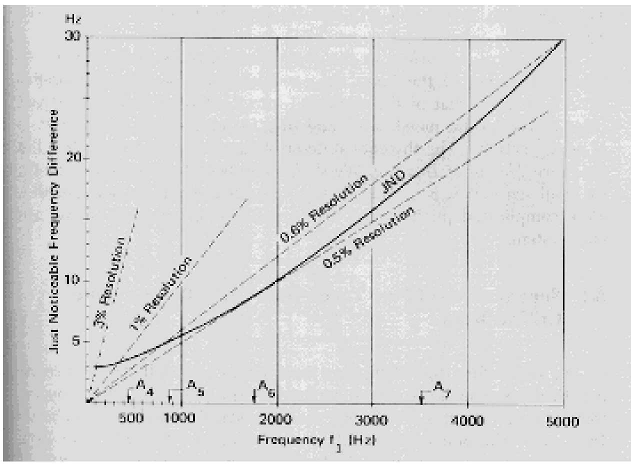
According to the ANSI definition, pitch is “that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high”. Pitch is the perceptual dimension related with wave frequency or periodicity.

In humans, the frequency hearing range (absolute threshold) is considered to extend from 20 Hz to 20 kHz. Pitch sensation is a subset of these frequencies, approximately going from 30 Hz to 5 kHz.



Audible range, music range and speech range

Regarding differential thresholds, the Just Noticeable Difference (JND) depends on the frequency. It is considered to be about 0.5% of the frequency, for frequencies above 1 kHz. For frequencies below, it is considered to be below 5 Hz.



Pitch Just Noticeable Difference

For a complex spectrum, with more than one partial, there are there are two different defined concepts:

- *Spectral pitch* is the pitch sensation that would produce one only partial
- *Periodicity pitch* is the pitch sensation that produce the whole harmonic spectrum sounding together

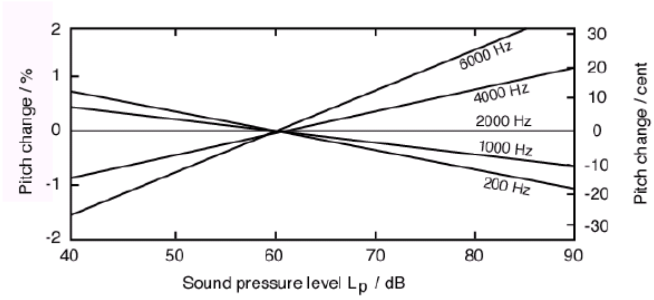
If the partials are ordered approximately in a integer multiples sequence, what is called an *harmonic spectrum*, we can perceive a *definite pitch*; this periodicity pitch will correspond to the spectral pitch of a pure tone placed in the lower partial (the *fundamental frequency*). If the spectrum is not harmonic, the sound will produce an ambiguous or *indefinite pitch* sensation.

There is a psychological effect happening in the Auditory Cortex called the *virtual pitch* or *missing fundamental*, that allows to hear a periodicity pitch in a harmonic spectrum, even if the fundamental frequency is not physically present. The resultant periodicity pitch depends on the difference frequency between partials, and on the highest common divisor. When these two coincide (as in the case of partials at 800, 1000 an 1200 Hz), the periodicity pitch is clearly placed at 200 Hz (both difference and hcd). When they do not coincide (in a non harmonic relation, such as 850, 1050 and 1250 Hz), the periodicity pitch is neither 200 Hz (frequency difference) nor 50 Hz (hcd), but a somehow ambiguous pitch above 210 Hz (the most matching fundamental possibility).

It is important to notice that, since virtual pitch is processed in the Auditory Cortex, it does not suffer for effects happening in the Cochlea, such as masking.

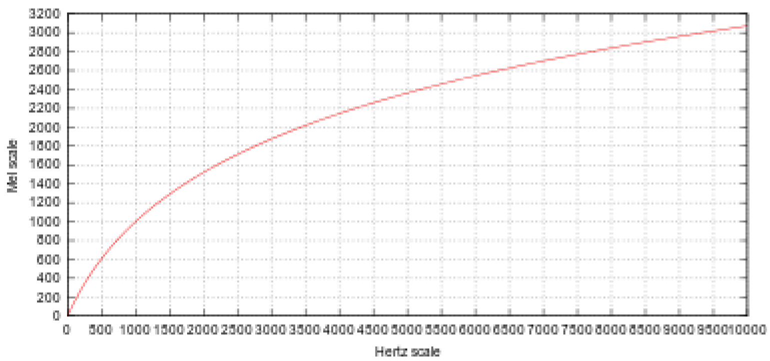
Pitch sensation depends also on other physical characteristics of sound:

- *Duration and pitch*: In order to perceive a definite pitch, sound duration must be bigger than a threshold. This threshold is dependent on the number of periods present in the sound. The higher the frequency, the bigger the number of periods needed to define the pitch; conversely, the higher the frequency, the smaller the duration of a period. Therefore, higher frequencies will need more periods, but less time, to define the pitch.
- *Intensity and pitch*
  - High frequency tones (above 2 kHz) tent to be perceived higher at levels above 60 dB_{SPL}, and lower below
  - Low frequency tones (below 2 kHz) tent to be perceived lower at levels above 60 dBSPL, and higher below



Effects of intensity on pitch perception

As in the case of loudness, we can create a scale with quantifies the perception of pitch. This scale is know as the *Mel scale*. By definition, a 1 kHz tone at 40 dB_{SPL} have a frequency of 1 kMel. A doubling in pitch height perception will correspond to 2 kMel, and so on. The conversion formula is $m = 2595\log_{10}(1+\frac{f}{700})$, with *f* in Hertzs.



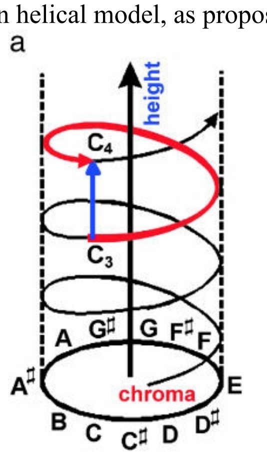
Mel scale versus frequency (Krishna Vedala)

Pitch sensation can be ordered in a two dimensional scale, according to two parameters:

- *Pitch height* is a continuous scale, arranging pitch from low to high as logarithmic function of frequency.
- *Pitch class* or *pitch chroma* divides the continuous pitch height into a set of discrete classes, ranged into one *octave* or frequency doubling. This octave- organization characteristic is found in every musical culture.

According to that, pitch can be represented in an helical model, as proposed by Sheppard

:



We tend to perceive pitch class in as a *categorical perception*, with well defined boundaries between instances.

There are two abilities related to pitch discrimination:

- *Relative pitch* is the ability of recognize the pitch relationship between two tones
- *Absolute pitch* is the ability of recognize the pitch of a single instance of a tone. According to Levitin, this ability not only present in musicians, and is divided into two different ones:
  - *Pitch memory*, as the ability of storing pitch information (developed by both musicians and non-musicians)
  - *Pitch labeling*, as the ability of naming this concrete pitch (typically developed by musicians)

When two tones are played together, they are referred as an *interval*. We can qualify them in terms of *consonance* and *dissonance*. There are two different types of consonance:

- *Tonal consonance* or *sensory consonance*: it is cultural-independent, and depends on the *roughness* of the produced sound. Helmholtz proposed that intervals that can be expressed as small integer ratios are considered more consonant.
- *Musical consonance:* it is cultural-dependent, and depends on the context (both musical and cultural), as well as in the tonal consonance

Dowling proposes four levels of abstraction for pitch perception, order by growing abstraction:

- *Psychophysical scale*: corresponds to the pitch height
- *Tonal material:* corresponds to the pitch class
- *Tuning system:* subset of the used tonal material. Typically a range of 5 to 7 pitch classes (which is related to the short-term memory limitations in handling simultaneous classes), selected in a more or less consonant way.
- *Mode:* implies the creation of a *tonal hierarchy*, which gives more importance or stability to some notes. The tonal hierarchies are perceived also for non-musicians.

A way of measuring tonal hierarchies is the *probe-tone method:* a sequence of notes is presented to a listener, and then he or she has to select the most fitting note that continues the sequence. This experiment will yield different results depending on the cultural context.

Tonal hierarchies are the base to create harmonic hierarchies and musical expectations. The mental representation of these hierarchies can be proved by the *chord priming* experiment. It shows that, in a sequence of two chords, we can determinate their consonance faster and more accurate than their dissonance. This feature is shown by both musicians and non-musicians, and shows the *implicit knowledge* about tonal relationships.

Regarding the neural coding of pitch, there are several theories that try to explain it:

- *Place theory:* pitch information comes from the maximum excitation place in the BM. This theory cannot explain effects such virtual pitch, since in that case there is no excitation present in the location associated with the periodicity pitch
- *Timing theory:* it is based in the *phase-locking* effect: neurons will spike always at a certain phase of the tone, but not necessarily at each cycle. It assumes that pitch is mainly inferred by unresolved high-frequency partials. However, not accordingly to this model, pitch sensation is more pronounced with low-frequency partials.
- *Pattern-matching theory:* it assumes that pitch is mainly inferred from low resolved partials. The brain computes the best fitting harmonic series that can be applied to the partials. This theory has the problem that does not explain pitch sensation from unresolved partials.

There is a consensus in considering pattern-matching theory as the main pitch theory, using the other presented theories as secondary or complementary mechanisms.

1.2.4 – Timbre

Timbre is the perceptual category of sound which allows to distinguish between sounds, when they coincide in all other perceptual variables. It is usually described as the *color* or *texture* of the sound.

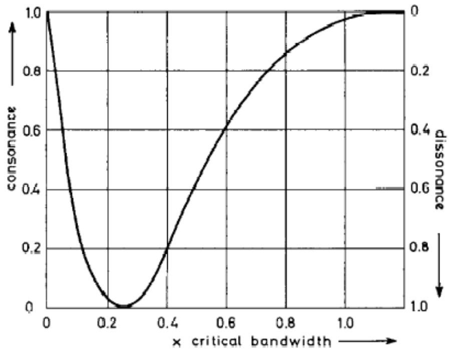
There are many physical variables from which timbre depends, from with the most predominant are:

- Spectral content and *spectral shape*: level relationship between partials
- Time envelope for each partial, with special emphasis on attack

Timbre is usually described by means of contrasting adjectives (bright/sharp, dull/mellow, fast/slow, soft/rough...). It is usually described in terms of spectral methods, such as *spectral centroid* or *spectral flux*.

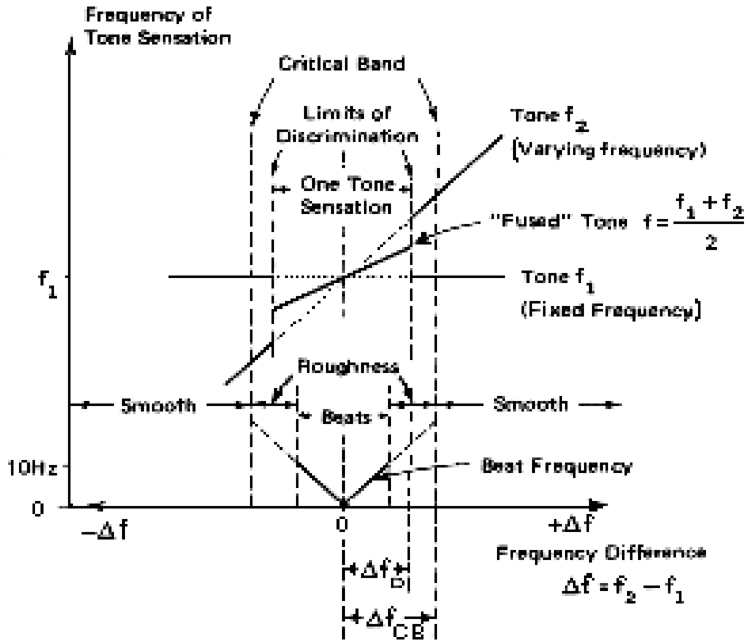
By combination of these methods, there are two terms which are typically used to describe sound:

- *Sharpness*: it is usually associated with brightness. In mathematical terms, it is related to the *spectral centroid*, or the "gravity center" of the spectrum. It is directly proportional to the frequency and intensity of the tone. It is measured in *acums*.
- *Roughness:* it is produced by partials sharing the same CB, and separated in frequency more than their 10%. It is the base for the consonance/dissonance classification between tones. The more roughness, the more dissonant the interval. Its unit is the *asper*.



Roughness as function of the frequency distance

Together with roughness, there is another psychophysical effect produced by partials sharing the same CB. When they are separated less than their 10%, they produce *beatings*, which are variations on the perceived intensity due to the produced interferences. It can be perceived with one single tone (when the frequency difference between partials is less than the JND), or with the two tones (in the other cases).



Beating and roughness in function of the distance between partials

2 – MUSIC COGNITION

2.1 – Perceptual organization

According to Pomerantz and Kubovy, “Perceptual organization is the process by which particular relationships among potentially separate elements are perceived and guide the interpretation of those elements… in sum, how we process sensory information in context”.

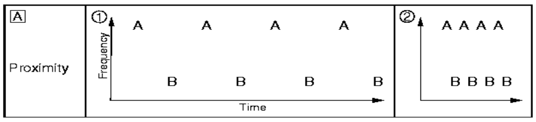
In order to understand perceptual organization in our brain, we use following terms:

- *Source* is the physical entity which produces the acoustic waves
- *Stream* is the perceptual grouping of several sources, in a way hat we perceive them as a whole. It is only possible to attend fully to one stream at the same time (*figure-ground effect*); furthermore, three or more simultaneous streams will form a *background*, not being possible to discriminate their individual characteristics (Murch's Law). Composition of streams can vary along time as it does the sound.

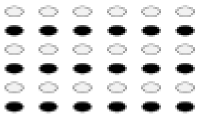
This ability of streaming the sources makes possible to differentiate a specific source in a “noisy” or environment. Examples of this are the *cocktail party effect*, and the ability of following a specific instrument within a band or orchestra. However, it is still not clear the way the auditory system performs this organization.

An early attempt to explain it was made by the *Gestalt* psychology. Gestalt argues that there are some innate principles of perceptual organization, which are, applied to sound:

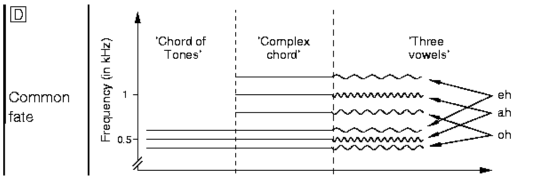
- *Proximity*: elements which appear close in time tend to be grouped together



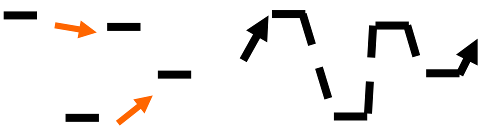
- *Similarity:* elements which have in common any perceptual category tend to be grouped together



- *Common fate:* elements with the same evolution in time tend to be grouped together

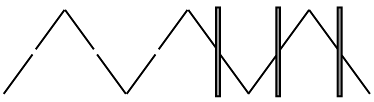


- *Good continuation:* elements with small variations over time tend to be grouped together



- *Belongingness*: we tend to separate elements in a figure-ground way; form will be the *smaller* element

- *Closure:* even when small gaps exist, we tend to perceive an event continuously

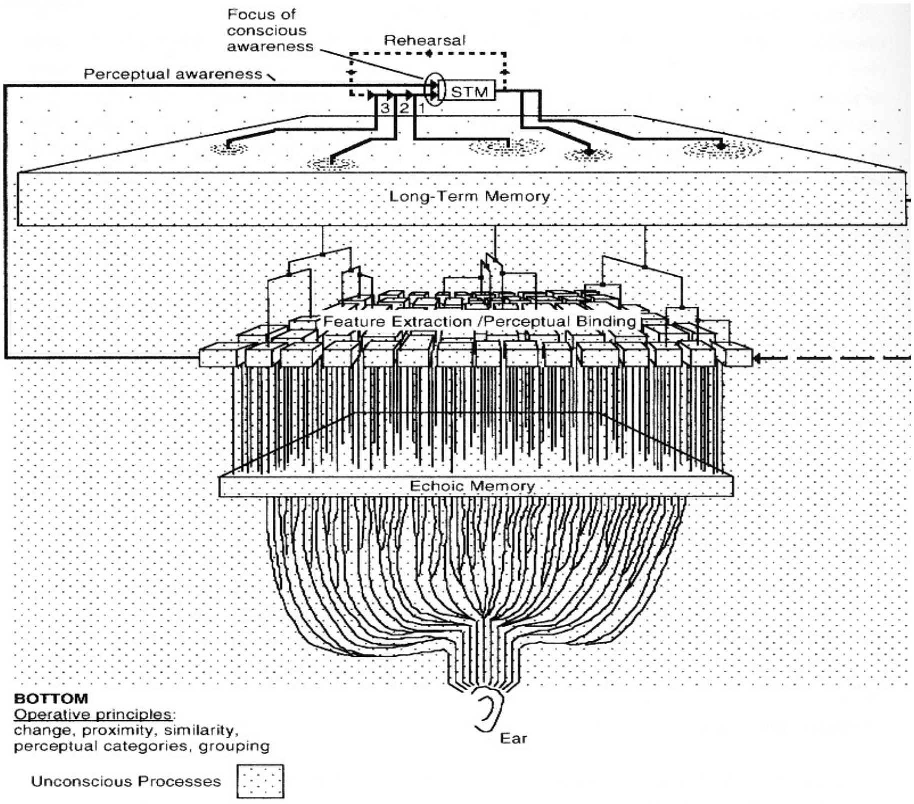


In Auditory Scene Analysis, Bregman reinterprets the Gestalt principles from the point of view of acoustic organization. According to him, there are mainly two grouping methods:

- *Simultaneous grouping*
  - *Simultaneous frequency*: two sources are easier to separate if they produce different fundamental frequencies
  - *Spectral regularity*: partials tend to be grouped together when presented in a harmonic relationship, and rejected if not
  - *Onset disparities*: sounds tend to be grouped if they have synchronous onsets. If the delay is less than 30 ms, we tend to separate this sound, even if consciously we perceive them as synchronous
  - *Correlated changes in frequency or amplitude*: sounds with same temporal evolution tend to be grouped (common fate Gestalt principle)
  - *Sound location*: sounds coming from different spatial positions tend to be separated
- *Sequential grouping:*
  - *Periodicity cues*
  - *Spectral cues*
  - *Level cues*
  - *Spatial cues*

2.2 – Music and memory

Human memory can be seen as a modular process in which three types of memory act: *echoic memory, short-term memory (STM) and long-term memory (LTM)*.



- *Echoic memory*

  Is in charge for the feature extraction and its binding. It is active up to only 4 seconds after the stimuli.

- *Short-term memory*

  It provides a temporal storage, facilitating the encode in the long-term memory. It is mainly the *conscious feeling of present*. It has a limited capacity (around 7 items), and also specific sensory blocks for parallel computing.

  In order to overcome the limited capacity, some "tricks" are used:

  - *Segmentation and grouping:* as seen in the perceptual organization, some cues are used in order to build a structure upon the plane stimulus.
  - *Chunking* is a method by which several elements are encoded together, and therefore act as a single item. This is the kind of processing done, for encoding scales and chords.

- *Long-term memory*

  It is the "memory" as commonly understood. It can be divided by:

  - *Declarative/explicit memory:* Is the content of the memory which is consciously available. Learning in this memory is a fast process, and the information is retrieved in "real time". There are two types:
    - *Episodic memory*: specific moments and situations (autobiographical memory)
    - *Semantic memory:* general information about the world ("common sense")
  - *Non-declarative/implicit/procedural memory*: Is the kind of memory which cannot be easily verbalized. Its learning process is slow, and acquired normally in a practical way. Once learned, it turns into a chunk in which is not possible to introspect.

2.3 – Rhythm

Rhythm is defined as the organization of events along time. Typically, although not required, those events are presented in a recurrent way.

We define an *event* as a whatever happening between *onsets.* We then define an *onset* as a detected change in one of the perceptual parameters (pitch, loudness, etc). The *Inter-Onset Interval (IOI)* is then defined as the time between successive events. IOI for rhythmic process is found approximately between 200 and 1800 ms. Below that, sounds are perceived as “fused”, and above the STM is not able to stream together the events.

Simple integer relations between IOI, and alternation between them, are preferred. In fact, due to categorical perception, it is difficult to perceive and produce non-isochronous IOI.

We call *pulse* to the succession of short-duration periodic *beats.* A *beat* is defined as the basic unit in time. Beats are perceptual features which serves as the repetitive temporal mental reference, and have no duration. *Beat induction* is the process by which a listener infers a pulse from a rhythmic pattern; it is an individual and cultural-dependent process, biased by several physical and perceptual parameters.

Tempo is derived from the pulse, and represents the repetition rate at which they occur. It is measured in *beats per minute (bpm)*, and ranges normally from 30 to 300 bpm (around the IOI range). 100 bpm is considered as “neutral” tempo, and provides the best timing accuracy. There are two related concepts:

- *Preferred tempo*: is the tempo preferred for the listeners
- *Spontaneous tempo*: the tempo at which a listener would tap spontaneously without musical input

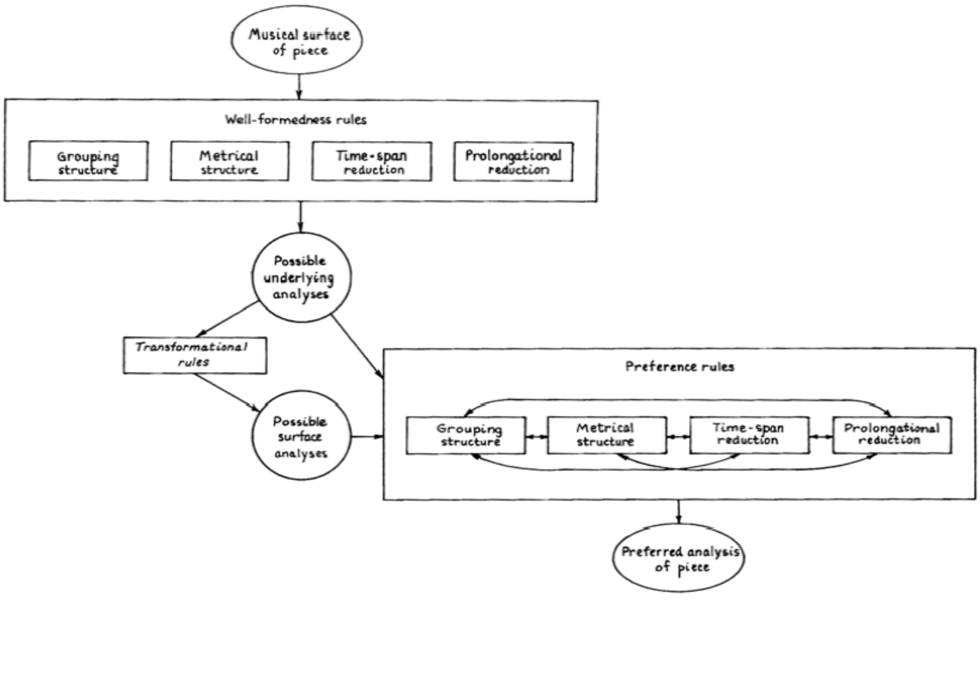
Both aforementioned tempos are located normally around 600 ms (again 100 bpm).

Tempo JND is situated around 5% for percussionist, and 10% for other instrumentalists and non-musicians.

When a series of events are presented, we tend to perceptually mark some of them in a regular basis; this marking is know as *metrical stress* or *accent*. Accents are favored when an event have some salient difference (pitch, timbre, loudness, etc). This process is know as *subjective rhythmization*, and tends to be produced in small integer ratios (normally 2:1 or 3:1); at fast tempos, the grouping is normally bigger.

Accents are the basis for the *meter* perception. Meter is characterized for the creation of a temporal hierarchy, the *metrical hierarchy*, in the form of strong/weak beats. When this hierarchy is established, the effect of an accent will be bigger when placed on a weaker beat.

As said before, Lerdahl and Jackendorff proposed in GTTM a theory that claims to imitate the human perception of rhythm. It works on both meter (retrieving beats) and grouping (organizing beats into phrases). They give a set of *well-formed rules* which creates potential candidates from raw information, and then a set of *preference rules* to actually select the preferred analysis.



2.4 – Music and emotion

Emotion is primitively a survival response of our brain. It consists on complex interactions by the neural and hormonal systems, and produce effects on affective experiences, cognitive and physiological processes, and behavior.

Emotion generates in the *limbic system*. The main systems involved are:

- *Nucleus accumbens*: rewards system
- *Amygdala:*survival values
- *Hypothalamus:*hormonal releases

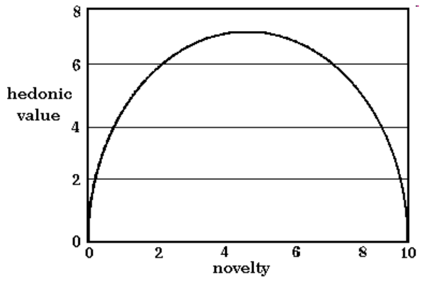
Emotions can be categorized by their neural path:

- *Primary*: generated inside the limbic system (thalamus → amygdala). It is the short way, and requires external inputs
- Secondary: generated across limbic system and cortex (thalamus → cortex → amygdala). It is the long way, and can be activated without external input.

Music can both *convey* and *induce* emotions:

- *Convey:* music can imitate they way emotions are encoded into speech (by means of timbre, loudness, etc) to transmit emotions. Furthermore, cultural situations with emotive component, in which music is used, can also be exploited to reproduce these emotions.
- *Induce*: music induces emotions by means of several paths:
  - *Musical expectation:*tension/stress, rewarding.
  - *Arousal*: cognitive activation can be produced by means of speed, complexity, etc.
  - *Mood contagion*: it is possible to exploit similarities between emotional language and music; this contagious can be also propagated by the audience.
  - *Associations:*music activates and recalls self-learned associations with emotional states
  - *Music imagery:* we can induce emotions without even an acoustic input

For some of these features, such as musical expectation and arousal, it is accepted that more joyful experiences happen at moderate levels, by following the U-Shape model.



Emotions have been classified according to many characteristics. One of the most used is the 2D representation, with axis representing valence (pleasant/unpleasant) and arousal (activation/deactivation).

Finally, we must take into account the social dimension of music, both in ancient times and nowadays. In fact, music preferences are one of the main ways of personal presentation and representation, specially between young people.