Perception of Music

By:
Daniel Williamson

Perception of Music

By:
Daniel Williamson

Online:
< http://cnx.org/content/col10697/1.1/ >

C O N N E X I O N S

Rice University, Houston, Texas

Collection structure revised: April 28, 2009
PDF generated: October 26, 2012
For copyright and attribution information for the modules contained in this collection, see p. 15.

Table of Contents

Chapter 1

Processing Sounds[1]

Introduction
Hearing is a very complex process involving a very long tract of highly specialized mechanisms. No matter whether the stimulus is a complex sound or a simple pure tone, extensive processing takes place both before and after the sound reaches the cerebral cortex.

For the scope of this project, the focus will be on research examining the initial coding of sounds that takes place in the spiral ganglion structure of the cochlea as well as the processing which occurs after the representation of a sound has reached the auditory cortices of the cerebral cortex. Due to the presence of extensive cross-connections along the ascending auditory pathway, the majority of research focuses on the the spiral ganglion located near the beginning of the pathway. It is very difficult to test other areas of the pathway because of the presence of these extensive cross-connections, and in order to successfully explore other segments, bilateral lesions to the pathway are necessary but rarely occur naturally.

1.1 The Ascending Auditory Pathway

Before sounds can be processed by the auditory cortices of the cerebral cortex, they must first reach the cortex. In order to do so the sound must be encoded and re-encoded several times as the trains of neuronal action potential ascend the auditory pathway. The ascending auditory pathway consists of many different structures that are involved in the processing of sounds. Figure 1.1 shows a simplified representation of the auditory pathway. The first structure involved with processing sounds is the cochlea, a structure that functions by means of detecting pressure changes associated with the acoustic wave of any given sound. The structure is filled with watery liquid that moves in response to the acoustic vibrations. The animation in Figure 1.2 demonstrates how the cochlea might respond to various forms of tonal stimuli.

[1] This content is available online at <http://cnx.org/content/m22651/1.3/>.

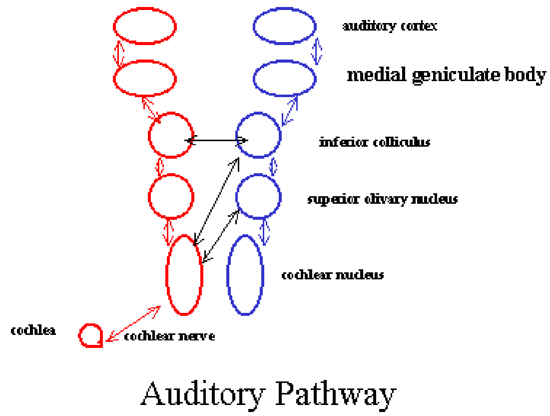

Figure 1.1: Simple representation of the Auditory Pathway.

This media object is a Flash object. Please view or download it at
<http://www.youtube.com/v/dyenMluFaUw&hl=en&fs=1>

Figure 1.2

Contained within the spiral structure of the cochlea is the spiral ganglion, a group of nerve cells that encode and transmit a representation of sounds from the cochlea to the brain(see Figure 1.3. It is estimated that between 35,000 and 50,000 neurons exist in the spiral ganglion. The axons of the spiral ganglion cells respond to information from sensory receptors in the cochlea and these trains of action potentials continue through the other structures of the ascending auditory pathway, until they reach and are processed by the auditory cortices.

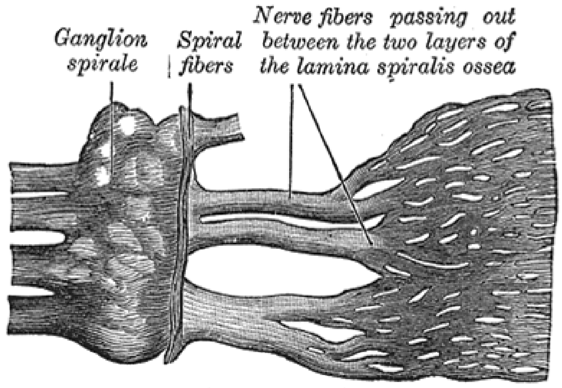

Figure 1.3

1.2 The Auditory Cortex

The auditory cortex is divided into three parts:

- Primary Auditory Cortex
- Secondary Auditory Cortex
- Tertiary Auditory Cortex

These three structures are situated concentrically to one another with the primary auditory cortex (PAC) located in the middle and the tertiary auditory cortex located on the outside. The cortex is found on the posterior superior temporal gyrus and portions of the planum temporal and Heschl's gyrus. The PAC is located in Broadmann areas 41 and 42.

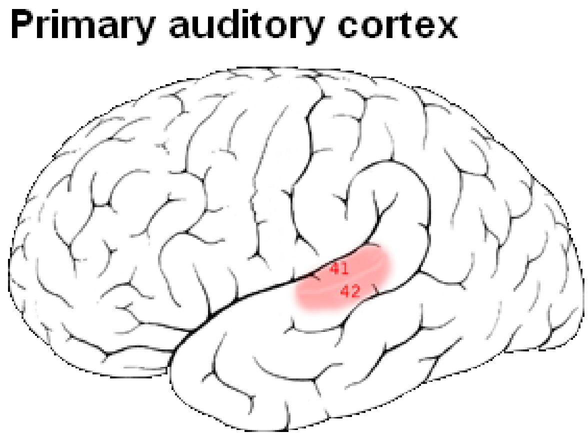

Figure 1.4

1.2.1 Tonotopic Organization

At the cortical level, perception of pitch is arranged according to a frequency spectrum. Each pitch frequency corresponds to a "best frequency"[1] site. According to Liégeois-Chauvel et. al. (2001) "for most auditory neurons, a "best frequency" (or BF) can be deteremined at which low-intensity auditory stimulus evoke the greatest electrophysiological response in a given region." This means that each neuron corresponds to a specific frequency.

The auditory cortex is arranged systematically according to these best-frequency regions. There is an ordered change in these sites with low frequencies (500 HZ) represented laterally or closer to the surface of the cortex and high frequencies represented medially or closer to the center of the brain.[1] Figure 1.5 created by Liégeois-Chauvel et al. (2001) shows a schematic representation of the tonotopic organization of the auditory cortices of both the left and right hemispheres.

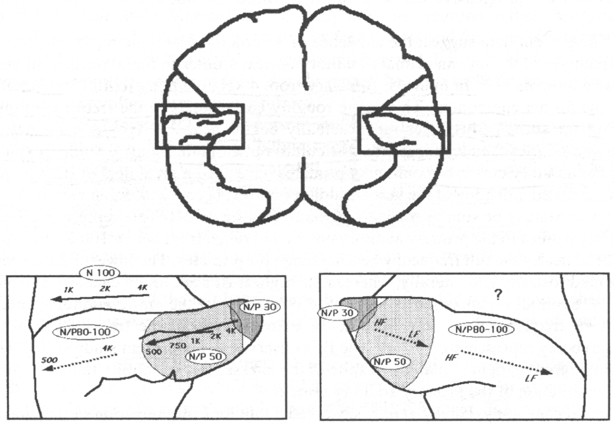

Figure 1.5

Information about the organization of the human auditory cortex has been gathered using a variety of techniques. In particular, positron emission tomography (PET) and magnetoencephalography (MEG) provided the initial clues as to how the cortex was organized.

PET Study

A PET study carried out by Lauter et al. (2001) observed an increase in regional cerebral blood flow to a deeper and more posterior location in the temporal lobe when the stimulus was a high frequency than when the stimulus was a low frequency.[1] This provided some evidence as to the location of neurons that process specific ranges of frequencies, but with PET the spatial and temporal resolution was not powerful enough to conclude anything with certainty.

MEG Study

Similarly, some MEG studies have shown that the processing of higher frequencies involved more medial regions of the Heschl's gyrus, whereas lower frequencies involve more lateral areas.[1]

1.2.1.1 Intracerebrally Recorded Auditory Evoked Potentials

At the date of the study by Liégeois-Chauvel et al. that is about to be explored in some depth, most information that had been gathered about the functional organization of the human auditory cortex "had been limited to the use of either scalp-recorded auditory evoked potentials (AEPs), which have relatively poor spatial resolving power, or functional imagery techniques, which have poor temporal resolving power."[1] The current study records intracerebrally evoked potentials in the auditory cortices of both hemispheres of the human brain. The objective of the study was to investigate the tonotopic organization of the auditory cortices in both hemispheres and examine any differentiation between the function and/or organization of the different hemispheres.

The study involved 45 adults with medically intractable partial epileptic seizures. 31 of the subjects had the origin of the seizures in the right hemisphere and the other 14 had seizures originating in the left hemisphere. In order to gather the data, electrodes were implanted orthogonally in the lateral part of the Heschl's gyrus and planum temporale. The subjects were awake and alert throughout the procedure, which involved auditory stimuli of 30-millisecond tone bursts of frequencies ranging from 250 Hertz to 4 kiloHertz with an intensity of 70 decibels.[1] The results gathered from the experiment exposed some specific differences between the organization of the different hemispheres.

The right hemisphere demonstrated clear spectrally organized tonotopic maps with distinct separations between different frequency processing regions. Also, the auditory evoked potentials (AEPs) for high frequencies were recorded medially whereas AEPs for low frequencies were recorded laterally. In the left hemisphere the tonotopic organization was less evident, and pitch specificity was less well defined. In addition the greatest electrophysiological response was observed at a range of frequencies between 600 Hertz and 2 kiloHertz.[1]

Conclusion

Based upon this evidence the researchers concluded that the auditory cortices are composed of "frequency dependent tonotopic maps."[1] Yet, these maps were concluded to be more complex and more hemisphere specific than previously thought, especially in relation to the medial versus lateral representations of different frequencies. Still, there was a conclusion that the best frequency sites are organized with high frequencies located in medial regions of the PAC and low frequencies in more lateral regions. These best frequencies "were highly stable from patient to patient in spite of intersubject variability in localization and orientation of the PAC."[1]

Perhaps the most significant finding of this study was the differentiation in response selectivity between neurons in the right versus the left hemisphere. According to this study, "neurons in the right auditory cortex were more sharply tuned to frequency than neurons in the homologous region of the left hemisphere."[1] The researchers suggest that this might be a clue indicating hemispheric specialization, with spectral or frequency related information being processed primarily in the right Heschl's gyrus.

Chapter 2

Dimensions of Music[1]

Music has a vertical and a horizontal dimension. The vertical dimension is composed of relationships of notes occurring simultaneously. By musical convention, a note is a pitch frequency in the musical scale, and "harmonic interval" refers to the distance between two notes sounded simultaneously. This is commonly referred to as harmony. Harmony is a basic element of music, and is one of the most highly developed elements in western music. Intervals are often described as being either consonant or dissonant. Consonant means simply that the interval seems stable and pleasant, while dissonant implies the opposite: unstable and unpleasant. Obviously there is some subjectivity in declaring an interval to be consonant or dissonant, but whether by convention or natural biological response, to the western ear there is a general consensus about which intervals are consonant and which are dissonant.

Intervals Commonly Considered Consonant

- unison
- octave
- perfect fourth
- perfect fifth
- major third
- minor sixth
- minor third
- major sixth

Intervals Commonly Considered Dissonant

- minor second
- tritone (augmented fourth or diminished fifth)
- major seventh
- major second
- minor sixths
- minor seventh

The horizontal dimension of music is an even more basic element, composed of relationships among a succession of notes and silent pauses. This is the temporal element of music.

According to Timothy Griffiths, there are three basic categories of musical features that are processed by the brain when listening to music: • simple acoustic features; • complex acoustic features; • semantic features. Simple acoustic features are things such as intensity (loudness), frequency (pitch) and onset, while complex acoustic features are patterns of frequencies, onsets, intensities, or a combination of any or all of the above as a function of time. Finally the semantic features are "learned associations of sound patterns and

[1]This content is available online at <http://cnx.org/content/m22649/1.4/>.

meanings."[4]. In relation to the horizontal and vertical dimensions of music, simple features would constitute the vertical dimension where the horizontal dimension would be represented by the complex acoustic features.

2.1 The Vertical Dimension Represented Physiologically

> *As long as several simple tones of a sufficiently different pitch enter the ear together the sensation due to each remains undisturbed in the ear probably because entirely different bundles of [auditory] nerve fibers are affected. But tones of the same, or of nearly the same pitch, which therefore affect the same nerve fibers, do not produce a sensation which is the sum of the two they would have separately excited, but new and peculiar phenomena arise which we term interference... and beats... Rapidly beating tones are jarring and rough... the sensible impression is also unpleasant. Consonance is a continuous, dissonance an intermittent tone sensation. The nature of dissonance is simply based on very fast beats. These are rough and annoying to the auditory nerve.*

Hermann von Helmholtz (1863 and 1885) *On the Sensations of Tone as a Physiological Basis for the Theory of Music*

Above is a quote by a German physician and physicist who made significant contributions to the our understanding of the perception of sound. The research presented below concerning the perception of consonances and dissonances confirms von Helmholtz's assumptions almost exactly.

2.1.1 A Breif Explanation of the Overtone Series

Before we continue on, there is a concept we need to understand in order to fully comprehend the research presented below. A pitch produced by a musical instrument or voice is "composed of a **fundamental** frequency and **harmonics** or **overtones**."(Doering[2]) These components of sounds establish the timbre of a particular instrument, based on the relative strength and number produced by the instrument.(Jones[3]) The reason a trumpet sounds different from a clarinet or a human voice, even when all are producing the same note, is due to the relative strength and number of harmonics (or overtones) present in the sound. Despite the presence of harmonics the listener perceives only a single tone, but the presence of these harmonics greatly affects our perception of harmonic intervals and chords as being consonant or dissonant.

2.1.2 Examining the Neurological Representation of Consonance and Dissonance

The study by Tramo et al. (2001) compares the acoustic representation of several intervals and the trains of action potentials produced by the auditory nerve fibers of a population of 100 cats hearing the same intervals. It is important to keep in mind the information presented about the ascending auditory pathway (Section 1.1: The Ascending Auditory Pathway) . Auditory nerve fibers are the central axons of the spiral ganglion cells that transmit synaptic information to the cochlear nucleus neurons in the brainstem. When an interval is sounded the nerve fibers corresponding to the frequencies present in that interval will fire. Virtually all information about sound is transmitted to the brain through trains of action potentials produced by these synapses.[6]

First the researchers examined the acoustic waveforms and their corresponding autocorrelations (the cross-correlation of a signal with itself, to find disguised repeating patterns) for four different intervals. Two of the intervals, the perfect fourth and fifth, are considered consonant, while the other two, the minor second and tritone, are considered dissonant. When looking at the acoustic representations of the consonant harmonic intervals there is a clear pattern of peaks and the autocorrelation is perfectly periodic. The period of the autocorrelation corresponds to the fundamental of bass of the interval. Yet with the dissonant intervals, there was no true periodicity. Instead the acoustic spikes occurred inconsistently.

[2]http://cnx.org/content/m15439/latest/
[3]http://cnx.org/content/m11654/latest

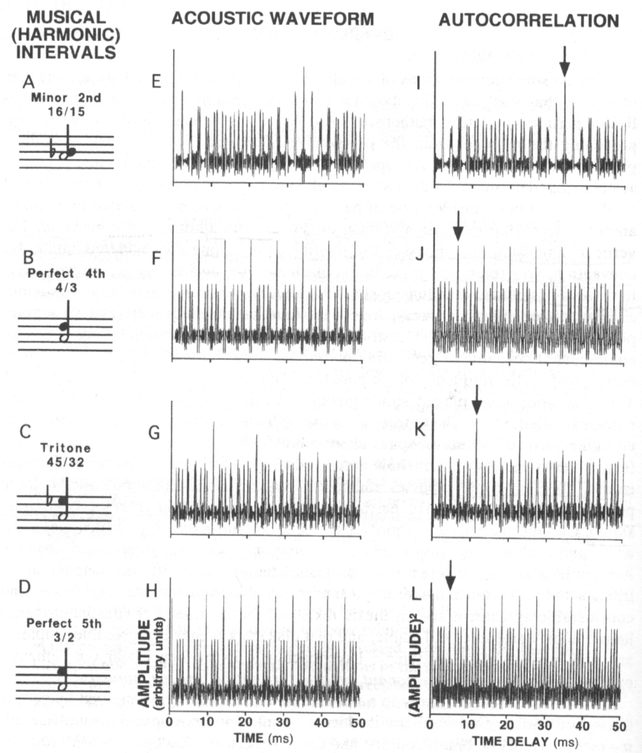

Figure 2.1: The acoustic representation of the sound waveforms and the corresponding autocorrelations for each interval. This image was created by Tramo, Cariani, Delgutte & Braida. (2001)

Using 100 cats, the experimenters measured the firings of the auditory nerve fibers on the cat's cochlear nucleus using electrodes implanted in the brainstem. Similar to the best frequencies in the section on tonotopic organization (Section 1.2.1: Tonotopic Organization) there is frequency selectivity present with the spiral ganglion cells and the nerve fiber "will only increase the number of action potentials it fires if it is sensitive to the frequencies present in the interval."[6]

Using the data collected, the researchers measured the the intervals between the spikes and then constructed an all-order histogram for the interspike intervals (ISIs). This all-order histogram is equivalent to the autocorrelation of the acoustic waveforms. In examining the histogram for the consonant intervals, it is obvious that the major peaks appear with clear periodicity corresponding to the fundamental bass. Dissonant harmonic intervals' ISIs are irregular and contain little or no representation of pitches corresponding to notes in the interval or the fundamental bass.[6] As was seen in the autocorrelation, there is again no clear periodicity of spikes. The neural response appears to mirror the acoustic representation of the intervals.

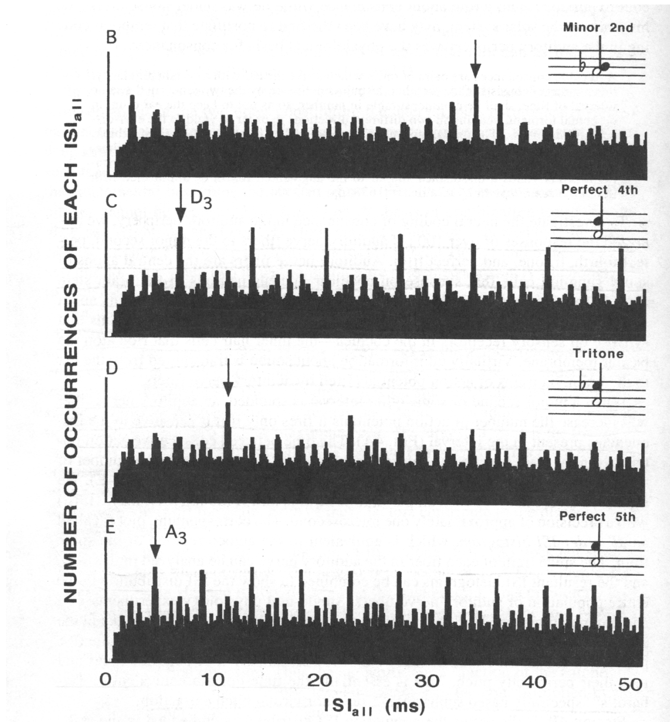

Figure 2.2: The all-order histogram produced by Liégeois-Chauvel et al. showing the the interspike intervals produced by the auditory nerve fibers of 100 cat subjects.

The reason that the consonant intervals' autocorrelations and histograms exhibit a clear periodicity and the dissonant intervals fail to demonstrate this relates back to the overtone series. Existing with each note of an interval are the inherent harmonics. The notes in consonant intervals, such as the perfect fourth and fifth, have some of the same overtones. Thus, the harmonic series of each note reinforces the other. However, the harmonics of the notes in the dissonant intervals, such as the tritone and minor second, do not reinforce each other in this way. This results in periodic amplitude fluctuations, known as beats, which make the tone combination sound rough or unpleasant. The audio example below presents a comparison of a perfect fifth followed by another fifth with the top note slightly lowered resulting in a clear example of beats.

This media object is a video file. Please view or download it at
<http://www.kylegann.com/32-4027.mp3>

Figure 2.3: A comparison of two intervals: the first is a perfect fifth that is in tune, while the second has the fifth of the interval slightly lowered resulting in clearly audible beats.

2.1.3 Lesion Study

In another study, a patient, MHS, who had bilateral lesion of the auditory cortices, was examined. In this study the experimenter presented the patient with either a major triad or a major triad with the top note slightly flattened, similar to the audio example presented above. After the chord was presented the patient was asked to indicate whether the chord was out of tune or properly tuned. The patient answered with an accuracy of 56% which is two standard deviations below the norm of 13 controls. [6]

There are two possible interpretations of MHS's results.

- MHS was having difficulty extracting the pitches of chord frequency components and analyzing their harmonic relationships
- He heard more roughness in the chords than the controls because his effective critical bandwidths were wider. (less precision of pitch perception)[6]

In addition to his inability to recognize the tuning of the chords, in another study he also demonstrated an impairment in his ability to detect the direction of pitch change between two notes. This was determined using a pitch discrimination test where the subject must identify whether the pitch is going up or down. The changes in frequency get progressively smaller. MHS "performed poorly in the final third of the test where the ΔFs [changes in frequency] were smallest."[6]

Conclusion

Drawing on the evidence presented by these studies we can conclude that representation of roughness or dissonance exists in the patterns of neural activity at both the level of the auditory nerve fibers, as seen in the trains of action potentials of the spiral ganglion cells, and in the cerebral cortex, specifically the auditory cortices and the planum temporale. In addition there is a clearly inverse relationship between the amount of temporal fluctuation in amplitude and perception of consonance, or put more simply, the more beat present in a sound, the less consonant it seems. Finally, bilateral lesions to the auditory cortices can lead to severe impairments in perception of consonance with a particular bias to perceive well-tuned chords as out of tune.

2.2 The Neurological Representation of the Horzontal Dimension of Music

As mentioned before the horizontal dimension of music is encompassed by patterns of frequencies, onsets, intensities or a combination of any or all of the above as a function of time. Several studies have been performed at the level of the cortex to determine where processing of temporal information occurs. The results of these studies lack any definite conclusions as to where temporal information is processed, with the exception of the processing of rapid temporal information.

Efron (2001) performed an experiment in which two tones of different frequencies separated by a silent pause were presented to brain damaged patients. The patients were asked to judge the temporal order of the two tones. The results showed that "aphasic patients with left-hemisphere lesions required longer intervals [of time](between 140 and 400ms) to discriminate temporal order than nonaphasic patients"[10]. The control subjects took on average only 75ms to identify the temporal order.

Similarly, Tallal and Newman examined the effects of selective damage to either the left or right hemisphere of the cortex. Their study concluded that damage to the left hemisphere disrupted the patients' ability to process two tones that were separated by a short interval of 300ms or less. Yet, damage to either side of the cortex had no effect on the patients' ability to process the two tones when longer silent intervals were used.[10]

When measuring intracerebrally evoked potentials resulting from the stimuli of syllables, Liégeois-Chauvel et al. (2001) "demonstrated a specialization of the left auditory cortex for speech perception that depends on rapid temporal coding."[10] This information supports the hypothesis that the left-hemisphere structures are predominantly involved in the processing of rapid sequential information, and thus it can be hypothesized that these same structures serve a similar function with the processing of rapid temporal information contained in music.

2.2.1 The Effect of Tempo on Irregularity Discrimination

Samson et al. (2001) conducted an experiment that confirms that the left hemisphere does in fact process rapid temporal information in music. The objective of the experiment was to "test the effect of tempo on irregularity discrimination in patients with unilateral lesions to the temporal lobe."[10] The experimenters played two successive sequences of a single tone repeated five times. One of the sequences presented contained regular silent intervals between the onsets of the tones, but the other contained irregular silent intervals. The patients were asked to judge which of the two sequences was regular or irregular. The duration of the silent intervals between the tones of the sequence were referred to as the "interonset intervals" and the various durations used were 80, 300, 500, 800, or 1000ms. This procedure was carried out on a total of 24 subjects: 8 with right hippocampal atrophy, 10 with left hippocampal atrophy, and 6 normal controls.[10]

The test confirmed the results of previous studies. At the rapid tempo, the interonset interval of 80ms, irregularity discrimination thresholds were significatly reduced for patients with damage to the left hemisphere.[10] Of course the opportunity for error was higher at the faster tempos no matter whether the patient had cerebral damage or not, but the results show a marked deficit for patients with left hemisphere damage as opposed to the other two subject categories. Again in agreement with previous studies, this data shows that damage to left hemisphere affected only the processing of fast temporal information, while slow temporal information processing was spared. In addition, damage to the right hemisphere resulted in very little difference in the ability to process rapid temporal information, and slow temporal processing was also normal.

Conclusion

From the information presented the only thing that is consistently confirmed is the role of the left hemisphere in processing of rapid temporal information. Mentioned previously, Griffiths refers to temporal information as a complex feature of music. Perhaps drawing on this idea of complexity we can hypothesize that with the exception of rapid temporal processing, sound patterns occurring as a function of time presented at a relatively slower rate require multiple structure to be successfully processed. Since the sounds are occurring in time and are heard as phrases rather than individual notes and silences, short term memory could be utilized to process the sound phrase as a unit. This could explain why only rapid temporal information, the information needing immediate processing, has a specific localization, whereas the slower phrase could utilize an extensive neuronal web, employing many different brain structures to processes specific pieces of the whole. Unfortunately, I did not come across any papers on this subject while doing my research, and certainly further research in this direction is needed to substantiate this hypothesis.

Bibliography

[1] Jean-Michel Badier Patrick Marquis Catherine Li[U+FFFD]ois-Chauvel, Kimberly Giraud and Patrick Chauvel. Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Annals of the New York Academy of Sciences*, 930:117– 132, 2001.

[2] Jean-Michel Badier Patrick Marquis Catherine Li[U+FFFD]ois-Chauvel, Kimberly Giraud and Patrick Chauvel. Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Annals of the New York Academy of Sciences*, 930:117– 132, 2001.

[3] Timothy D. Griffiths. The neural processing of complex sounds. *Annals of the New York Academy of Sciences*, 930:133–142, 2001.

[4] Timothy D. Griffiths. The neural processing of complex sounds. *Annals of the New York Academy of Sciences*, 930:133–142, 2001.

[5] Bertrand Delgutte Mark Jude Tramo, Peter A. Cariani and Louis D. Braida. Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Sciences*, 930:92–116, 2001.

[6] Bertrand Delgutte Mark Jude Tramo, Peter A. Cariani and Louis D. Braida. Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Sciences*, 930:92–116, 2001.

[7] Isabelle Peretz. Brain specialization for music: New evidence from congenital amusia. *Annals of the New York Academy of Sciences*, 930:153–165, 2001.

[8] Isabelle Peretz. Brain specialization for music: New evidence from congenital amusia. *Annals of the New York Academy of Sciences*, 930:153–165, 2001.

[9] Nathalie Ehrl[U+FFFD]verine Samson and Michel Baulac. Cerebral substrates for musical temporal processes. *Annals of the New York Academy of Sciences*, 930:166–178, 2001.

[10] Nathalie Ehrl[U+FFFD]verine Samson and Michel Baulac. Cerebral substrates for musical temporal processes. *Annals of the New York Academy of Sciences*, 930:166–178, 2001.

Index of Keywords and Terms

Keywords are listed by the section with that keyword (page numbers are in parentheses). Keywords do not necessarily appear in the text of the page. They are merely associated with that section. *Ex.* apples, § 1.1 (1) **Terms** are referenced by the page they appear on. *Ex.* apples, 1

Attributions

Collection: *Perception of Music*
Edited by: Daniel Williamson
URL: http://cnx.org/content/col10697/1.1/
License: http://creativecommons.org/licenses/by/3.0/

Module: "Processing Sounds"
By: Daniel Williamson
URL: http://cnx.org/content/m22651/1.3/
Pages: 1-6
Copyright: Daniel Williamson
License: http://creativecommons.org/licenses/by/3.0/

Module: "Dimensions of Music"
By: Daniel Williamson
URL: http://cnx.org/content/m22649/1.4/
Pages: 7-12
Copyright: Daniel Williamson
License: http://creativecommons.org/licenses/by/3.0/

Perception of Music
This collection is a term research project for the class "Linguistics 411: Neurolinguistics" at Rice University. The project focuses on current research concerning the neuronal structures and processes involved with the perception of music.