

Physiology of Music Perception and Cognition

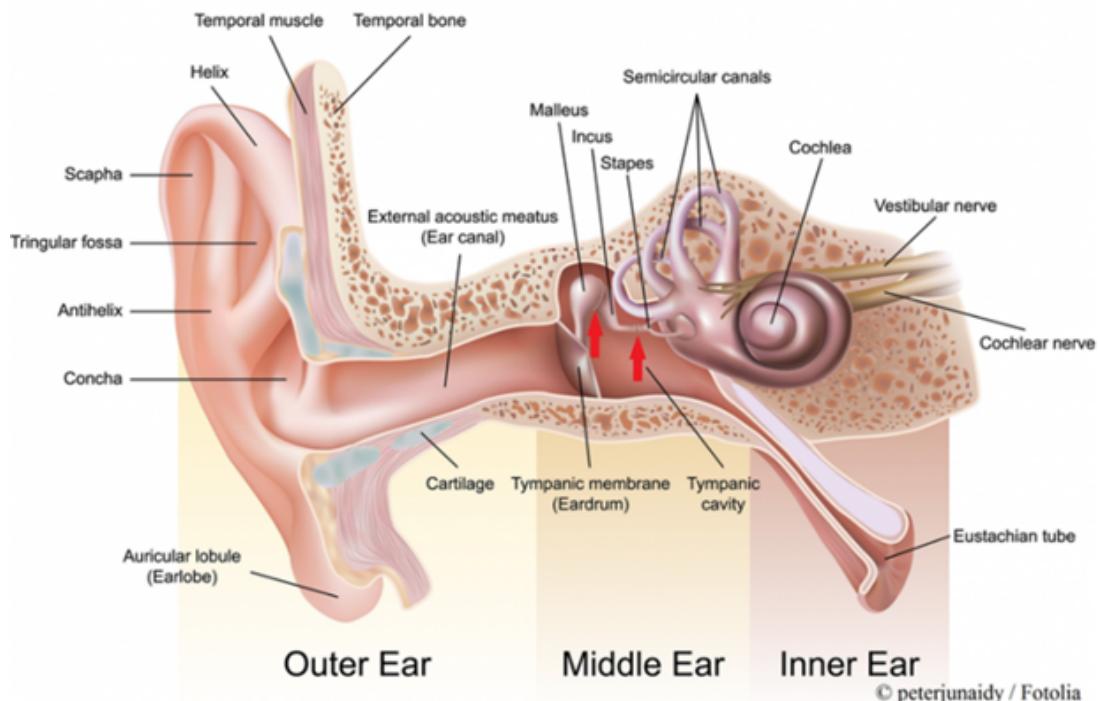
Enumerate the sections and elements of the whole hearing system

Outer ear: pinna and ear canal

Middle ear: eustachian tube, eardrum, tympanic cavity, malleus, incus and stapes

Inner ear: cochlea, semicircular canal

Anatomy of the Ear



Explain the functions of outer ear, middle ear and inner ear

Outer ear: sound collection, localization of sounds and protection

Middle ear: amplification, pressure compensation and protection

Inner ear: transduction (=transformation of sound into electric impulses) and equilibrium

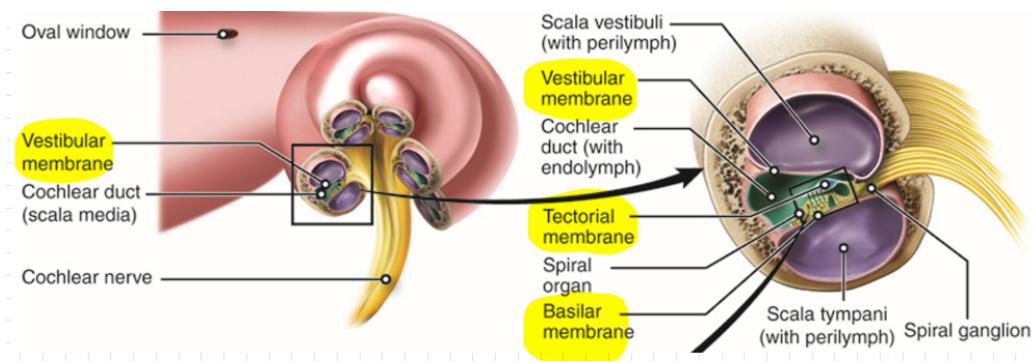
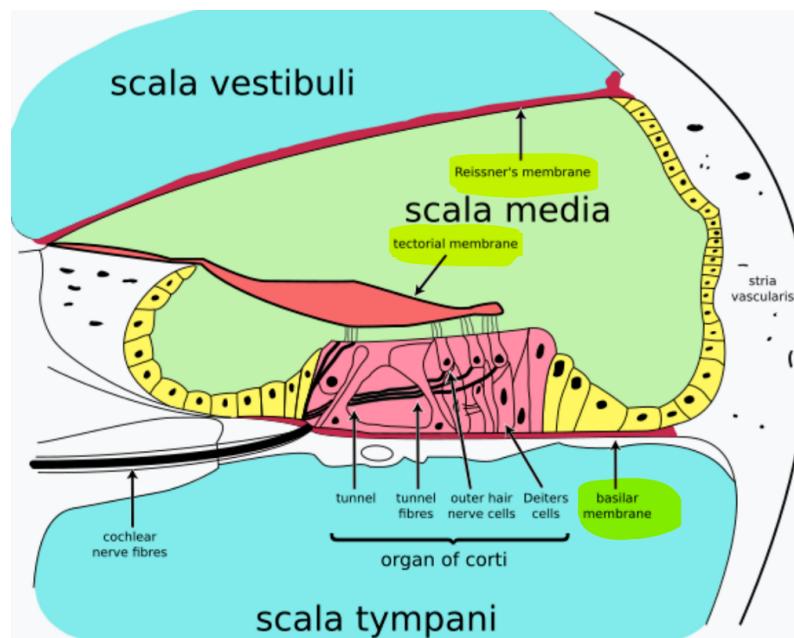
Enumerate and explain the functions of the different membranes involved in hearing

Tympanic membrane: Its main function is to transform the sound waves into mechanical vibrations that stimulate the inner ear

Basilar membrane: It provides information about the frequency of the sounds. It possesses graded mass and stiffness properties over its length, and its vibration patterns have the effect of separating incoming sound into its component frequencies that activate different cochlear regions.

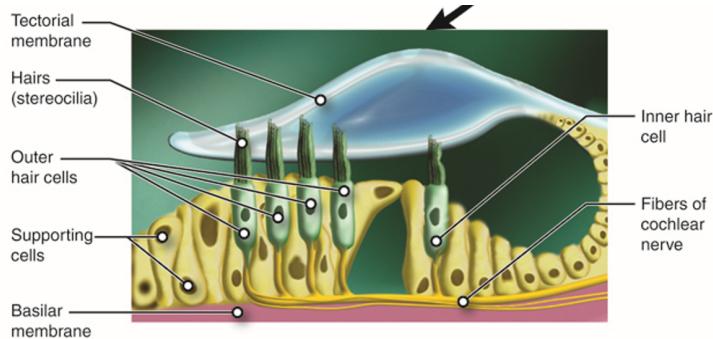
Vestibular membrane: The vestibular membrane, vestibular wall or Reissner's membrane, is a membrane inside the cochlea of the inner ear. It separates the cochlear duct from the vestibular duct. It helps to transmit vibrations from fluid in the vestibular duct to the cochlear duct. Together with the basilar membrane, it creates a compartment in the cochlea filled with endolymph, which is important for the function of the spiral organ of Corti.

Tectorial membrane: The function of this membrane is not clear yet. It overlies the sensory inner hair cells and electrically-motile outer hair cells of the organ of Corti and during acoustic stimulation stimulates the inner hair cells through fluid coupling, and the outer hair cells via direct connection to their tallest stereocilia.



Explain how the organ of Corti is structured, how it works and what results from its normal working

The main function of the organ of Corti is the transduction of auditory signals into electrical signals. Shifts between the tectorial and basilar membranes move the inner hairs and activate or deactivate receptors on the hair cell surface. This way the brain recognizes the frequencies of the sound.



How does the ear “decompose” or “analyze” acoustic information?

There are several strategies that the ear uses to code incoming information in each part of the ear:

The frequency analysis of environmental sounds begins in the **external ear**. The outer portion of the external ear reflects sound towards the ear canal. Once in the ear canal, the pressure waves are aligned so they strike the eardrum at right angles. The reflection of sounds of different frequencies is not the same and as a result the relative amplitude of some frequencies is greater than others. The result is that the relative amplitude of different frequencies at the eardrum differs, even if sound begins at the same intensity for all frequencies. Modification of the original sound by the external ear is a type of analysis that your brain learns to interpret. The frequency composition of familiar sounds aids your auditory system in determining where a sound is coming from.

The **middle ear** bones conduct sound from the eardrum to the fluids of the inner ear. The eardrum is bigger than the oval window. The decrease in these two membranes leads to an increase in pressure. The middle ear bones act as mechanical levers and further increase the pressure of the sound at the entrance to the cochlea. All of this is necessary to maximize the sound energy that gets to the fluids of the inner ear.

Finally, the **inner ear** determines how much energy is contained in the different frequencies that compose the sound. Auditory hair cells are specialized along the length of the cochlea to respond to the different sound frequencies the human ear receives.

These different locations then transmit the information to the brain. If the brain receives an increase in activity from a certain frequency location it then knows that the original sound contained energy at that frequency.

What happens in the cochlea when a square waveform is played? And if the waveform is a sawtooth? And if the sound is that of a snare?

Square and sawtooth waveforms are composed by a fundamental frequency and harmonics, therefore they will have resonances at those concrete positions in the cochlea. However, a snare is not harmonic and will resonate with a noisy distribution throughout the cochlea. (María Aguado says: I think this is the answer but i'm not sure lol).

(Additional info: a sine sound does not only excite the part of the basilar membrane that corresponds to its frequency but also other frequencies as well (**traveling waves**). The traveling wave moves from the **base** (high frequencies) to the **apex** (low frequencies), and declines after passing the resonance frequency. Therefore, we expect that a sound at a given frequency also affects the detection of a sound at another frequency.)

Explain what the characteristic frequency of a neuron in the auditory nerve is. Does it send spikes only when a sound is played? (this question is not answered)

The characteristic frequency of a neuron is the most sensitive frequency for an auditory nerve fiber. Each neuron is more sensitive over a narrow range of frequencies.

(Amaia) I think this has to do with the phase-lock. I wrote this in class but it's not suuper clear:

Neurons send spikes at the same phase as the waveform. But this spike can be sent frequently or not so much. The separation (in time) of the spikes will help the system to discover the periodicity (pitch). Meaning that periodicity of waveform = periodicity of spikes. Overall, neurons are phase-locked!

We will see that many spikes happen at multiples of T (see picture of slide 15), because of that we observe the frequency. For encoding low freq the total amount of spikes is smaller.

(Santi): chatGPT response about this. *The characteristic frequency of a neuron in the auditory nerve is the frequency of sound that the neuron is most sensitive to. This means that the neuron will respond most strongly to sounds at this particular frequency, and will send a larger number of spikes (or action potentials) in response to these sounds.*

However, the characteristic frequency of a neuron in the auditory nerve is not the only factor that determines how it will respond to a sound. Neurons in the auditory nerve are also sensitive to the intensity of the sound, and they will respond more strongly to louder sounds, regardless of the frequency. In addition, neurons in the auditory nerve will also respond to other factors, such as the duration and timbre of the sound.

Therefore, a neuron in the auditory nerve will not necessarily only send spikes when a sound is played. It will also respond to other stimuli, such as changes in the intensity or duration of the sound. However, the neuron will be most sensitive to sounds at its characteristic frequency, and will respond most strongly to these sounds.

What is phase locking, referred to the neuron's in the auditory nerve?

Phase locking refers to the phenomenon where the firing of neurons in the auditory nerve is synchronized with the cycles of a sound wave. This means that the neurons will produce spikes, or action potentials, at regular intervals that match the frequency of the sound wave. This allows the brain to accurately interpret the frequency of the sound and distinguish different pitches. Phase locking is important for the perception of sound, as it allows the brain to extract important information about the sound from the neural activity in the auditory nerve. It is also important for the ability to locate sounds in space, as the timing of the neural spikes can be used to determine the direction of the sound source.

Discuss if the function of the auditory pathways from the cochlea to the auditory cortex is just “transmission”

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.

Auditory messages are conveyed to the brain via two types of pathways: the primary auditory pathway which exclusively carries messages from the cochlea, and the non-primary pathway (also called the reticular sensory pathway) which carries all types of sensory messages.

(Amaia) I have something like this (my boyfriend helped me a little, he studies neuro-engineering):

It's not just transmission, some information is extracted along the way of the Auditory pathways before arriving to the Auditory Cortex.

- 1) Starting at the cochlea, the cochlea nuclei receives the information of one side of the body.
- 2) Then in the olfactory nucleus the first connection is done with the info of the other side of the body à combines information of sound pressure and delays for spatialization
- 3) Then the inferior colliculus nucleus, interprets and combines information connected to other facial muscles like eyes.
- 4) Finally they arrive to the auditory cortex

In the auditory cortex we have 3 regions:

- 1) Primary auditory cortex (AI), composed of neurons involved in decoding the cochleotopic and tonotopic spatial representation of a stimulus.
- 2) Secondary auditory cortex (AII), which doesn't have clear tonotopic organisation but has an important role in sound localisation and analysis of complex sounds: in particular for specific animal vocalisations and human language. It also has a role in auditory memory.

3) The tertiary or associative region, surrounding AI and AII, which helps to **integrate hearing with other sensory systems.**

(Chris Morse): After the cochlea, the first function of the auditory system is to determine sound source. The secondary function is to detect and classify the sound system. The tertiary function is to learn from the sound and its meaning. Along the auditory pathway are structures to determine specific qualities of the sound which get processed in parallel BEFORE reaching the auditory cortex. The whole system computes the essential functions of sound identification, NOT just the upper parts of the brain. For example, loudness and pitch information are available before reaching the auditory cortex.

What's the purpose of crossing information from the left and right ears? Where does it happen? Why so "early" in the auditory pathway?

The purpose of combining information from the left and right ears is to help us determine the direction of a sound source. This process, known as binaural hearing, occurs in the brainstem, specifically in an area called the superior olivary complex. This area is "early" in the auditory pathway because it is one of the first places that auditory information is processed after it is received by the ears. Binaural hearing is important for helping us to locate the source of a sound in our environment, which is a crucial ability for survival.

Explain what “tonotopic organization” (or tonotopical mapping) is and how far up in the hearing system can it be observed. Has this type of organization had anything to do with the sensitive and motor “homunculus” representation?

In the Cochlea, information from different frequencies is placed separately. The brain receives, up to its primary region A1, the info in a tonotropic way. Tonotopy is the spatial arrangement of where sounds of different frequency are processed in the brain. Tones close to each other in terms of frequency are represented in topologically neighbouring regions in the brain (Low freq go together, high together, they reach the auditory cortex in an organized way)

Tonotopy in the auditory system begins at the cochlea and ends at A1.

Santi (chatGPT): Tonotopic organization, also known as tonotopical mapping, is a characteristic feature of the auditory system that refers to the way that different frequencies of sound are represented in different regions of the auditory pathway. In the cochlea, for example, low frequencies are represented at the base of the cochlea, while high frequencies are represented at the apex. This pattern of organization is preserved as the sound signals are transmitted through the auditory nerve and into the auditory cortex, where different regions are specialized for processing different frequencies of sound.

In which ways the nervous system “codes” the acoustic information?

Acoustic information is coded in electrical signals thanks to neurons that create electrical impulses that travel through the nervous system to the brain.

Mention to phase-locking.

(Additional info: There are efferent and afferent neurons: depending on the direction in which information travels across the nervous system. Afferent neurons carry information from sensory receptors of the skin and other organs to the central nervous system (i.e., brain and spinal cord), whereas efferent neurons carry motor information away from the central nervous system to the muscles and glands of the body.)

Explain the basic/rough “modular” organization of the brain

Modularity of the brain: evidence of brain modularity is the empirical body of findings that demonstrates that the brain is organized into semiautonomous groups of neurons and brain regions that interact with one another through relatively sparse and long-distance pathways in the brain. Modular models of the brain aid the understanding of a complex system by decomposing it into structural modules.

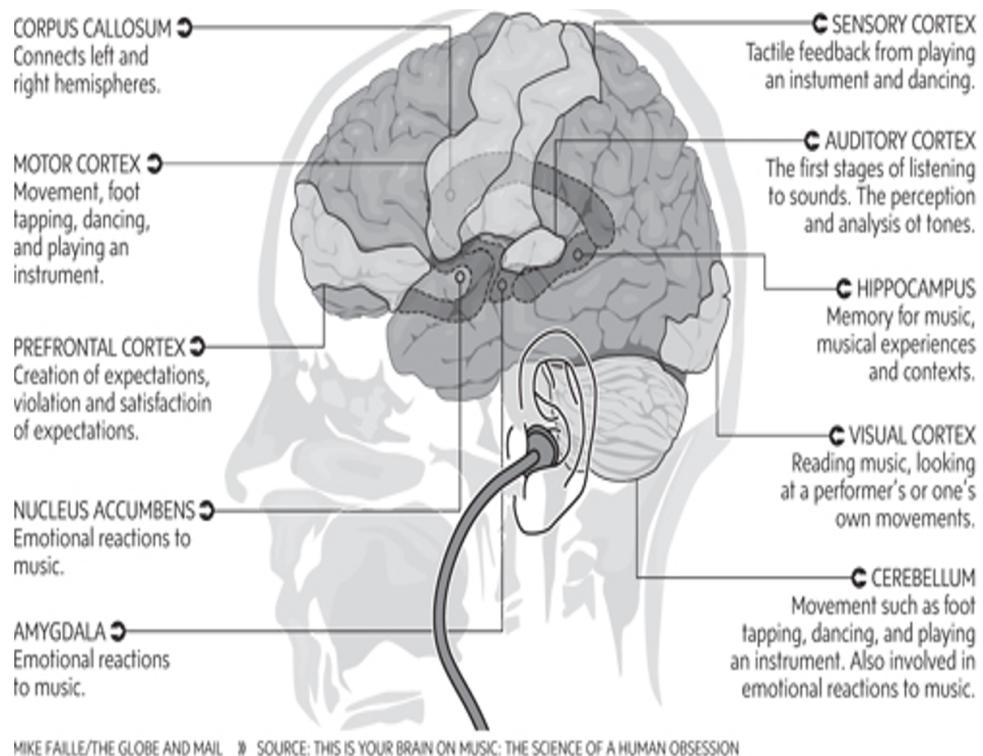
For example:

Temporal lobe we locate the area where the auditory nerve is

Occipital lobe: visual info

Parietal lobe: touch info

Name at least 6 areas in the brain involved in music processing and understanding, and their respective roles in it



MIKE FAILLE/THE GLOBE AND MAIL » SOURCE: THIS IS YOUR BRAIN ON MUSIC: THE SCIENCE OF A HUMAN OBSESSION

1. The auditory cortex is involved in the initial processing and interpretation of auditory information, including music.
2. The insula is involved in emotional responses to music.
3. The basal ganglia are involved in the motor control and movement associated with music, such as playing an instrument or singing.
4. The hippocampus is involved in the formation of memories related to music.
5. The amygdala is involved in the emotional processing of music.
6. The prefrontal cortex is involved in higher-level cognitive processes related to music, such as analyzing and making judgments about musical structures and patterns.

Speculate/discuss where our musical memories could be in the brain

It is difficult to speculate on the specific location of musical memories in the brain, as the exact mechanism by which the brain processes and stores musical information is not fully understood. However, it is likely that musical memories are stored in a network of brain regions involved in sensory perception, emotion, and memory. These regions may include the auditory cortex, which is responsible for processing sound, as well as the hippocampus, which plays a key role in forming and storing memories.

Speculate/discuss about how veridical/faithful is the representation that our brain builds about the acoustic/musical environment?

It is difficult to say how veridical or faithful the representation of the acoustic environment that the brain builds is, as this can vary greatly from person to person and can depend on a number of factors, such as a person's individual sensory abilities, their past experiences, and the specific context in which they are listening to music. Additionally, the brain is not a perfect recorder of sensory information, and can alter or distort incoming sensory information in various ways. For example, the brain may fill in gaps in the sensory information it receives, or it may prioritise certain aspects of the information over others. Overall, while the brain is capable of building a representation of the acoustic environment, the accuracy and fidelity of this representation can vary.

Speculate/discuss the pros and cons of biologically plausible emulations of the auditory processing that happens in the ear and early stages of the auditory pathways.

One potential benefit of biologically plausible emulations of auditory processing is that they could potentially provide a more natural and intuitive way for machines to interpret and understand sounds. This could be especially useful in applications where it is important for the machine to be able to recognize and understand complex or subtle sounds, such as in speech recognition or environmental monitoring.

On the other hand, there are also potential drawbacks to using biologically plausible emulations of auditory processing. One potential issue is that such emulations may be computationally intensive, which could make them less practical for use in real-time applications. Additionally, there is also the risk that such emulations may not be able to accurately capture all of the nuances and complexities of auditory processing, which could limit their effectiveness in certain situations.

Overall, the use of biologically plausible emulations of auditory processing is likely to be most effective when it is used in combination with other approaches, such as more traditional signal processing techniques, to provide a more comprehensive and accurate representation of sound.

Can you give examples of techniques or algorithms that we use in music processing that are inspired or try to emulate something of the early auditory processing?

There are several techniques and algorithms used in music processing that are inspired by or try to emulate aspects of early auditory processing in humans. One example is the use of filters to mimic the way the human ear processes sound. Filters can be used to remove unwanted noise or to emphasize certain frequencies, much like the ear naturally does. Another example is the use of feature extraction algorithms, which try to identify relevant characteristics of a piece of music, such as its rhythm or pitch. These algorithms are often based on how the brain processes

and recognizes music. Additionally, many music generation algorithms use models of how the brain processes music to generate new melodies or harmonies.

The fourier series analysis done by identifying the frequencies in an audio is similar to what is done in the cochlea (frequencies recognized from base to apex).

Research Methods

Explain the differences between observational studies, correlational studies and experimental studies

Observational: Researchers observe participants but do not attempt to influence them.

Correlational: Researchers examine how two or more variables are related. It is not possible to tell which variable is affecting the other(s). As you have probably heard, "correlation is not causation."

Experimental: Researchers control conditions to examine how one variable affects the other(s). Participants are assigned to random groups (at least two). There is a control group that does not experience or receive the variable being examined and a experimental group that does experience or receive the variable being examined. The groups are compared to examine the effect of the variable being investigated.

Explain the different variable types involved in an experiment

Independent variables: these are the variables that you can control on an experiment. (x)

Dependent variables: the type of variables you cannot control because they are fully dependent on the independent variables. They respond to changes of the independent variables. (y(x))

Control variables: they are variables you keep the same or do not change. They must be well known and fixed in order to establish trustworthy dependencies between the previous types of variables.

What is a hypothesis? Can it be confirmed or refuted using only experiments?

A hypothesis is an assumption, an idea that is proposed for the sake of argument so that it can be tested to see if it might be true.

What is a pilot study? Why should you include it for your thesis work?

A pilot study is a small-scale preliminary study conducted before any large-scale quantitative research to evaluate the potential for a future, full-scale project.

Describe different approaches to study music perception and cognition phenomena

Surveys (not an experimental procedure)

Physiological measures: Heart rate, skin conductance, blood pressure, temperature, muscle tension.

Behavioral measures: Errors solving a task, reaction time, relatedness or similarity (5- or 7-point Likert-type scale), choice (forced, not forced, 2/3/4 choices), eye-fixation, motion capture, Quantified Self.

Electro-Encephalography (EEG)

Event-Related Potentials (ERP): ERP: neural signal that reflects **coordinated activity of an ensemble of neurons**, observed “after” certain events or stimuli have been processed.

Frequency-Following Response (FFR): 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 population of neural elements. It is often phase-locked to the individual cycles of the stimulus waveform and/or the envelope of the periodic stimuli.

Magneto-Electro-Encephalography (MEG):

Positron Emission Tomography (PET): Radioactive contrast substance injected.

Measurements of local blood flow, metabolism, neuro-receptor bindings...

Functional Magnetic Resonance Imaging (fMRI): fMRI measures brain activity by detecting changes associated with blood flow. This technique relies on the fact that cerebral blood flow and neuronal activation are coupled. When an area of the brain is in use, blood flow to that region also increases.

Describe different tools to study music perception and cognition phenomena

Music perception and cognition can be studied through a variety of tools and methods. Some common tools used to study these phenomena include surveys and questionnaires, which can be used to gather self-reported data from participants about their experiences with and perceptions of music; experimental studies, in which participants are exposed to specific musical stimuli and their responses are measured; and neuroimaging techniques, such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), which allow researchers to study the brain activity of participants while they listen to or perform music. Other tools that are commonly used in the study of music perception and cognition include behavioral measures, such as reaction time tasks and ratings of emotional responses to music, and computational models, which can be used to simulate and analyze musical data.

What is an ERP? Is it the only type of electrophysiological measurement that can be taken from the brain?

An ERP, or event-related potential, is a type of electrophysiological measurement that is used to assess the electrical activity of the brain in response to specific stimuli. This type of measurement is typically obtained by using electroencephalography (EEG), which involves placing electrodes on the scalp to measure the electrical activity of the brain.

ERPs are not the only type of electrophysiological measurement that can be taken from the brain. Other examples include magnetoencephalography (MEG), which measures the magnetic fields produced by the brain, and functional magnetic resonance imaging (fMRI), which

measures changes in blood flow to different areas of the brain. Each of these methods has its own advantages and disadvantages, and they can be used alone or in combination to study different aspects of brain function.

What is fMRI? Discuss advantages and disadvantages in comparison to electrophysiological measurements

Functional magnetic resonance imaging (fMRI) is a neuroimaging technique that is used to measure brain activity. It works by detecting changes in blood flow to specific areas of the brain. This allows researchers to infer which brain regions are being used during a particular task or experience.

One advantage of fMRI is that it is a non-invasive technique, meaning that it does not involve inserting any electrodes or other devices into the brain. This makes it a safe and relatively comfortable way to study the brain. fMRI is also relatively good at providing high-resolution images of brain activity, which can help researchers to identify specific brain regions and their functions.

However, there are also some disadvantages to using fMRI. One of the main disadvantages is that it has low temporal resolution, meaning that it is not very good at measuring rapid changes in brain activity. This is in contrast to electrophysiological measurements, which can provide a much more detailed picture of the timing of brain activity. Additionally, fMRI is limited in its ability to measure activity deep within the brain, so it may not be as useful for studying certain brain regions or processes.

Imagine that you want to study the emotional responses to music. Give several examples of approaching the problem using different research strategies (do not forget to include one that includes behavioral measures).

One way to study the emotional responses to music would be to use a survey or questionnaire to ask participants about their emotional experiences while listening to music.

This could be a self-report measure, where participants are asked to rate their emotional experiences on a scale or choose from a list of emotions. Another way to study the emotional responses to music would be to use a more objective measure, such as facial expression analysis or physiological measurements (such as heart rate or skin conductance).

For example, participants could listen to different pieces of music while their facial expressions are recorded, and then a trained observer or computer program could analyze the facial expressions to determine the emotional responses to the music. Another approach could be to use a behavioral measure, such as having participants perform a task (such as a cognitive task or motor task) while listening to music, and then measuring the effect of the music on performance. Participants could listen to different types of music and then complete a cognitive task, such as a memory task or a problem-solving task, and researchers could compare the

performance on the task across different types of music to see if certain types of music have a positive or negative effect on performance.

Search for some papers on one specific topic in music perception and cognition: one of the papers should be using behavioral techniques, one using physiological techniques, one using fMRI, one using ERPs

Frequency resolution and masking

Explain the concept “frequency resolution”.

It is the ability to hear frequencies separately. Frequency resolution ability can be interpreted as the filtering capability of the auditory system. It can be studied by the shape of the underlying filter – called the auditory filter – that is centered at the frequency component of interest.

Describe the limitations humans have on such “frequency resolution” and its bases.

The ability of humans to perceive differences in frequency, or pitch, is limited by a number of factors. One of the primary factors is the range of frequencies that the human ear can detect, which is typically between 20 and 20,000 Hz for most people. Another factor is the sensitivity of the ear, which can vary from person to person and can be affected by factors such as age and hearing loss. Additionally, the brain plays a role in the perception of frequency, and certain cognitive processes can affect how well an individual can discern differences in pitch. For example, if two sounds are presented close together in time, it can be difficult for the brain to distinguish between them.

Define masking.

It's the effect by which a faint but audible sound becomes inaudible in the presence of another sound.

Describe the different “asymmetries” to be found when observing masking.

When two sounds are presented simultaneously, the louder or more intense sound can mask the quieter one, making it more difficult or impossible to hear. This effect can be observed in a number of different ways, and there are a number of asymmetries that can be found when studying masking. One common asymmetry in masking is known as the forward masking effect, in which the masker (the louder sound) precedes the maskee (the quieter sound) in time. This can make it more difficult to detect the maskee, especially if it is presented close in time to the masker. Another asymmetry is known as backward masking, in which the masker follows the maskee in time. This can also make it more difficult to detect the maskee, especially if the two sounds are

presented in close proximity. Another asymmetry in masking is known as simultaneous masking, in which the masker and maskee are presented at the same time. In this case, the masker can make it more difficult to detect the maskee, even if the two sounds are presented at different frequencies. Finally, there is the phenomenon of temporal masking, in which the masker and maskee are presented in close proximity in time, but not simultaneously. This can also make it difficult to detect the maskee, especially if the two sounds are presented at similar frequencies.

Describe different types of masking. Relate some of them with physiological constraints in our hearing and neural mechanisms.

Simultaneous masking by tones

Upward spread of masking: masking increases strongly at high frequencies if masker level >50 dB.

Masking strongest near masker frequency

Temporal masking

Raise probe level by 10 dB for duration < 200 ms

Backward masking: masking before sound onset

Forward masking: masking decays over 200 ms after the end of masker.

AI said:

There are several different types of masking in the field of auditory perception, which refers to the ways in which one sound can interfere with the perception of another. One type of masking is called "**forward masking**," which occurs when the presence of one sound makes it harder to detect a subsequent sound. This can occur because the first sound activates the auditory neurons in the ear, temporarily reducing their sensitivity to subsequent sounds. This type of masking is thought to be a result of physiological constraints in the ear and the neural mechanisms involved in auditory perception.

Another type of masking is called "**backward masking**," which occurs when the presence of one sound makes it harder to detect a preceding sound. This is thought to be due to the way that the brain processes and integrates auditory information over time, with later sounds effectively "masking" the perception of earlier ones.

In addition to these temporal forms of masking, there is also "spectral masking," which occurs when the presence of one sound makes it harder to detect another sound that occurs at the same time, but at a different frequency. This is thought to be due to the way that the auditory system processes sounds within different frequency bands, with sounds in one band masking the perception of sounds in other bands.

Overall, these different types of masking reflect the complex ways in which the auditory system processes and integrates sounds, and the physiological and neural constraints that shape our ability to perceive the sounds around us.

Under which conditions we expect maximal masking when listening to a complex musical stimulus (i.e., think about that as a performer, DJ/producer and/or listener).

The maximal masking happens when the masker and the signal are the same frequency and this decreases as the signal frequency moves further away from the masker frequency.

Comment on the idea that “the peripheral auditory system can be conceived as a bank of bandpass filters with overlapping bands”.

Auditory filters do work as bandpass filters and they play a big role in the frequency resolution ability. Bandpass filters emulate accurately the functioning of our auditory.

AI said:

The idea that the peripheral auditory system can be conceived as a bank of bandpass filters with overlapping bands is a widely accepted model of how the ear processes sounds. A bandpass filter is a type of electronic filter that allows only a certain range of frequencies to pass through, while attenuating or blocking other frequencies. In the case of the auditory system, these filters are thought to correspond to the different frequency bands or "octaves" that the ear can perceive, with each filter tuned to a specific range of frequencies.

One of the key features of this model is that the filters have overlapping bands, meaning that the same frequencies can be detected by more than one filter. This is thought to be important for our ability to detect subtle differences in the frequency of sounds, as well as for our ability to combine information from different filters to create a more complete representation of the sounds we hear.

Overall, this model provides a useful framework for understanding how the ear processes sounds, and how different frequencies are represented in the auditory system. It also highlights the importance of the overlap between different frequency bands in our ability to perceive sounds accurately.

Explain the concept of critical bandwidth and connect it with that of “frequency resolution”. Explain how the critical bandwidth changes depending on frequency.

The width of the auditory filters is what we call 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 it would be only one. This effect is called unresolved frequencies, and is the base of different effects, such as beating and roughness.

AI said:

The concept of critical bandwidth is a fundamental concept in the field of auditory perception, and is related to the idea of "frequency resolution" in the auditory system. Critical bandwidth refers to the range of frequencies that are just barely distinguishable as separate sounds by the ear. In other words, it is the smallest change in frequency that can be detected by the ear, and is therefore a measure of the ear's ability to resolve or distinguish between different frequencies.

The critical bandwidth is known to vary depending on the frequency of the sounds being considered. At low frequencies, the critical bandwidth is relatively large, meaning that the ear is less sensitive to small changes in frequency. At higher frequencies, the critical bandwidth is smaller, meaning that the ear is more sensitive to small changes in frequency. This is thought to reflect the way that the auditory system processes sounds within different frequency bands or "octaves," with higher frequency bands being more finely tuned than lower frequency bands.

Overall, the concept of critical bandwidth is important for understanding how the auditory system processes and distinguishes between different frequencies, and how this ability varies depending on the frequency of the sounds being considered.

JND Pitch perception graph shows at low frequencies change needed is less and at higher frequencies change in frequency needed is more. (not to be confused with critical band where at low frequencies critical band is large and high frequencies it's critical band is small)

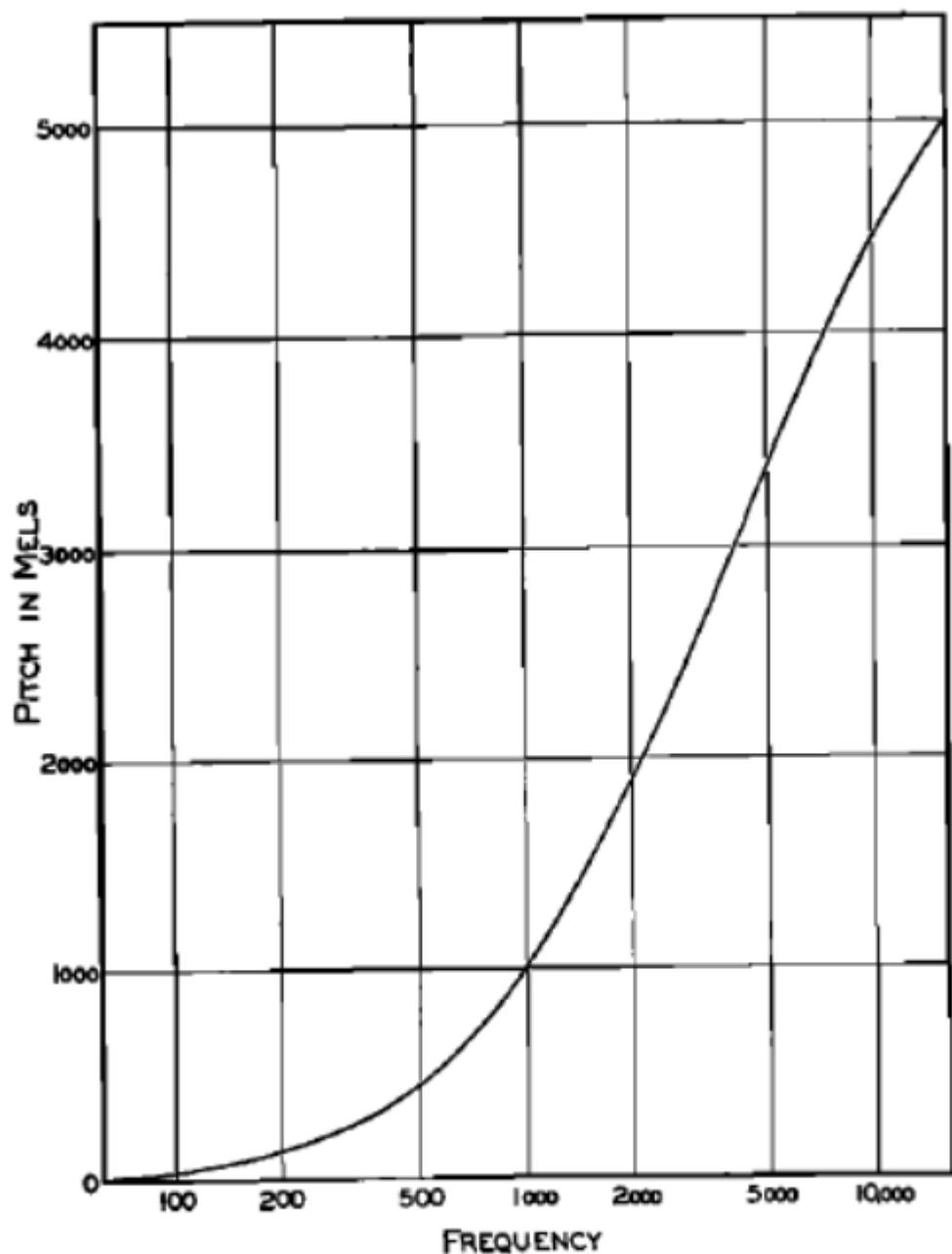


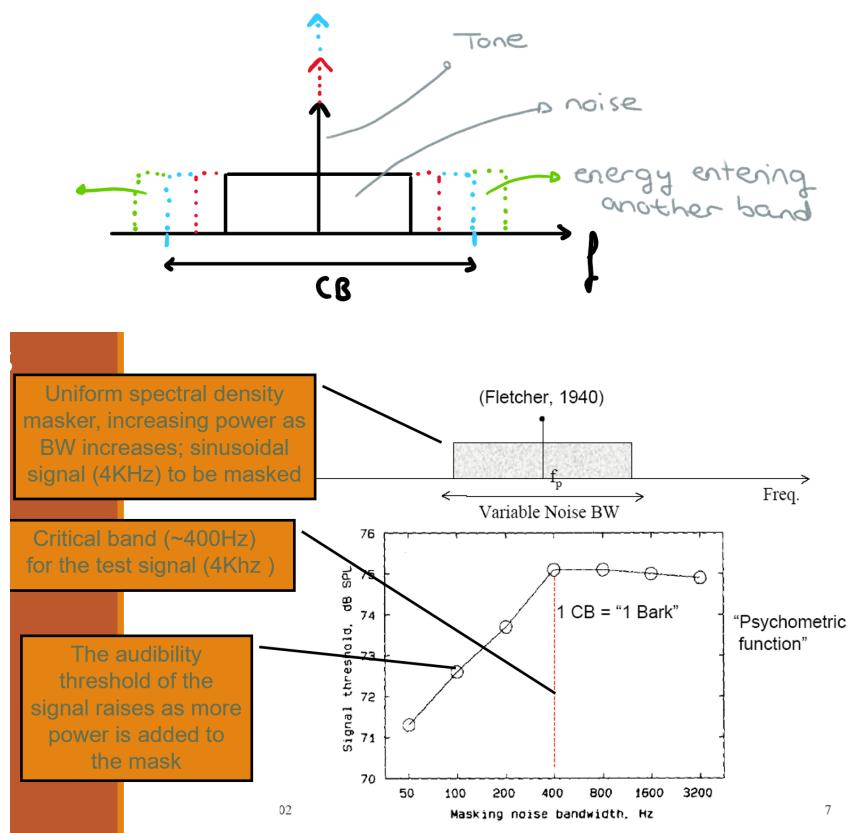
FIG. 2. The pitch function, showing the relation of perceived pitch (in mels) to the frequency of the stimulus. The values in mels are derived from the curve of Fig. 1; the 1000-cycle tone is arbitrarily assigned the value of 1000 mels.

Explain the classical experiment by Fletcher on measuring the critical bandwidth of a tone.

In 1940, Fletcher postulated that the auditory system behaves like a bank of pass-band filters with overlapping passband. He measured the threshold for detecting a sinusoidal signal as a function of the bandwidth of a bandpass noisemaker. The noise was always centered at the signal frequency and the noise power density was held constant. Thus, the total noise power increased as the bandwidth increased.

(María Aguado) My explanation of Fletcher's experiment:

A user is asked to listen to a tone with a noise band centered at it. The noise band is increased in steps, and thus the user needs to increase the tone loudness in order to continue hearing it. However, the user reaches a threshold from which the increasing steps in the noise bandwidth do not affect the hearing of the tone anymore. It can be directly inferred from this result that there exists a critical bandwidth for the specific tone. When the length of the bandwidth is exceeded by the noise, the energy goes into the contiguous bands.



Name and describe the workings of the usual scales or units that are available to measure the critical bandwidth.

Bark: a frequency scale on which equal distances correspond with perceptually equal distances. Above about 500 Hz this scale is more or less equal to a logarithmic frequency axis.

Below 500 Hz the Bark scale becomes more and more linear. The scale ranges from 1 to 24 and corresponds to the first 24 critical bands of hearing

ERB: The ERB of the auditory filter is assumed to be closely related to the critical bandwidth, but it is measured using the notched-noise method, rather than on classical masking experiments involving a narrow-band masker and probe tone. As a result, the ERB is said not to be affected by the detection of beats or intermodulation products between the signal and masker. Since this scale is defined analytically, it is also more smoothly behaved than the Bark scale data.

Loudness

Indicate where the absolute threshold of hearing is, and how “absolute” and how “linear” it is.

(María Aguado) Info from wikipedia I found useful:

The **absolute threshold of hearing (ATH)** is the minimum sound level of a pure tone that an average human ear with normal hearing can hear with no other sound present. The absolute threshold relates to the sound that can just be heard by the organism. The absolute threshold is not a discrete point, and is therefore classed as the point at which a sound elicits a response a specified percentage of the time. This is also known as the **auditory threshold**.

The threshold of hearing is generally reported as the sound pressure of 20 micropascals, i.e. 0 dB SPL. It is approximately the quietest sound a young human with undamaged hearing can detect at 1,000 Hz.

The threshold of hearing is frequency-dependent and it has been shown that the ear's sensitivity is best at frequencies between 2 kHz and 5 kHz, where the threshold reaches as low as -9 dB SPL.

Taking this into account, my answer would be:

- The threshold reference is around 20 micropascals for 1kHz, but this threshold is frequency-dependent.
- It is not “very absolute” in the sense that it is frequency-dependent and it varies from one person to another.
- Not linear. It is a curve. It can only be taken as linear in small regions.

What is a relative threshold? Which names does it take, alternatively?

A relative threshold in hearing is the minimum level of sound intensity that a person can detect at a specific frequency. It is also known as the minimum audible level or minimum audible sound pressure. The relative threshold is important because it helps determine the range of

frequencies that a person can hear, as well as their overall sensitivity to sound. It is typically measured in decibels (dB) and is determined by conducting a hearing test. In some cases, the relative threshold may also be referred to as the hearing threshold level or the hearing sensitivity.

(Amaia) I think that relative threshold explanation is:

Absolute threshold is the minimum sound level pressure needed to hear a pure tone, when there is no other sound simultaneously. However, once a sound or several sounds are already present, a new “absolute threshold” is generated, called “relative threshold”, which determines a new minimum sound level required for pure tones to be heard over that background “noise”. This threshold is created taking into account the critical bands that the background noise created (the masking effect).

What is the physical property of a waveform that mostly influences our loudness sensation?

The physical property of a waveform that mostly influences our loudness sensation is its amplitude, or the magnitude of its oscillations. In other words, the louder a sound is, the greater its amplitude will be.

Explain the purpose and meaning of the dB SPL scale

It is the unit we normally use to measure the Sound Pressure Level in the air. It's measured in decibels (dB) relative to a pressure of 20 micro pascals.

Explain the purpose and meaning of the dBHL scale

dBHL means dbs Hearing Level and it is used to measure hearing loss. The reference level varies with frequency according to a minimum audibility curve as defined in ANSI and other standards, such that the resulting audiogram shows deviation from what is regarded as 'normal' hearing.

0 : normal level of hearing

-20 : Crazy good level of hearing : hearing the sound at 20dB less than normal -> big difference.

Discuss if we can measure negative dB SPL and under which circumstances

Negative dB SPL can be measured and they mean that the sound is below the average hearing threshold (or below the reference value).

(María Aguado)

The dB SPL is negative under the condition $P < P_0$. If $P < P_0$ then $P/P_0 < 1$ and the logarithm of a number in the interval $(0,1)$ is always negative.

$$dB_{SPL} = 20 \log(P/P_0)$$

Notice that for $P = P_0$ the dB SPL is 0 since the quotient is 1.

Compare the loudness of a sinusoid, a square wave, and a sawtooth wave, all three peaking at -6dBFS and discuss the differences, if any.

Square waves sound louder than the other two. First of all, human hearing perceives things as louder if they are loud for longer, even if the maximum volume is the same, also, a square wave with the same peak amplitude as a sine wave will have a higher rms level. But it will also have numerous odd harmonics. There is no reason they should travel further. If it has a fairly low frequency, the harmonics will sound louder.

Is there any direct connection or equivalence between dB SPL and dBFS?

They are both measurements of different things and they are independent of one another. That said, a dB is dB, so a 10dB loss @ 10kHz measured by SPL will translate into a 10dB cut @ 10kHz in digital Full Scale. The only thing you would need to know would be the bandwidth used in the measurement to recreate it properly.

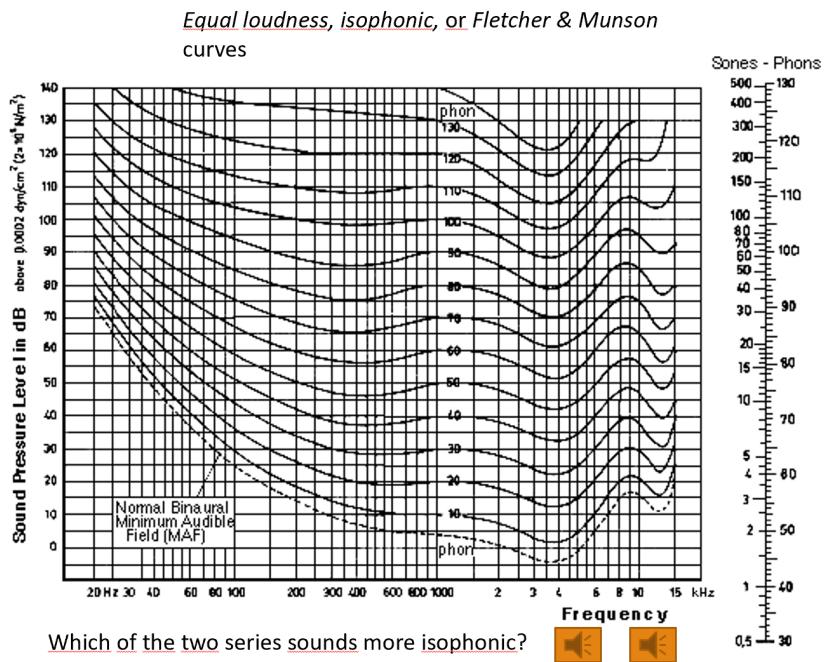
Define what does it mean that a certain sound generates 1 Sone and explain how you could get a 2 Sones sensation from the same sound.

1 sone is defined as the loudness of a sinusoid at 40 dB SPL and 1 kHz. An increase of 10dB gives a doubling in loudness. The same tone at 50 dB causes a doubling in loudness and will have a loudness of 2 sone, and so on. (60dB - 4 sones, 70dB - 8 sones).

Explain what equal-loudness curves tell us about our sound perception and how they can be obtained

Equal Loudness Curves, also known as isophonic curves, provide information of loudness across frequencies. You can obtain them with phones. The phon is another loudness unit, used for loudness comparison between sounds. 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.

(María Aguado) An equal-loudness curve is a curve that connects points with the same perceived loudness in a plane of dB SPL (y-axis) against frequency (x-axis). This shows us how loudness varies with frequency. It can be measured as we did in the lab, selecting the dB SPL for a fixed frequency and then varying frequency and choosing their respective dB SPL for the same perceptual loudness outcome.



Are sones and phones the same? Can they be converted into each other?

They are not the same. The unit of phon is based on sound pressure measured on a logarithmic (dB) scale. Sone is a more linear scale of perception. According to the sone scale, a 1 sone sound is defined as a sound whose loudness is equal to 40 phones.

sone	1	2	4	8	16	32	64
phon	40	50	60	70	80	90	100

Explain the 20 phones curve: what is it telling us about our sound perception?

They are a set of experimentally determined graphs that show how loud in dB SPL a sound at one frequency must be to be perceived as equally loud as a sound at another frequency. What it shows is that our hearing system is not linear. High pressure curves are a bit flatter than the low pressure curves.

What are the most usual loudness weighting scales, and which is their purpose?

dB(A) = 40 phon curve (approx) Normally used for soft sounds. To account for the relative loudness perceived by the human ear, as the ear is less sensitive to low audio frequencies. Originally defined for the measurement of low-level sounds (around 40 phon), is now commonly used for the measurement of environmental noise and industrial noise. However, if applied to noisy situations, it is not so good (it makes a wrong weighting), but it is very much used everywhere.

dB(B) = 70 phon curve

dB(C) = essentially flat, for high pressure levels

(María Aguado) A loudness weighting scale is a curve that gives us weights to adjust the loudness when measuring with instruments/devices, since loudness is frequency-dependent.

Found these next paragraphs on Google:

“A - weighting: a measurement scale that approximates the “loudness” of tones relative to a 40-dB sound pressure level, 1,000-Hz reference tone. A-weighting is said to best fit the frequency response of the human ear: when a sound dosimeter is set to A-weighting, it responds to the frequency components of sound much like your ear responds.

B-weighting: is similar to A-weighting but with less attenuation. B-weighting was an attempt to approximate human perception of loudness for moderately high sound pressure levels. It is now outdated and no longer used.

C-weighting: a measurement scale that approximates the “loudness” of tones relative to a 90-dB sound pressure level, 1,000-Hz reference tone. C-weighting has the added advantage of providing a relatively “flat” measurement scale that includes very low frequencies. Linear weighting: A weighting most commonly found on upper model sound level meters, typically used when performing octave band filtering analysis.”

Give a rough estimation of our JND for loudness and comment on its variability depending on some factors.

1dB decrease is close to the JND for loudness for mid and high-frequency tones or broadband noises.

(María Aguado) Answer generated by the ChatGPT AI:

The Just Noticeable Difference (JND) for loudness is typically estimated to be around 1-3 decibels. However, this value can vary depending on factors such as the individuals sensitivity to sound, the background noise level, and the specific frequencies being perceived. For example, individuals with high sensitivity to sound may have a lower JND for loudness, while those in a noisy environment may have a higher JND. Additionally, certain frequencies may be more easily detected than others, leading to differences in JND for different sounds.

jnd for intensity

- Jnd = Just Noticeable Difference (also Intensity Difference Limen)
- Methods
 - e.g. intensity discrimination between stimuli
- Weber fraction $\Delta I/I = \text{constant}$
(This means jnd is a constant value in dB, independent of the level!)
- Typical normal results
 - Wide band noise
20 - 100 dB; 0.5 - 1.0 dB
 - Tones
20 - 80 dB SL;
1.5 reducing to 0.3 dB
- “Near miss” to Weber law

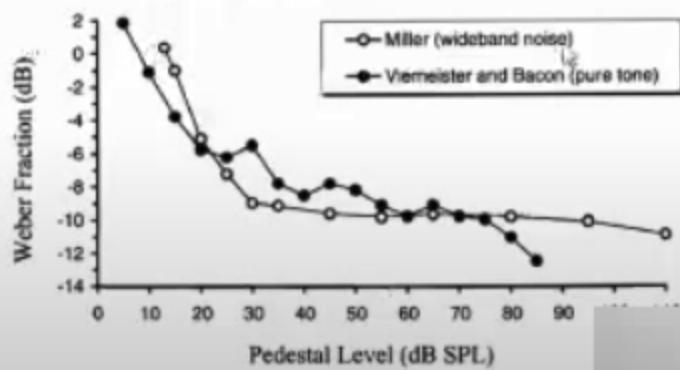


FIG. 6.6. Intensity discrimination for a wideband noise and for pure tone, as a function of level. In both experiments listeners required to detect a brief increment in a continuous sound. Data Miller (1947) and Viemeister and Bacon (1988).

Discuss the interaction of sound physical properties as duration and frequency with our loudness sensation

The human ear has a non-linear response across both volume and frequency. Greater amplitude waves have more energy and greater intensity, so they sound louder. The human ear is most sensitive to sounds from about 2 to 5 kHz and will perceive those as louder than a sound of lower or higher frequency but equal sound pressure.

(María Aguado) Answer generated by the ChatGPT AI:

Duration and frequency are two important physical properties of sound that interact with our perception of loudness. Duration refers to the length of time that a sound lasts, while frequency refers to the number of vibrations per second that a sound produces.

When it comes to our perception of loudness, duration and frequency play a key role. For instance, a sound with a short duration but high frequency will be perceived as louder than a sound with a longer duration but lower frequency. This is because the high frequency sound will have more vibrations per second, which will be more easily detected by our auditory system and thus perceived as louder.

In contrast, a sound with a long duration but low frequency will be perceived as less loud. This is because the low frequency sound will have fewer vibrations per second, which will be less easily detected by our auditory system and thus perceived as quieter.

Overall, the interaction of duration and frequency with our perception of loudness is complex and multifaceted. While these physical properties play a significant role in our ability to perceive loudness, they are not the only factors that influence our perception of sound. Other factors such as intensity and distance can also play a role in our perception of loudness.

Discuss the connection between loudness and the working of auditory filters

(Corcu, Maria, Sofia and ChatGPT AI)

The loudness of a sound is related to its amplitude, or the height of the sound waves.

Auditory filters, also known as "critical bands," are frequency-selective processes in the auditory system that help the brain to process sounds. These filters allow the brain to identify and focus on specific frequencies within a complex sound, such as a person's voice within a noisy environment.

In general, loud sounds are more likely to activate the auditory filters and be processed by the brain than softer sounds. This is because the auditory filters are more sensitive to higher-amplitude sounds. Therefore, the louder a sound is, the more likely it is to be noticed and processed by the brain. This is why loud sounds are often able to capture our attention more easily than softer sounds.

If you play a chord where all its notes are (roughly) concentrated in the same octave and then you play it, with the same intensity as before, but you spread the notes across 3 octaves, what could be happening with your loudness sensations?

(Chris) When you play a chord where all notes are roughly concentrated in the same octave some notes may share the same critical bandwidth, and as a consequence could contribute to a masking effect. When the same chord is spread out over three octaves they do not share the same critical bandwidth. As such they will be processed through separate filters and contribute to a greater sensation of loudness.

Elaborate a list of some practical consequences (for casual listening, for mixing music, for measuring sensations, for audio coding, etc.) of the facts you have dealt with in this lesson

Pitch

Define pitch and discuss if it is a physical property of sounds. Clarify about physical properties and sensations regarding pitch.

Pitch is a term used to describe the perceived highness or lowness of a sound. It is commonly associated with the frequency of a sound wave, with higher-frequency sounds being perceived as having a higher pitch and lower-frequency sounds being perceived as having a lower pitch.

However, it is important to note that pitch is not a physical property of sound. This is because pitch is a sensation that is created in the brain in response to certain physical properties of sound, such as frequency and amplitude. In other words, pitch is a psychological phenomenon, not a physical one.

Physical properties of sound, such as frequency and amplitude, are objective and can be measured using scientific instruments. In contrast, sensations, such as pitch, are subjective and can only be experienced by an individual.

Therefore, while pitch is related to the physical properties of sound, it is not a physical property itself. It is a sensation that is created in the brain in response to the physical properties of sound.

Is pitch a unidimensional sensation as loudness was? Define pitch chroma and pitch height.

Pitch is not a unidimensional sensation like loudness. While loudness is a single-dimensional sensation that can be described by a single number (the loudness level), pitch is a multi-dimensional sensation that is influenced by several factors, including the frequency and amplitude of a sound wave.

Pitch chroma refers to the "color" or "timbre" of a pitch, which is determined by the specific frequencies that are present in a sound wave. Pitch chroma is what allows us to distinguish between different musical notes, even when they have the same pitch height.

Pitch height, on the other hand, refers to the perceived highness or lowness of a pitch. Pitch height is determined by the frequency of a sound wave, with higher-frequency sounds being perceived as having a higher pitch and lower-frequency sounds being perceived as having a lower pitch.

Therefore, while pitch chroma and pitch height are related to one another, they are distinct aspects of the sensation of pitch. Pitch chroma allows us to distinguish between different pitches, while pitch height allows us to perceive the relative highness or lowness of a pitch.

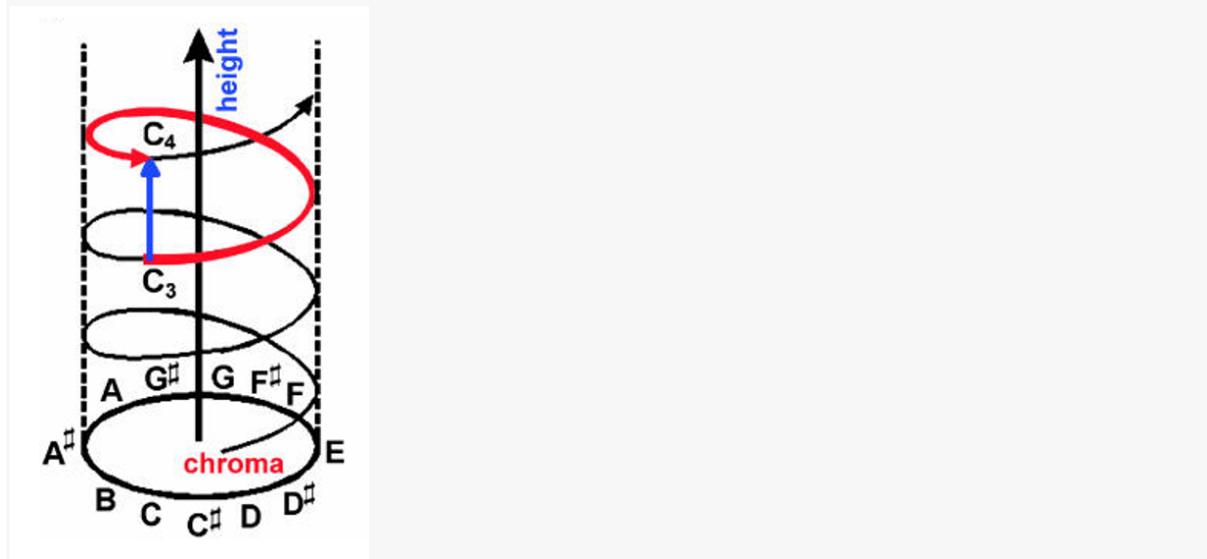
(María Aguado) Answer generated by the ChatGPT AI:

Pitch chroma refers to the specific identity of a pitch within a musical scale. It is the measure of how closely a pitch matches the specific pitches of a scale, such as C, D, E, F, G, A, and B in the Western music system.

Pitch height, on the other hand, refers to the perceived highness or lowness of a pitch. This is determined by the frequency of the pitch and is often measured in hertz (Hz). A higher pitch

will have a higher frequency and will be perceived as higher in pitch height, while a lower pitch will have a lower frequency and will be perceived as lower in pitch height.

In summary, pitch chroma refers to the specific identity of a pitch within a scale, while pitch height refers to the perceived highness or lowness of a pitch.



Discuss why many musical cultures do not use fundamental frequencies higher than 5kHz.

There's not many harmonics up to 5KHz so not many rich sounds.

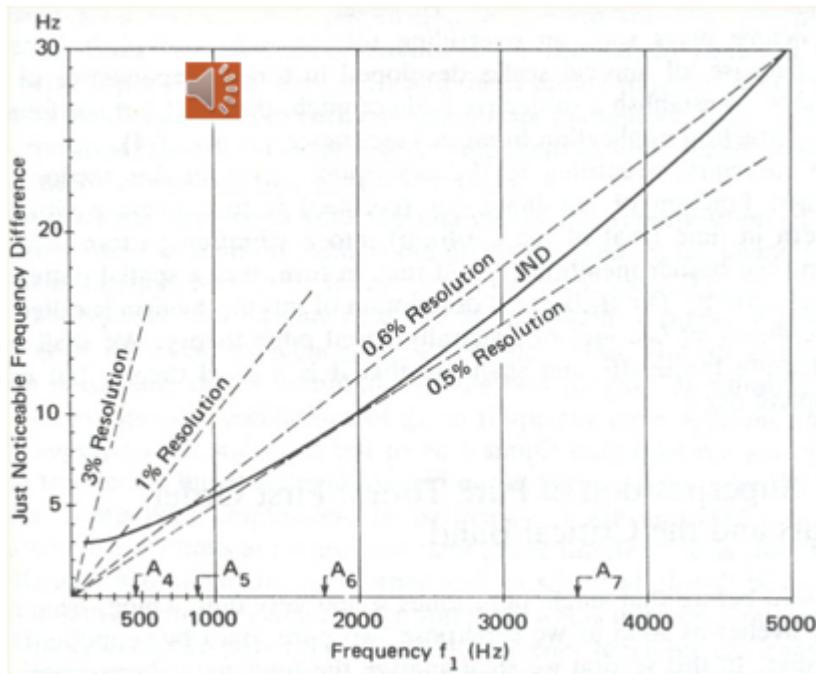
(María Aguado) Answer generated by the ChatGPT AI:

Many musical cultures do not use fundamental frequencies higher than 5kHz because the human ear is not sensitive to frequencies above this range. The range of frequencies that the human ear can hear is typically between 20Hz and 20kHz, with the highest sensitivity around 1-4kHz. Beyond 5kHz, the ear's sensitivity decreases significantly, making it difficult for us to perceive and enjoy sounds at these frequencies. Additionally, the production of high-frequency sounds requires a great deal of energy and skill, which may be difficult for some musical instruments or vocalists to achieve. As a result, many musical cultures have evolved to focus on the range of frequencies that are most easily perceived and produced by the human ear.

Describe the function depicting the JND for pitch. Are there flat sections? Is it a monotonically increasing one?

(María Aguado) Answer generated by the ChatGPT AI:

The JND for pitch is a monotonically increasing function that depicts the minimum detectable difference in pitch between two sounds. It typically has a U-shaped curve, with flatter sections at the lower and higher end of the pitch range. This indicates that the ability to detect small differences in pitch is greater at mid-range pitches compared to very low or very high pitches.



Explain the Mel scale and discuss the paradox inherent in the graph shown in the slides.

It's a perceptual scale of pitches judged by listeners to be equal in distance from one another.

The **mel scale** (after the word *melody*)^[1] is a perceptual scale of **pitches** to be equal in distance from one another. The reference point between the **frequency** measurement is defined by assigning a perceptual pitch of 100 Hz tone, 40 dB above the listener's threshold. Above about 500 Hz, **increasing intervals** are judged by listeners to produce equal pitch increments.

Contents [hide]

- [1 Formula](#)
- [2 History and other formulas](#)
- [3 Criticism](#)
- [4 See also](#)
- [5 References](#)

Does pitch show influences of physical properties of the sound in addition to those related to fundamental frequency and periodicity?

Pitch is determined by both the physical properties of a sound and by the way our auditory system processes that sound. In terms of the physical properties of a sound, pitch is primarily determined by the frequency of the sound wave. Sounds with a higher frequency are perceived as having a higher pitch, while sounds with a lower frequency are perceived as having a lower pitch. In addition to the frequency of a sound, the harmonic content of a sound wave can also affect its pitch. Sounds with a complex harmonic content, such as those produced by musical instruments, may be perceived as having a different pitch than a pure tone with the same frequency. Overall, pitch is influenced by both the physical properties of a sound, such as its frequency and harmonic content, as well as by the way our auditory system processes that sound. This interplay between the physical properties of sound and our auditory system allows us to perceive pitch in a wide range of sounds.

Describe absolute pitch and its influencing factors.

Its the ability of recognize the pitch of a single instance of a tone.

Pitch memory, as the ability of storing pitch information (developed by both musicians and nonmusicians)

Pitch labeling, as the ability of naming this concrete pitch (typically developed by musicians)

(María Aguado) Answer generated by the ChatGPT AI:

Absolute pitch, also known as perfect pitch, is the ability to identify or produce a specific musical pitch without the use of a reference tone. This ability is rare and is thought to be influenced by a combination of genetic and environmental factors. Some studies have suggested that early musical training, exposure to a tonal language, and innate cognitive abilities may play a role in the development of absolute pitch.

What is the so-called “missing fundamental”? How is it possible? Why is it challenging, with respect to the brain mechanisms involved in the pitch sensation?

It's a psychological effect that happens 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

(María Aguado) Answer generated by the ChatGPT AI:

The "missing fundamental" is a phenomenon in which the brain perceives a pitch that is not actually present in the sound signal. This is possible because the brain uses information from the harmonics (overtones) of a sound to construct a perception of its pitch. If the fundamental frequency (the lowest frequency component of a sound) is absent, the brain may still use the harmonics to create a perception of pitch. This can be challenging for the brain because it requires additional processing to determine the missing fundamental and create a consistent pitch perception.

Discuss some situation where your pitch strength sensation would be a weak one, and another one where it would be of strong pitch.

Pitch strength is determined by:

The microstructure of the waveform (the simpler, the stronger pitch)

The amplitude envelope (random phase -> weak pitch)

The strength of the pitch sensation can be influenced by many factors, including the volume and harmonic content of a sound. These factors can either enhance or weaken our perception of pitch.

(Chris) Pitch Strength is determined by two things: the micro structure of the waveform and the amplitude of the waveform. In terms of the micro structure, the more simple a waveform the stronger the pitch will be. If a waveform has many "ripples" it will make the pitch sensation less strong. In terms of amplitude, having a more constant amplitude across the waveform will contribute to a stronger sense of pitch. When there is amplitude modulation this causes the pitch to "blur." Additionally when considering amplitude we must consider the noise level across a spectrum. For example, when listening to band noise with a very narrow band we will start to perceive a pitch. When we widen this band we lose the sense of pitch and begin to hear white noise.

Explain the basic ideas under different theories of pitch sensation: place theory, timing theory, pattern matching theory.

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.

(María Aguado) Answer generated by the ChatGPT AI:

1. Place theory suggests that the perception of pitch is determined by the location of the firing of specific nerve fibers in the auditory nerve. This theory proposes that different frequencies stimulate different regions of the cochlea, and the brain interprets this spatial information as pitch.
2. Timing theory proposes that the perception of pitch is determined by the timing of nerve impulses in the auditory nerve. This theory suggests that different frequencies cause different patterns of firing in the auditory nerve, and the brain interprets this temporal information as pitch.
3. Pattern matching theory proposes that the perception of pitch is determined by the brain's ability to recognize and match patterns in the incoming auditory information.

This theory suggests that the brain compares the incoming sound patterns to stored patterns in order to identify pitch.

Timbre

Why is it said that timbre is a multidimensional sensation?

It is determined by multiple factors and can be perceived in different ways. Unlike pitch, which is primarily determined by the frequency of a sound wave, timbre is determined by the complex combination of frequencies that make up a sound. This combination of frequencies, known as the sound's spectrum, is what gives a sound its unique character or quality.

Enumerate different “dimensions” of timbre sensations and discuss their importance

- Roughness
- Fluctuation strength
- Sharpness (Brightness)

Why, when synthesizing a note of a musical instrument, it is not enough using the spectral shape or the addition of all the partials found in its spectrum?

(Chris): The timbre of a sound consists of a set of sensory attributes, some of which are continuously varying over time. For example, the amplitude envelope (the level of attack, decay, sustain, and release) for a particular partial can vary greatly across a sound. Other attributes are discrete and categorical. For example the ‘blatt’ of a trombone’s sforzando note. These attributes are not able to be modelled simply by the spectral shape.

Are timbre sensations involved in the identification of sound sources only?

Is there anything else in timbre?

(Chris): Timbre can also allow us to identify different classes of sounds. Formants are Spectral regions of high-energy, that impart distinctive characteristics to the sounds. They allow us to distinguish the vowels in human speech across many different languages.

Explain what formants are, their acoustic origin and the different ways we can take advantage of them (perceptually and technologically).

Formants are spectral regions of high-energy, that impart distinctive characteristics to the sounds. Allows us to distinguish the vowels. Prominent harmonics with prominent energy. Two formants can be enough to identify different classes of sounds.

For example, a singer’s formant is a skill developed by male singers to stand in front of the orchestra.

Formants can be used in a variety of ways. For example, formants can be used in speech processing applications, such as speech recognition and speech synthesis, to improve the accuracy and naturalness of these systems. Formants can also be used in music production, where they can be manipulated to create different timbres and to shape the sound of a musical instrument

What are the so-called Tartini tones? Are they physically in the waveform?

Tartini tones are a combination of tones.

At loud and very loud levels, 2 tones that are close in frequency may create a “ghost” tone corresponding to the sum and to the difference (Tartini Tone) of the frequencies.

Other extra tones may also appear, all of them obeying to a non-linear distortion process happening in the middle-ear or/and in the early auditory processing

Define beating and its applications in musical contexts.

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).

Define sensory consonance. How is it connected with the appreciation of different combinations of musical notes?

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.

Define roughness and name the unit to measure this subjective sensation; give an example for high and low roughness. Discuss its connection with the critical bandwidth. Discuss its connection with beating.

Expresses how rough a sound is. Musically, could be equivalent to “dissonance” Roughness indicates the rough-sounding perception of stimuli with AM or FM rates around 70 Hz. 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.

Define brightness and name the unit to measure this subjective sensation; give an example for high and low brightness.

It is a subjective sensation that describes the perceived sharpness or clarity of a sound, and is typically measured using the unit of brightness. Sounds with a high brightness are typically perceived as sharp and clear, while sounds with a low brightness are typically perceived as dull and muffled.

What is a timbre space? How is it created? What is the information that we can extract from it?

A timbre space is a mathematical representation of the timbre of a sound, and can be used to extract information about the relative position, distance, and overall structure of sounds within the space. This information can be used to identify sounds, compare the timbre of different sounds, and gain insight into the characteristics that define the timbre of a particular class of sounds.

Discuss possible techniques to do research on the description of timbres by humans.

There are several techniques that can be used to do research on the description of timbres by humans. These techniques include self-report measures, stimulus classification tasks, and similarity ratings.

Self-report measures are a commonly used technique for studying the description of timbres by humans. In this technique, participants are asked to listen to a series of sounds, and to describe the timbre of each sound using words or phrases. For example, participants might be asked to describe the timbre of a sound as "bright," "dark," "hollow," or "smooth." This technique can provide useful information about the words and phrases that people use to describe the timbre of sounds, and can help to identify the dimensions or qualities that are most important for timbre perception.

Stimulus classification tasks are another technique that can be used to study the description of timbres by humans. In this technique, participants are asked to listen to a series of sounds, and to classify the timbre of each sound into one of several pre-defined categories. For example, participants might be asked to classify the timbre of a sound as "percussion," "strings," "brass," or "woodwinds." This technique can provide more objective information about the way that people categorize and classify the timbre of sounds, and can help to identify the dimensions or qualities that are most important for timbre perception.

Similarity ratings are a third technique that can be used to study the description of timbres by humans. In this technique, participants are asked to listen to pairs of sounds, and to rate the similarity of the timbre of each pair of sounds. For example, participants might be asked to listen to a sound that has a "bright" timbre and a sound that has a "dark" timbre, and to rate how similar the two sounds are in terms of their timbre. This technique can provide information about the way that people perceive the similarity of different timbres, and can help to identify the dimensions or qualities that are most important for timbre perception.

Overall, there are several techniques that can be used to do research on the description of timbres by humans, including self-report measures, stimulus classification tasks, and similarity ratings. These techniques can provide useful information about the way that people perceive and describe the timbre of sounds, and can help to identify the dimensions or qualities that are most important for timbre perception.

Discuss why we have not mentioned absolute and relative thresholds in this lesson. Could we have dealt with some thresholds anyway? Develop your ideas.

We have not mentioned absolute and relative thresholds in this lesson because these concepts are not directly related to the topics that we have discussed. Absolute and relative thresholds are concepts in psychology and neuroscience that refer to the minimum level of stimulation that is required to produce a detectable sensation or response. These concepts are used to study the sensitivity of our senses, and to understand the limits of our perception.

However, absolute and relative thresholds are not directly related to the topics that we have discussed in this lesson, which include the structure and organization of musical sounds, the way that our brains process and perceive musical information, and the emotional responses that we have to music. While absolute and relative thresholds are important concepts in psychology and neuroscience, they are not central to the study of music perception and cognition.

However, we could have discussed some thresholds in the context of music perception and cognition. For example, we could have discussed the minimum level of stimulation that is required to produce a detectable musical pitch or timbre, or the minimum level of contrast that is required for us to perceive a change in the musical sound. These thresholds can be important for understanding the limits of our musical perception, and for studying the way that our brains process and perceive musical information.

Overall, while we have not mentioned absolute and relative thresholds in this lesson, these concepts could have been discussed in the context of music perception and cognition. Absolute and relative thresholds are important concepts in psychology and neuroscience, and can provide useful information about the sensitivity of our senses and the limits of our perception.

Perceptual organization

Explain the concept of stream or streaming, applied to our hearing

Stream is a percept of a group of successive and/or simultaneous sounds as a coherent whole appearing to come from a single source (for example in an orchestra, a brass section)

Explain the differences between spectral and sequential integration

Simultaneous/Spectral grouping: the grouping together of the simultaneous frequency components that come from a single source.

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

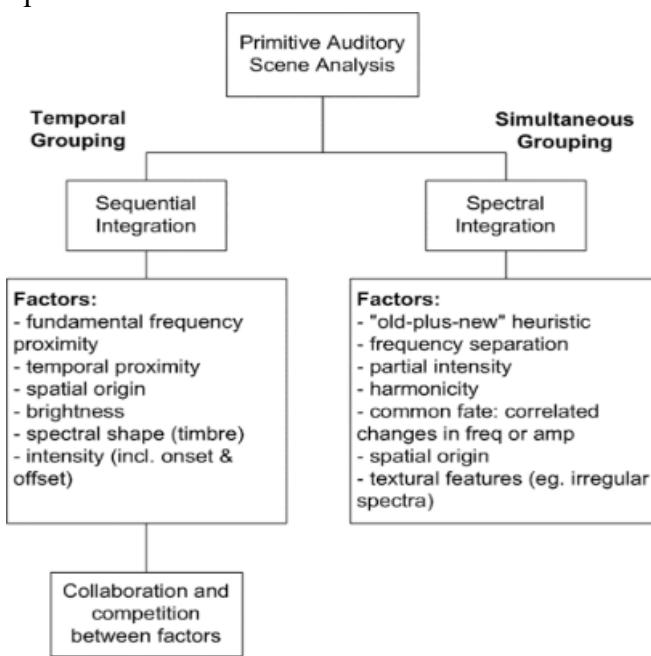
Sequential grouping: the connecting over time of the changing frequencies that a single source produces from one moment to the next.

Periodicity cues

Spectral cues

Level cues

Spatial cues



What are the bases used for building up the hypotheses that try to organize our auditory scenes?

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. Gestalt argues that there are some innate principles of perceptual organization, which are, applied to sound: proximity, similarity, common fate, good continuation, belongingness and closure.

How far the hypotheses that guide our auditory scene analysis can be consciously controlled?

Auditory scene analysis is the process by which the brain organizes and interprets the sounds that we hear. The hypotheses that guide this process are based on our past experiences and the

context in which the sounds are heard. While we may not be consciously aware of the specific hypotheses that guide our auditory scene analysis, we can influence the process to some extent by paying attention to the sounds we hear and consciously trying to make sense of them. For example, if we are trying to follow a conversation in a noisy room, we can focus our attention on the speaker's voice and try to ignore other distracting sounds. This can help our brain more effectively analyze the auditory scene and allow us to understand the conversation better.

Enumerate some Gestalt laws and discuss its relevance in listening to music

A description of auditory and visual Gestalt Laws.

Name	Audition	Vision
Proximity (Belongingness)	Sounds arriving from places <i>close in space</i> tend to be grouped	Elements <i>close together</i> in space tend to be grouped
Similarity	Sounds with <i>similar timbre and pitch</i> tend to be grouped	Elements <i>shaped alike</i> tend to be grouped
Good Continuation	Sounds that follow a <i>regular pitch contour</i> tend to be grouped	Elements that follow a <i>regular spatial contour</i> tend to be grouped
Closure	Interrupted auditory stimuli tend to be perceived as <i>continuous</i> when plausible	<i>Borders are interpreted/completed to specify shapes</i>
Simplicity (Pragnanz)	Frequencies with <i>simple harmonic ratios</i> tend to be grouped	<i>Prototypical shapes</i> tend to be regular, simple, symmetric
Common Fate	Sounds with <i>synchronous rhythm patterns</i> tend to be grouped	Elements that <i>move together</i> tend to be grouped

Discuss the “figure-ground” phenomena when listening to natural sounds and when listening to music

The gestalt notion "figure-ground phenomenon" refers to the **characteristic organization of perception into a figure that 'stands out' against an undifferentiated background**. What is figural at any one moment depends on patterns of sensory stimulation and on the momentary interests of the perceiver. Figure-ground relationship is an important element of the way we organise reality in our awareness, including works of art. The momentary interests of the perceiver can be manipulated to a considerable extent by the performer, by rather evasive cues: in different performances different threads of the patterns of sensory stimulation may be foregrounded, by mild shifts of attention to and fro, as, in the visual arts.

Discuss how far can go our capacity to perceive “separated” instrumental lines. Does it resemble what source separation algorithms usually do?

I don't know how I would answer this, but it would be something like:

I guess our capacity to perceive separated instrumental lines depends on our abilities in terms of distinguishing different timbres, frequencies...(JNDs in general, the farthest limit is where our physical limitations are). Our musical abilities, our past experiences and our knowledge also influence.

However, Perfe mentioned in class that in terms of <<simultaneity>>, normal human beings can at most follow 2-3 different separated streams. Maybe musicians can reach up to 5 lines of melodies followed at the same time.

Can you give some examples of musical performance/composition practices that may be grounded on principles of perceptual organization?

It is difficult to say for certain which songs may be specifically grounded on principles of perceptual organization, as this can vary depending on the interpretation and perception of the listener. However, here are a few examples of songs that could potentially be seen as incorporating principles of perceptual organization:

"Bohemian Rhapsody" by Queen: This classic song is known for its complex structure and use of repetition and variation. The song is divided into several distinct sections, each with its own unique melody and arrangement, but many of the sections also include repeated elements (such as the famous "Galileo" melody) that help to create a sense of unity and coherence.

"Stairway to Heaven" by Led Zeppelin: This iconic song is known for its use of dynamics and timbre to create a sense of contrast and development within the piece. The song starts with a quiet, acoustic guitar-based introduction, but gradually builds in intensity and complexity, eventually reaching a climactic electric guitar solo at the end. The use of different instrumentation and timbres throughout the song helps to create a sense of progression and evolution.

"Rhapsody in Blue" by George Gershwin: This classic composition is known for its use of timbre to create contrast and variety within the piece. The piece is written for a solo piano, but also incorporates a full orchestra, and the use of different instruments and timbres helps to create a rich and complex soundscape. Additionally, the piece is known for its use of classical compositional techniques, such as theme and variation, which can also be seen as a form of perceptual organization.

Provide a generic explanation for some (or most) musical illusions. Discuss the role of “ambiguity” in many of them

The auditory system is very prone to illusion. Considering the paucity of information that is supplied by the peripheral auditory system it is not surprising. There are around 16,000 peripheral receptors in each ear, compared with 125 million receptors in each eye. In addition, the amount of brain tissue devoted to hearing is far less than is devoted to vision. As a further consideration, sound signals are mixed together and are subject to numerous distortions as they

travel from their sources to our ears, so that the brain is faced with the very difficult task of reconstructing the original signals from the mixture of sounds that arrive at the ears.

What does it mean “bi-stability” in the context of these topics?

It is the phenomenon in which perception switches between different interpretations of an unchanging stimulus, are characterised by very similar properties across a wide range of qualitatively different paradigms.

Music and memory

Enumerate the different “big” memory storage systems, according to their functions, and discuss their properties and interactions

Iconic memory: Iconic memory involves the memory of visual stimuli. The word iconic refers to an icon, which is a pictorial representation or image. Icon memory is how the brain remembers an image you have seen in the world around you. ... Iconic memory is a type of sensory memory that lasts just milliseconds before fading.

Echoic memory: sensory memory that registers specific to auditory information. Once an auditory stimulus is heard, it is stored in memory so that it can be processed and understood. Stored for slightly longer periods of time than iconic memories (visual memories).

Short term memory (STM) (working 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, and also specific sensory blocks for parallel computing.

Long term memory (LTM): It is the “memory” as commonly understood.

Is long-term memory a single system? Enumerate the most relevant subsystems that play a role in our musical experiences

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.

Are musical memories localized in a single brain region (mostly in the A1 cortex)?

There is no single brain region that is responsible for musical memories. Instead, musical memory is thought to involve a distributed network of brain regions, including the A1 cortex

(part of the primary auditory cortex) as well as other areas of the cortex involved in auditory processing, such as the inferior frontal gyrus and the superior temporal gyrus. Additionally, musical memories can also be influenced by other brain regions involved in emotion, memory, and spatial processing. For example, the amygdala and hippocampus are known to play a role in emotional and declarative memory, respectively, and may be involved in the formation of musical memories. Similarly, the parietal lobe is known to be involved in spatial processing, and may play a role in our ability to remember the structure and arrangement of a piece of music. Overall, the formation and recall of musical memories is a complex process that involves the interaction of multiple brain regions.

Discuss the role of segmentation, grouping and phrasing in our online parsing of musical input (i.e., as we hear music in real-time)

To overcome the limited capacity of the short-term memory, some “tricks” are used in order to build a structure upon the plane stimulus. The continuous musical texture is broken into shorter sequences using “**segmentation cues**”:

- Closure and change detection
- Pauses, silences, stretching of notes
- Instrument changes
- Cadences
- Accents and other metrical elements
- Tendency changes, contrasts (up-down melody, long-short notes, etc)

Phrasing is the way notes are joined/separated/played into a meaningful “unit”, which separates from the previous one and from the next one. It is a musical resource to help segmentation in STM.

Give examples of declarative and non-declarative musical memories

Declarative memories are memories for facts and events that can be consciously recalled and verbalized. Non-declarative memories, also known as procedural memories, are memories for skills and habits that are learned and performed automatically, without conscious effort. In the context of musical memories, some examples of declarative memories might include:

- The title and composer of a particular piece of music
- The lyrics to a song
- The names of the musicians in a band

Examples of non-declarative musical memories might include:

- The ability to play a musical instrument
- The ability to sing a song without looking at the lyrics
- The ability to tap along to a beat or keep time with a metronome

Overall, both declarative and non-declarative memories play an important role in our ability to process and understand music. Declarative memories help us to remember specific facts and

details about a piece of music, while non-declarative memories allow us to perform and appreciate music on a more automatic and intuitive level.

What is implicit memory? What is its role in our memory experiences?

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.

Which are the different operations that make possible to “extract” knowledge from our memories?

Types of retrieval:

- a) Recognition: acknowledgement that a pattern in STM is stored in LTM
- b) Recall or Recollection: activation of a LTM encoded pattern by a “diffuse” effort of will
- c) Reminding: activation of a LTM

encoded pattern by a pattern in STM (which memories are activated by listening to this song?)

What is the role of motives and themes in the working of our musical memories?

They're musical strategies to increase the memorability and association of musical materials.

Describe the role and functioning of the hippocampus in our encoding and decoding of musical memories

For encoding:

Segmenting: The continuous musical texture is broken into shorter sequences using “segmentation cues”

Grouping: The Gestalt laws of grouping are a set of principles to account for the observation that humans naturally perceive objects as organized patterns and objects. Five categories: Proximity, Similarity, Continuity, Closure and Connectedness.

Chunking: Encoding or consolidation of small groups of elements into a compact larger or more abstract element, which is then encoded, recognized or remembered.

For decoding:

Recognition: acknowledgement that a pattern in STM is stored in LTM (did you hear this song before? Does it sound familiar?) Recognition may happen without further recall or reminding.

Recall or Recollection: activation of a LTM encoded pattern by a “diffuse” effort of will (what is the title of this song? Who was the composer?)

Reminding: activation of a LTM encoded pattern by a pattern in STM (which memories are activated by listening to this song?)

Discuss the possibility that many of our musical memories have an "absolute" quality (i.e., they are mostly literal instead of having just stored a summary, prototype, partial information, etc.)

There is some debate among researchers about the extent to which our musical memories have an "absolute" quality. Some studies have suggested that our musical memories may be more literal and detailed than memories for other types of information, such as verbal or visual information. For example, people often have very accurate and detailed memories for the melodies and harmonies of songs that they have heard, and can accurately reproduce these melodies even after many years. This suggests that our musical memories may be particularly well-preserved and detailed.

However, it is also possible that our musical memories are not always "absolute" in the sense of being completely literal and detailed. It is well-known that our memories can be influenced by various factors, such as the context in which the information is presented, our emotional state, and our prior knowledge and experiences. As a result, our musical memories may be subject to similar influences, and may not always be completely literal and accurate. For example, if we hear a song in a noisy environment, or if we are not paying close attention to the music, our memories of the song may be less detailed and accurate than if we had heard the same song in a quieter, more focused setting.

Overall, while it is possible that many of our musical memories may have an "absolute" quality, it is also likely that our musical memories are subject to the same kinds of influences and biases as our memories for other types of information. As a result, the "absolute" quality of our musical memories may vary depending on the specific circumstances in which the music is experienced and remembered.

How the experience of disentangling scrambled melodies reveals our perceptual and mnemonic mechanisms?

Frequency separation progressively improves recognition of the melodies
Intensity first, then timbre are used to favor groupings of notes, improving the identification of the melodies.

What/When Hierarchies

Indicate the differences, if any, between "music perception" and "music cognition"

Music perception and music cognition are two related but distinct areas of study within the field of music psychology. Music perception refers to the process by which we become aware of and understand the various components of music, such as melody, harmony, and rhythm.

This includes both the physiological processes that take place in the auditory system, as well as the psychological processes that allow us to make sense of the sounds that we hear.

In contrast, music cognition refers to the mental processes involved in creating, comprehending, and responding to music. This includes higher-level cognitive processes such as memory, attention, and emotion, as well as more specific musical processes such as tonality, meter, and harmony. Music cognition is concerned with understanding how the brain processes and represents musical information, and how this information is used to guide our musical behaviors and preferences.

Therefore, while both music perception and music cognition are concerned with understanding how the brain processes and responds to music, they differ in the specific aspects of this process that they focus on. Music perception focuses on the physiological and psychological processes involved in becoming aware of and understanding musical sounds, while music cognition focuses on the higher-level cognitive processes involved in creating, comprehending, and responding to music.

Why do we use the term “hierarchy” in our account of the musical experience?

The term "hierarchy" is used in discussions of the musical experience to refer to the way in which different elements of music, such as melody, harmony, and rhythm, can be arranged in a hierarchical structure, with some elements being more prominent or important than others. This hierarchical structure can affect the overall experience of listening to a piece of music, with certain elements drawing the listener's attention and influencing their perception of the music. The use of the term "hierarchy" in this context allows us to describe and analyze the musical experience in a systematic and organized way, by breaking it down into its constituent parts and considering the relationships between them.

Discuss the role of statistical or implicit learning in the development of the “what” and the “when” hierarchies

Statistical or implicit learning plays a crucial role in the development of both the "what" and the "when" hierarchies in music. The "what" hierarchy refers to the way in which different musical elements, such as melody, harmony, and rhythm, are arranged in a hierarchical structure, with some elements being more prominent or important than others. Through statistical learning, the brain is able to make predictions about which elements are likely to occur in a piece of music based on previous experiences and exposure to similar music. This allows the brain to identify the most important elements and to focus on them when listening to the music.

Similarly, the "when" hierarchy refers to the way in which different musical events unfold over time, with some events being more significant or noteworthy than others. Through implicit learning, the brain is able to predict the timing of musical events based on previous experiences and exposure to similar music. This allows the brain to anticipate and prepare for significant

events, such as the arrival of a new melody or a change in the harmony, and to better understand the overall structure and progression of the music.

Overall, statistical and implicit learning play a key role in the development of both the "what" and the "when" hierarchies in music, allowing the brain to make predictions and anticipate musical events, which in turn enables us to experience and understand the music in a more profound and meaningful way.

Enumerate some of the elements or layers of the “when” framework and their role in listening to music. Are they operating similarly when performing music?

Rhythm hierarchies (when): pulse, tempo, meter make certain “moments” more important than others. “When” is related to: rhythm, temporal organization...

Discuss the role of anticipation/expectation in music, and the role of the so-called what/when frameworks as facilitators of hypothesis about the real-time experience of music

Anticipation and expectation play a significant role in the experience of music, as they allow us to make predictions about what is likely to happen next in the music and to prepare for upcoming events. This can help us to better understand and appreciate the music, as we are able to follow its structure and progression and to anticipate significant events, such as the arrival of a new melody or a change in the harmony.

The what/when frameworks, also known as the "what" and "when" hierarchies, are two conceptual frameworks that are often used to describe and analyze the real-time experience of music. The "what" hierarchy refers to the hierarchical arrangement of different musical elements, such as melody, harmony, and rhythm, with some elements being more prominent or important than others. The "when" hierarchy, on the other hand, refers to the way in which different musical events unfold over time, with some events being more significant or noteworthy than others. These frameworks provide a useful framework for making predictions about the real-time experience of music and for facilitating hypotheses about the structure and progression of a piece of music.

Overall, anticipation and expectation play a key role in the experience of music, and the what/when frameworks provide a useful tool for making predictions and hypotheses about the real-time experience of music. These frameworks can help us to better understand and appreciate the music by allowing us to anticipate and prepare for upcoming events, and by providing a framework for understanding the hierarchical arrangement of different musical elements.

Do we need to take formal music lessons for the hierarchies to develop?

It is not necessary to take formal music lessons in order for the "what" and "when" hierarchies to develop. These hierarchies are a product of the brain's ability to make predictions and to learn from experience, and can develop through exposure to music through listening and other means. However, taking formal music lessons can provide a structured and systematic approach to learning about music, and can help to develop a more nuanced understanding of the hierarchical structure of music and the relationships between different musical elements. Formal music lessons can also provide opportunities to learn about musical notation and theory, which can further deepen one's understanding and appreciation of music.

Define the concept of IOI and explain its role in the “when” framework

IOI (Inter-Onset Interval) is a term used in music psychology to refer to the duration of time between two successive onsets (or attacks) in a musical piece. IOIs can be measured in a variety of ways, depending on the specific characteristics of the music being analyzed. For example, IOIs can be measured in terms of the number of beats between onsets, the number of seconds between onsets, or in terms of the musical meter (e.g. as a fraction of a measure or bar).

IOIs play an important role in the "when" framework of musical perception, which refers to the way that we process and make sense of the temporal structure of a musical piece. In particular, IOIs can help us to identify the underlying rhythmic structure of a piece of music, and can also help us to perceive the relative distances between different events in the music. For example, if two onsets are separated by a longer IOI, we may perceive this as a longer pause or gap in the music, whereas if the IOIs are shorter, we may perceive the onsets as being closer together in time.

Overall, the concept of IOI is an important part of the "when" framework of musical perception, and can help us to understand how our brains process and make sense of the temporal structure of music.

Discuss how some limitations in the memory and information processing of our brain impacts on the

The memory and information processing capabilities of our brains are not perfect, and there are various limitations that can impact our ability to process and remember musical information. For example, our brains have a limited capacity for storing and processing information, which can make it difficult for us to remember large amounts of musical information. Additionally, our memories are subject to various forms of decay and interference, which can cause us to forget or confuse musical information over time.

These limitations can impact our ability to process and remember musical information in various ways. For example, if we are trying to learn a new piece of music, we may have difficulty remembering all of the details of the piece if it is very long or complex. Similarly, if we are trying to recall a piece of music that we have heard before, we may have difficulty remembering all of the details if it has been a long time since we last heard the piece.

Overall, the limitations of our memory and information processing systems can impact our ability to process and remember musical information in various ways. While our brains are capable of processing and remembering a great deal of musical information, there are limits to what we can store and recall, and these limitations can make it more challenging to fully appreciate and understand complex or unfamiliar pieces of music.

Relate different temporal scopes of our musical experiences with different musical elements of our what/when frameworks

The "what" and "when" frameworks are two important concepts in music psychology that refer to the way that we process and make sense of the various elements of a piece of music. The "what" framework refers to the way that we perceive and process the pitch, timbre, and other qualities of a piece of music, while the "when" framework refers to the way that we perceive and process the temporal structure of a piece of music, including its rhythmic and melodic elements.

Different temporal scopes of our musical experiences can be related to different musical elements in the "what" and "when" frameworks. For example, at the shortest temporal scope, we might focus on individual notes or sounds within a piece of music, and pay attention to their specific pitches, timbres, and durations. This would involve processing the "what" elements of the music, such as its specific pitches and timbres.

At a slightly longer temporal scope, we might focus on the structure and arrangement of the notes within a melody or phrase, and pay attention to the way that the notes are organized and related to each other. This would involve processing both the "what" and "when" elements of the music, such as the pitches of the notes and the rhythms and melodies that they create.

At the longest temporal scope, we might focus on the overall structure and form of a piece of music, and pay attention to the way that the different sections and movements of the piece fit together and relate to each other. This would involve processing both the "what" and "when" elements of the music, as well as the larger-scale relationships and patterns within the piece.

Overall, different temporal scopes of our musical experiences can be related to different musical elements in the "what" and "when" frameworks, and the way that we perceive and process these elements can vary depending on the specific focus and scope of our musical experiences.

Why the “what” hierarchy is built on musical categories such as “notes”, “scales”, “modes”, etc?

The "what" hierarchy in music psychology is built on musical categories such as "notes," "scales," and "modes" because these categories are the building blocks of musical structure and organization. Notes are the individual sounds that make up a piece of music, and form the basis of melody, harmony, and rhythm. Scales are sequences of notes arranged in a particular order,

and form the basis of tonality and key. Modes are variations of scales that are used in different types of music, and can give a piece of music a unique character and flavor.

Together, these musical categories form a hierarchical structure that allows us to organize and understand the various elements of a piece of music. For example, when we hear a piece of music, our brains use the "what" hierarchy to identify the individual notes, the scales and modes that they form, and the larger-scale patterns and relationships that these elements create. This allows us to make sense of the music and to perceive its structure and organization.

Additionally, the "what" hierarchy is built on these musical categories because they are universal and are used in many different types of music. Notes, scales, and modes are found in almost all styles of music, and are used to create the melodies, harmonies, and rhythms that we hear in a wide variety of musical genres. As a result, the "what" hierarchy provides a common framework that can be used to understand and analyze music from many different cultures and traditions.

Discuss the role of categorization in our music understanding

Categorization is an important process in our understanding of music. When we hear a piece of music, our brains use various cues to identify and categorize the sounds that we hear. For example, we may use the pitch, timbre, and duration of a sound to determine whether it is a note, a chord, or some other type of musical event. We may also use the context in which the sound is heard, as well as our prior knowledge and experiences, to determine which category the sound belongs to.

Categorization is important because it helps us to make sense of the music that we hear. By categorizing the sounds that we hear, we can organize and structure the music in our minds, and this can make it easier for us to understand and remember the music. Additionally, categorization allows us to communicate about music with others, by using common labels and categories to describe the sounds and structures that we hear.

Overall, the process of categorization plays an important role in our understanding of music. By using various cues and prior knowledge to identify and categorize the sounds that we hear, we can better understand and appreciate the music that we listen to.

Explain what a scale is and its importance to frame our experience of the musical “what” Define melody contour and melody schema. Are they the same?

A scale is a sequence of musical pitches arranged in a particular order. Scales are typically built on a series of intervals (the distances between consecutive pitches), and the specific intervals used in a scale can determine its overall character and quality. For example, a major scale consists of the intervals 1-1-1/2-1-1-1-1/2, while a minor scale consists of the intervals 1-1/2-1-1-1/2-1-1.

Scales are important to our experience of the musical "what" because they provide a framework for organizing and understanding the pitches that we hear in a piece of music. When we hear a piece of music, our brains use the pitches that we hear to determine the underlying scale or tonality of the piece. This allows us to make sense of the pitches and to perceive their relationships and patterns.

Melody contour and melody schema are two related concepts in music psychology. Melody contour refers to the overall shape and direction of a melody, as determined by the pitches and intervals of the melody. Melody schema refers to our mental representations of melodies, which can include information about their contour, as well as other characteristics such as their rhythm, tempo, and tonality.

While melody contour and melody schema are related concepts, they are not the same thing. Melody contour refers to the specific pitches and intervals of a melody, while melody schema refers to our mental representations of melodies, which can include information about their contour as well as other characteristics. Additionally, while melody contour is a property of a specific melody, melody schema is a more general concept that can apply to many different melodies.

Define tonal hierarchy and how does it play in the game of music

Tonal hierarchy is a concept in music psychology that refers to the way that different pitches and chords are organized and related to each other within a piece of music. In tonal music (which is music that is based on a specific tonality or key), the pitches and chords are organized into a hierarchical structure, with some pitches and chords being perceived as more stable and important than others. This hierarchy is based on the relationships between the pitches and chords, and can determine the overall structure and organization of the music.

In the game of music, tonal hierarchy plays an important role in the way that we perceive and understand the pitches and chords that we hear. When we listen to a piece of music, our brains use the tonal hierarchy to make sense of the pitches and chords, and to determine the overall structure and key of the piece. This allows us to perceive the music as coherent and organized, and to appreciate its melody, harmony, and rhythm.

Additionally, tonal hierarchy can also play a role in the way that we perform and compose music. For example, when we are improvising or composing a piece of music, we may use the tonal hierarchy to guide our choices of pitches and chords, and to create a sense of structure and coherence in the music. Similarly, when we are performing a piece of music, we may use the tonal hierarchy to interpret and express the music in a way that is consistent with its underlying structure and tonality.

Overall, tonal hierarchy is an important concept in the game of music, and plays a key role in the way

Is there any connection between the temporal and pitch-based elements of our what/when hierarchies for music understanding?

There is a connection between the temporal and pitch-based elements of our "what" and "when" hierarchies for music understanding. In particular, the pitches and rhythms of a piece of music are closely related, and the way that we perceive and process these elements can be influenced by each other.

For example, when we hear a melody, the pitches and rhythms of the melody are closely linked, and our brains use this information to determine the overall structure and organization of the melody. The pitches of the melody can help us to identify the underlying tonality or key of the piece, while the rhythms of the melody can help us to perceive its melody and phrasing.

Similarly, when we hear a rhythm, the pitches and durations of the individual notes within the rhythm can influence the way that we perceive and understand the rhythm. For example, if the pitches of the notes in a rhythm form a specific melody or chord progression, this can give the rhythm a particular character and flavor. Additionally, the durations of the notes in a rhythm can also influence the way that we perceive the rhythm, and can help us to determine its tempo and meter.

Overall, there is a connection between the temporal and pitch-based elements of our "what" and "when" hierarchies for music understanding, and the way that we perceive and process these elements can be influenced by each other.

How “arbitrary” are musical systems in the world? Are they, though, completely arbitrary?

Musical systems in the world are somewhat arbitrary, in the sense that the specific rules and conventions of a musical system are determined by the culture and tradition in which the music is created. For example, the rules for constructing melodies and harmonies, the scales and modes that are used, and the instruments and timbres that are employed can vary greatly from one culture to another. As a result, the specific rules and conventions of a musical system are not universal, but are instead determined by the specific cultural and historical context in which the music is created.

However, musical systems are not completely arbitrary. Even though the specific rules and conventions of a musical system can vary from one culture to another, there are also some underlying principles and structures that are common to many different musical systems. For example, most musical systems are based on some form of pitch organization, in which the pitches of the music are arranged in a particular sequence or scale. Additionally, most musical systems use some form of rhythmic organization, in which the durations of the notes are arranged in a particular pattern or meter.

Overall, while the specific rules and conventions of a musical system may be somewhat arbitrary, there are also some underlying principles and structures that are common to many

different musical systems. As a result, while the details of a musical system may vary from one culture to another, the basic principles of pitch and rhythm organization are found in many different musical traditions.

Putting this question in other words, how far and in which ways are they constrained by some aspects of the way our brain processes information and relates with the physical world?

Musical systems are constrained by some aspects of the way our brain processes information and relates with the physical world. For example, the pitches that we can hear and produce are limited by the range of frequencies that our ears and vocal cords can detect and produce. Additionally, the way that we perceive and process musical information is influenced by the structure and organization of our brains, and the neural mechanisms that are involved in auditory perception and cognition.

As a result, the specific rules and conventions of a musical system are influenced by the limitations and constraints of our brains and our physical abilities. For example, the scales and modes that are used in a musical system may be based on the range of pitches that we can hear and produce, and the rhythms and melodies of the music may be organized in ways that are easy for us to perceive and understand. Additionally, the instruments and timbres that are used in a musical system may be chosen based on the way that our brains process and perceive different sounds.

Overall, while the specific rules and conventions of a musical system may be somewhat arbitrary, they are also influenced by the limitations and constraints of our brains and our physical abilities. As a result, the music that we create and experience is shaped by the way that our brains process and relate to the physical world.

What is the “music syntax” that our brain learns when it is exposed to a musical culture?

The "music syntax" that our brain learns when it is exposed to a musical culture is the set of rules and conventions that govern the structure and organization of the music in that culture. For example, when we are exposed to a musical culture, we learn the scales and modes that are used in the music, the rhythms and melodies that are common, and the instruments and timbres that are characteristic of the culture. Additionally, we also learn the rules for constructing melodies and harmonies, and the conventions for combining these elements to create coherent and expressive pieces of music.

The music syntax that our brain learns when it is exposed to a musical culture is similar to the linguistic syntax that we learn when we are exposed to a language. Just as linguistic syntax provides the rules and conventions for combining words and phrases to create sentences and texts, music syntax provides the rules and conventions for

combining pitches, rhythms, and other elements to create melodies, harmonies, and other musical structures.

Overall, the "music syntax" that our brain learns when it is exposed to a musical culture is the set of rules and conventions that govern the structure and organization of the music in that culture. These rules and conventions help us to make sense of the music that we hear and to create music that is coherent and expressive.

Music and emotions

Define and discuss the concept of “emotion” and its adaptive value

Hard to define and explain, no precise agreed upon definition, agreed upon characteristics.

Relatively brief, intense, rapidly changing responses to potentially important events in the external or internal environment.

Usually of social nature

Response to internal or external events

Related “synchronized” subcomponents: cognitive changes, subjective feelings, expressive behavior, action tendencies.

Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can

- give rise to **affective experiences** such as feelings of arousal (activation at different behavior levels), pleasure/displeasure, etc.
- generate **cognitive processes** (e.g. increasing attention, appraisals, labeling processes, social bonding)
- activate widespread **physiological adjustments** (e.g., increasing heart-rate, sweating, crying...)
- lead to **behavior** that is often, but not always, expressive, goal-directed, and adaptive (e.g., running away, reiterating exposure...)

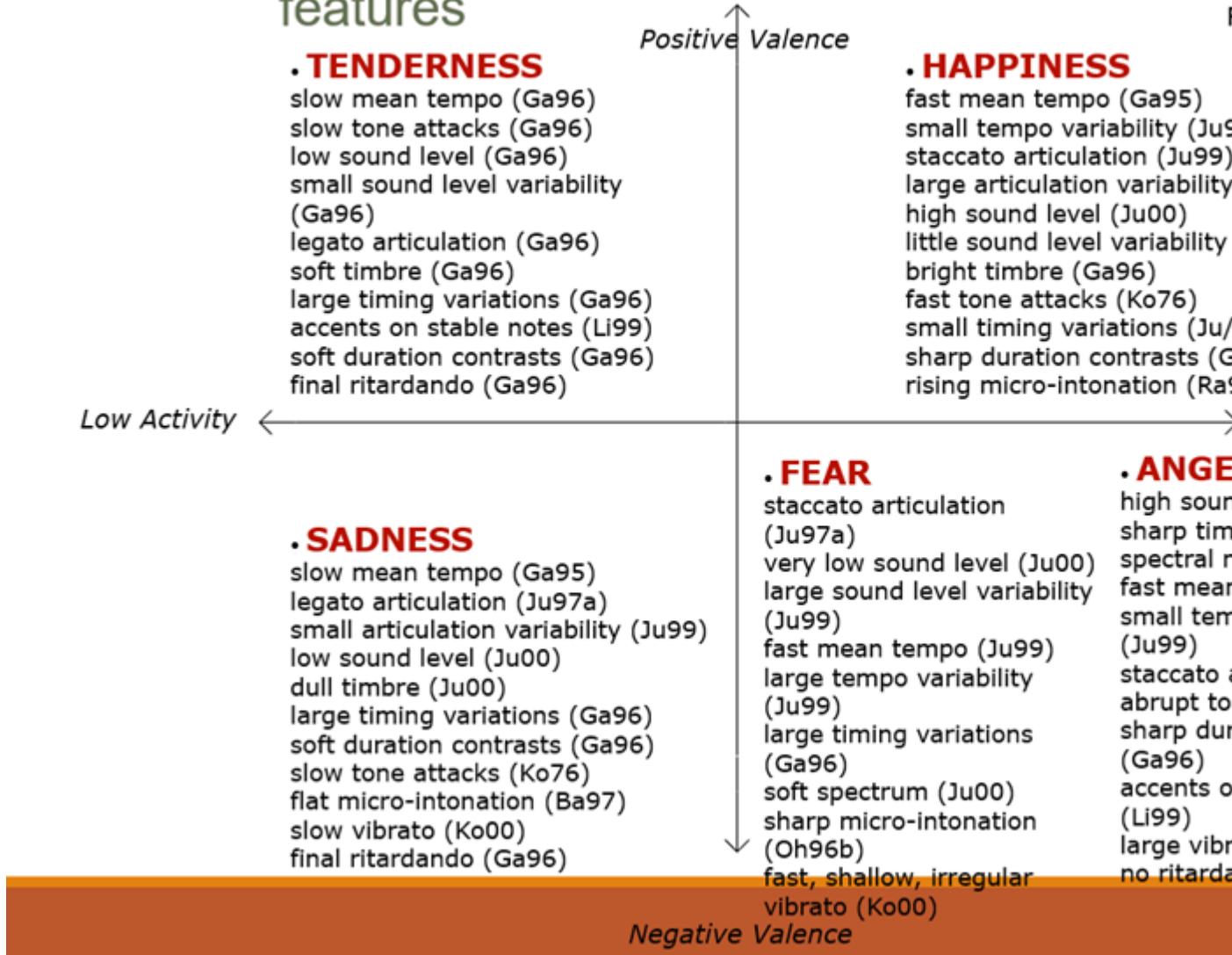
What are basic emotions? Can you identify other emotions beyond those basic? Can you decide if all of them can be communicated by music? Can all of them be induced by music?

In slide 6 it gives 8 emotions: angry, sad, joyful, ashamed, proud, elated, desperate

And I guess not all can be perfectly communicated or induced by music, at least not just with the instruments, if you add the lyrics yes, you can express any emotion.

The following emotions can be induced by music: tenderness, happiness, sadness, fear, anger.

Dimensional representation with determinants



Discuss the distinction between emotions communicated by music and emotions caused by music. Include personal, biological, and cultural factors in your discussion.

Emotions communicated by music: Depending on the type of scale that we use, or the harmony, or the rhythm, music can communicate different emotions. For example the Dorian Mode I is quite neutral, whereas the Hypodorian Mode II sounds sad, Lydian mode sounds happy.

However, we must take in consideration that these emotions have probably been associated to these modes (objective) in a subjective way, maybe running some experiments on what **Emotions are caused by music** to a group of subjects. The results of these experiments can be generalized to the whole population and create a “theory”. Nevertheless, we must keep in mind that emotions might be different among different people, or even they can change in one single person several times in the same day.

Another concept to highlight is **disassociation**: do you just listen to sad music when you are sad? Or do you listen to sad music in order to get sad? Not always, there exists a disassociation between what our emotions might be and those contained in the message a song is transmitting us.

How is it possible that music communicates about emotional content? Are there music features that can be more suitable for that?

Hearing resemblance between the music and the natural expression of the emotion (similarity to speech)

Accumulated connotations a certain musical phenomena acquire in a culture: we learn in our culture which musical cues correspond to which feeling

Features:

- Loudness and spectral dissonance found in an angry voice and in angry music
- Minor scale resembling spectra of subdued (depressed) speech
- Melodic contours may resemble questions, severe statements, etc.
- Big and fast melodic leaps analog of happy jumps
- Descending contours analog of body movements (arms, head)

So in general: tempo, major/minor, harmony, loudness, melodies, consonance/dissonance, spectral content...

How is it possible that music induces, modulates, or changes the emotional status of listeners?

Because when the listener hears a resemblance between the music and the natural expression of the emotion (similarity to speech), it associates the sound to the emotion.

For example:

- Loudness + spectral dissonance = angry voice and in angry music
- Minor scale has a spectra = depressed speech
- Melodic contours = questions
- Big and fast melodic = happy jumps
- Descending contours = body movements

Once the transmitted message is received and understood (we as listeners understand what emotion the sound is willing to express), we might get induced that emotion because of different reasonings:

- Simple activation of our system “arousal”
- Mood contagion “monkey see, money do”, “culo veo culo quiero” in Spanish
- Associations due to a memory (can be unconscious)
- Imaginery (multimodality)

Are there music features that can be more suitable for that?

Yes, there are certain music features that can be more suitable for inducing, modulating, or changing the emotional state of listeners. For example, certain melodies and harmonies can be perceived as having particular emotional qualities, such as happiness, sadness, fear, or anger. Additionally, the tempo, rhythm, and timbre of the music can also influence its emotional character, and can make the music sound more energetic, relaxed, tense, or suspenseful.

Additionally, the lyrics and other contextual information about the music can also influence its emotional character. For example, if the lyrics of a song are about a happy or sad event, this can influence the way that we perceive the emotional character of the music. Similarly, if we know that a piece of music was composed to celebrate a joyful occasion, or to mourn a tragic event, this can also influence the way that we perceive its emotional character.

Overall, there are certain music features that can be more suitable for inducing, modulating, or changing the emotional state of listeners, including the melody, harmony, rhythm, tempo, timbre, lyrics, and other contextual information about the music. By using these features in particular ways, composers and performers can create music that has a specific emotional character, and that can influence the emotional state of the listeners.

Explain the different models (dimensional versus categorical) used to study musical emotions

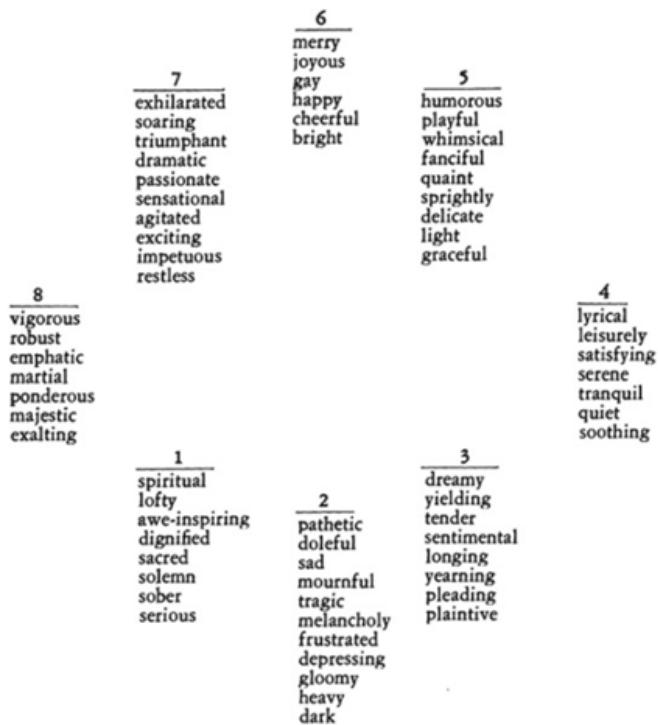
The dimensional representation is based on determining features. It is visually represented as an 2D axis of Valence(positive or negative) and Activity (high or low):

Dimensional representation with determining features

		From Ju
		↑ Positive Valence
		High A
Low Activity ←	→ High Activity	
. TENDERNESS	slow mean tempo (Ga96) slow tone attacks (Ga96) low sound level (Ga96) small sound level variability (Ga96) legato articulation (Ga96) soft timbre (Ga96) large timing variations (Ga96) accents on stable notes (Li99) soft duration contrasts (Ga96) final ritardando (Ga96)	. HAPPINESS fast mean tempo (Ga95) small tempo variability (Ju99) staccato articulation (Ju99) large articulation variability (Ju99) high sound level (Ju00) little sound level variability (Ju99) bright timbre (Ga96) fast tone attacks (Ko76) small timing variations (Ju/La00) sharp duration contrasts (Ga96) rising micro-intonation (Ra96)
. SADNESS	slow mean tempo (Ga95) legato articulation (Ju97a) small articulation variability (Ju99) low sound level (Ju00) dull timbre (Ju00) large timing variations (Ga96) soft duration contrasts (Ga96) slow tone attacks (Ko76) flat micro-intonation (Ba97) slow vibrato (Ko00) final ritardando (Ga96)	. FEAR staccato articulation (Ju97a) very low sound level (Ju00) large sound level variability (Ju99) fast mean tempo (Ju99) large tempo variability (Ju99) large timing variations (Ga96) soft spectrum (Ju00) sharp micro-intonation (Oh96b) fast, shallow, irregular vibrato (Ko00)
. ANGER		high sound level (Ga96) sharp timbre (Ju00) spectral noise (Ga96) fast mean tempo (Ju99) small tempo variability (Ju99) staccato articulation (Ju99) abrupt tone attack (Ju99) sharp duration contrasts (Ga96) accents on unstable notes (Li99) large vibrato extent (Ju99) no ritardando (Ga96)
Negative Valence		

On the other hand, Hevner's models belong to categorical models because the mood spaces consist of a set of discrete mood categories. Hevner designed an adjective circle of eight clusters of adjectives as shown in the figure, from which we can see:

- 1) the adjectives within each cluster are close in meaning;
- 2) the meanings of adjacent clusters would differ slightly; and
- 3) the difference between clusters gets larger step by step until a cluster at the opposite position is reached.



Explain the concepts of valence and arousal in musical emotions

The circumplex model of emotion suggests that emotions are distributed in a two-dimensional circular space, containing arousal and valence dimensions. Arousal represents the vertical axis and valence represents the horizontal axis, while the center of the circle represents a neutral valence and a medium level of arousal. In this model, emotional states can be represented at any level of valence and arousal, or at a neutral level of one or both of these factors. Circumplex models have been used most commonly to test stimuli of emotion words, emotional facial expressions and affective states.

Explain the role of the hippocampus, amygdala, and nucleus accumbens in the emotional processing of music

The hippocampus, amygdala, and nucleus accumbens are three brain regions that are involved in the emotional processing of music. The hippocampus is a brain region that is involved in memory and spatial navigation, and has been found to play a role in the emotional processing of music. For example, the hippocampus has been shown to be involved in the emotional impact of musical memories, and in the way that we perceive the emotional content of music.

The amygdala is a brain region that is involved in emotional processing, and has been found to play a role in the emotional response to music. The amygdala is activated by emotional stimuli, and can influence the way that we perceive and interpret the emotional content of music. For example, if we hear a piece of music that is perceived as happy or sad, the amygdala can influence the way that we feel and respond to the music.

The nucleus accumbens is a brain region that is involved in reward and pleasure, and has been found to be activated by music that we find pleasurable or rewarding. The nucleus accumbens can influence the way that we perceive the emotional content of music, and can influence our emotional response to the music. For example, if we hear a piece of music that we enjoy, the nucleus accumbens can cause us to feel pleasure and reward, and can motivate us to seek out more of the music.

Overall, the hippocampus, amygdala, and nucleus accumbens are three brain regions that are involved in the emotional processing of music. These brain regions are activated by the emotional content of music, and can influence the way that we perceive and respond to the music.

Discuss the advantages and disadvantages of different techniques used to study emotional responses to music

There are several different techniques that are used to study emotional responses to music, and each of these techniques has its own advantages and disadvantages. Some of the most common techniques include self-report measures, physiological measures, and neuroimaging.

Self-report measures are one of the most commonly used techniques for studying emotional responses to music. In this technique, participants are asked to listen to a piece of music and to report on their emotional response to the music, using a questionnaire or other rating scale. This technique is relatively easy to administer, and can provide useful information about the emotional content of the music and the participants' subjective experiences of the music. However, self-report measures can be subject to bias, and may not always accurately reflect the participants' true emotional responses to the music.

Physiological measures are another technique that is commonly used to study emotional responses to music. In this technique, participants are asked to listen to a piece of music, and their physiological responses to the music are measured using sensors or other physiological instruments. For example, heart rate, skin conductance, and facial expression can be measured to provide information about the participants' emotional responses to the music. This technique can provide more objective information about the participants' emotional responses to the music, but can be more difficult and expensive to administer than self-report measures.

Neuroimaging is a more advanced technique that is used to study emotional responses to music. In this technique, participants are asked to listen to a piece of music, and their brain activity is measured using a neuroimaging technique such as functional magnetic resonance imaging (fMRI). This technique can provide detailed information about the brain regions that are involved in emotional processing, and can help to identify the neural mechanisms that are involved in our emotional responses to music. However, neuroimaging is a more expensive and complex technique than self-report or physiological measures, and is not always practical for studying emotional responses to music.

Overall, there are several different techniques that are used to study emotional responses to music, and each of these techniques has its own advantages and disadvantages. Self-report measures are easy to administer, but can be subject to bias. Physiological measures are more objective, but can be more difficult and expensive to administer. Neuroimaging provides detailed information about the brain regions involved in emotional processing, but is more complex and expensive to administer.

How can anticipation be linked to emotional responses to music?

Anticipation can be linked to emotional responses to music by influencing the way that we perceive and interpret the emotional content of the music. When we listen to a piece of music, we often anticipate what will happen next in the music, based on our knowledge of the structure and conventions of the music. For example, if we know that a piece of music is in a particular key, or follows a particular chord progression, we may anticipate the next pitches and chords that will occur in the music.

This anticipation can influence our emotional response to the music in several ways. For example, if the music confirms our anticipation, this can create a feeling of satisfaction and pleasure, and can reinforce our emotional response to the music. On the other hand, if the music surprises us or goes against our anticipation, this can create a feeling of tension or suspense, and can change our emotional response to the music.

Additionally, anticipation can also influence our emotional response to music by influencing the way that we perceive the emotional content of the music. For example, if we anticipate that a piece of music will be happy or sad, this can influence the way that we perceive the emotional character of the music, and can affect our emotional response to the music.

Overall, anticipation can be linked to emotional responses to music by influencing the way that we perceive and interpret the emotional content of the music. By anticipating what will happen next in the music, we can influence our emotional response to the music, and can create a feeling of satisfaction, suspense, or other emotions.

Exam Question:

Use your thesis topic (or, in case not having a topic yet, take one that you like or that could be a candidate for your thesis topic; in any case you should have some basic familiarity with the

topic) to discuss how your thesis work could take advantage of knowledge, data and/or methods from music perception and cognition (or, maybe, your work could provide some additional data or knowledge to music perception and cognition). Be as much specific as possible. Your answer should include at least 600 words and can be extended up to 2 pages.

If you have doubts, or think that your topic does not have links to MP&C let me know and I'll try to help you to see them or, maybe, will pose an alternative question to you.