

# What We Know So Far About Music Cognition

Mladen Milicevic<sup>1</sup>

Loyola Marymount University Los Angeles, CA, USA

**Abstract.** The most recent research in neuroscience shows that the human brain “deconstructs” music patterns and then takes the elements of that “deconstruction” in order to store them in the brain’s multiple memory locations. The retrieval of memory from these multiple neural circuits is required in order to reconstruct in our mind the representation of those music patterns. The smoothness of this process is heavily dependent on repeated use of the same neural circuits for the similar, if not the same purposes. Increased smoothness also increases the emotional reward system and we have a pleasurable listening experience. This smoothness can also be defined in terms of fulfilled expectations of what might happen. One of the reasons why the human brain searches for previously identified and processed patterns is that such patterns can be very quickly reconstructed from the data stored in the brain’s long-term memory. Then that reconstruction can be compared with similar incoming information, giving it the most pragmatic interpretation that fits the situation at hand. So the function of memory is to ignore irrelevant details, while preserving the gist. This could be “good” or “bad” music but it can more importantly be related to the survival of the species in terms of “friend or foe.”

**Keywords:** Music, Brain, Cognition, Processing.

## 1. Introduction

What do we scientifically know today about how the human brain processes music? We do know much more than we did fifty years ago, but this is still just the beginning, and we have a long way to go. Let’s start with a quote from Steven Pinker:

The major faculties of the mind, with their feats no robot can duplicate, show the handiwork of selection. This does not mean that every aspect of the mind is adaptive. From low- level features like the sluggishness and noisiness of neurons, to momentous activities like art, music, religion, and dreams, we should expect to find activities of the mind that are not adaptations in the biologists’ sense. But it does mean that our understanding of how the mind works would be woefully incomplete or downright wrong unless it meshes up with our understanding of how the mind evolved [1].

As we go inside the human brain and try to figure out what is actually going on in respect to music processing, we always find out that what shaped the brain’s organization is best understood in evolutionary terms. Perhaps the best way to start may be by following what happens when any sound enters the human ear. I will not talk about the mechanics of how the sound wave energy gets transformed into a bioelectrical signal, but rather what the human brain does when the “information” gets to the auditory cortex.

## 2. Inside the Brain

### 2.1. Signal Processing

First, sound signals have to be separated by pitch. Following that, speech and music more likely have to move into separate processing circuits. The speech circuits will analyze the signal in terms of individual phonemes—the consonants and the vowels, which are the ingredients of a given language’s phonetic system. The music circuits will deconstruct the signal and independently dissect pitch, timbre, contour and rhythm. These neuronal circuits link their output to the regions of the frontal lobe, which are in the “business” of putting it all together and making sense of it. This is where the structural analysis takes place in an attempt to determine if there is any structural arrangement to the temporal patterning of the incoming information. The frontal lobe circuits will then pass the information to hippocampus processing areas in the interior temporal

---

<sup>1</sup> Mladen Milicevic Tel.: +1-310-338-4575 ; fax: +1-310-338-3030.  
*Email:* mmilicev@lmu.edu

lobe in order to examine the brain's memory storage, which may help in making sense of this auditory information. The neural circuits will try to determine if the brain has processed this (or a similar) particular sonic pattern before. This "before" may be 3-5-10 years ago, or it may be one minute ago. If it is something from a "distant past," there will be more than likely some kind of meaningful interpretation attached to it. If it is from the "near past," the brain will try to figure out if the pattern may be an element of a larger arrangement of events whose meanings are emerging right now as the brain processes these events.

Addressing the past experiences we have to also consider the way our brain handles its long-term memory. These memories are not stored in the brain photographically as intact individual events; and there are no warehouses of audiotapes, albums or pictures. This is completely unlike computer-based memory, which deals with exact reproductions. There has been significant disagreement whether human memory is relational or absolute. If our memory system stores information about relations between events and ideas, rather than all details about the events themselves, this would be considered a constructivist view. It implies that, in the absence of sensory details, the human brain will construct a representation of reality based on the available information. The constructivists would argue that in order to make sense from the vast sonic events that enter its auditory cortex, the brain has to become a master of simplification. This process is unlike filtering unwanted information, because such a mechanism would be tremendously complicated and utterly inefficient. In actuality, the brain searches for familiar devices and patterns [2]. It latches on to things that are in some respect already known, disregarding most of the unfamiliar information. The reason that the human (but also animal) brain is doing this lies in the fact that previously processed and already interpreted information can be very quickly reconstructed from the data stored in the brain's long-term memory. Then that reconstruction can be efficiently processed and compared with similar incoming information, giving it the most pragmatic interpretation that fits the situation at hand. So the function of memory is to ignore irrelevant details, while preserving the gist. On the other hand the record-keeping theory argues that memory is like a tape recorder which preserves all or most of human experiences accurately, and with almost absolute fidelity. This is very unlikely because of the inherent speed-inefficiency of memory retrieval if this would be the case, as I mentioned earlier.

However, the third, hybrid approach of the previously mentioned theories is a multiple-trace memory model, which seems to make the most sense by proposing that every experience is potentially encoded in the brain's memory. It is not encoded in a single particular place in the brain, (record-keeping theory) but rather scattered in groups of neurons each holding specific part of the information about a particular memory. For example the information about the structure of the organization of sonic events may be stored in one place, while the information about the timbre may be stored at another place, then the information about under what emotional circumstances these sonic events were originally "stored" may yet be stored in another place. Thus, when these neuronal groups get tuned to specific values and configured in a precise way, this may cause memory to be recovered and repeated in the consciousness of our minds. The main obstacle to recalling every previous sonic experience (and for that matter any experience) does not lie in the fact that it was not memorized, but rather in the fact that it needs a very specific "trigger" to activate the memory retrieval process. So, some musical "triggers" may have nothing to do with music but may be rather triggering memory storage retrieval by the "circumstances" related to the current "state of mind" of the listener.

According to the multiple-trace memory model, the human brain operates with a reconstructed version of the original an interpretation. In order to compile a musical tune, the brain has to fire a certain set of neural mappings as the means to recreate "the music." These firing patterns trigger the momentary reconstruction of an approximate representation of the "Star Spangled Banner," for example. Your interpretation of music today, depends on: who you are, what are you doing at that moment, as well as your past experiences stored in your long-term memory. But, soon you are going to change, what you will be doing is going to change, and your past experiences in long-term memory will be altered as well. The "mass of soothing sound" your mother made while singing lullabies to you in childhood, gets reduced to Twinkle, Twinkle Little Star, later on in your life. Our memory of a certain musical piece is influenced not only by previous knowledge but also by events that happen between the time an event is perceived and the time it is recalled [3]. Furthermore, we can only recall memories that are related to our present situation where you are and what are you doing. If

you are composing an orchestral piece your brain more likely focuses on recalling memories related to the instrumental ranges, rather than memories of how to change a flat tire on your car.

So our memories, as exact, recorded, fixed images of the past, are an illusion. We believe we are stable, but this is one of the built-in illusions of the mental system. We believe we remember specific events, surely. Yet we don't. We make them up on the fly. We change our minds all the time, from our estimate of the odds on a bet, to how we view our future. And we are unaware that the mind is doing this [4].

## 2.2. Musical Expectations

One of the main principles of being intelligent is to be able to have expectations and anticipate what may happen in the future. So, when listening to a piece of music there are certain expectations, which the human brain has. Violating these musical expectations would be a formation of a musical event, which is in conflict with what might be logically anticipated. This works along the same lines as when the human brain operates with everyday situations extrapolating common elements to multiple situations and constructing a framework within which to operate with them; this framework is known as schema. This information may allow some aspects of a future signal to be anticipated as it happens when we hear the first line of a familiar song. The brain's ability to extrapolate forwards on the basis of past experience is one form of that ability that we call "intelligence"; it can dramatically enhance an organism's chances of survival.

We constantly judge by comparison, and our judgment of any item depends upon what we are comparing it to at that moment [5].

Our musical expectations for the Western music, among other things, are based on our knowledge of the musical scales that are commonly used. This is one of the reasons why non-Western music sounds strange to us. Looking around the world we will find a preponderance of asymmetric scales, which by necessity provide a sense of tonal center. Now, we may wonder why should this be? We can speculate that having a tonal center creates a cognitive reference point for human perception of pitch, making it easier to process and memorize complex melodic events.

The brain has evolved to deal with expectations and their fulfillment, thus it is impossible to "turn off" the human tendency to anticipate. The purpose of expectation in a species is to predict events in the environment, and if those predictions are correct that can have a strong survival importance. Hence, presuming that the brain itself provides reward mechanisms for accurate predictions is highly plausible. One might well imagine that expectational failures would engender stress, whereas expectational successes would engender some feeling of satisfaction or enjoyment [6]. So, we can conclude that our engagement with the "expectation-fulfillment game" when it comes to music, will translate into our interest in it. If the music is highly predictable, listeners will inevitably get bored. However, this does not apply to meditative and trance-inducing type of musical stimuli, which actually function on the very basis of high predictability.

Even though repetitive and highly predictable sounds induce boredom and may have a desensitizing effect; those kinds of environments that carry no "new" information are typically considered safe for habituation. This is an evolutionary feature of allowing humans to reduce arousal levels by ignoring predictable and redundant environmental stimuli while preserving energy for important purposes such as sexual reproduction and feeding. It is understandable that organisms do not habituate when exposed to repetitive painful stimulations. For example Chinese water torture, or repeated loud sonic events will evoke frustration rather than a feeling of homeostasis and existential comfort. Therefore, when music is highly unpredictable we may feel stressed by such an experience. So, the question is how to reconcile the extremes between the musical experiences of stress and experiences of boredom? If positive experiences come from predictability, then the most enjoyable music should be extremely trite? It is fair to assume that when we go to listen to music in a concert situation the main reason is to get pleasantly excited. This means that preparation will be needed in anticipation of incoming musical stimuli. Entering a concert hall is like an "expectation-fulfillment game" where we want to be in an exciting zone between the stress and boredom.

## 2.3. Music Schemas

Through the exposure to music our brain creates cognitive schemes for musical genres and forms. This happens even with passive listening with no attempt to analyze the music. Very early in childhood, humans

establish what are the “proper ingredients” of music in their culture. We can observe that our musical taste (a cognitive scheme) is largely formed early in life where the music listening has the most profound influence. This does not mean that exposure to music of different cultures later on in life cannot make us acculturated to them, accepting new musical schemas as well. It all comes down to how much (how many times) the neural networks, which represent a certain aspect of a particular musical scheme, get recalled and refreshed in an attempt to reconstruct music memory from the “storage bin.” This is why early impressions of the “proper ingredients” of a particular cultures’ musical scheme get “engraved” strongly. It is a simple fact of how many times the similar neurons and synapses in the brain get to “fire”. Like anything else in life, the more repetition the longer it will stick.

So repeated music patterns and their repeated storage and retrieval inside the human brain create a particular cultural music scheme. I would like to emphasize two words here: pattern and repetition. No human culture can exist without the repetition. Thus in order to repeat something, that something has to be organized in a recognizable structure. When we talk about organizing musical events the most important elements of organization are melody and the timbre. Instrumentation, tempo, pitch and loudness may be considered quite irrelevant from a musical pattern recognition standpoint. Music is interesting to us because we can recall the tones we have just heard and can then associate the tones that we are hearing right now. The groups of tone patterns (phrases) generally come up repeated and varied later on in the musical piece. If this pattern variation becomes intriguing enough it will “positively” activate our emotional centers, giving us a sense of reward when our expectations get somewhat fulfilled.

Neuroimaging has also shown that our memory system is closely related to our emotional system. The amygdala, which handles emotions, is located next to the hippocampus, which is a crucial structure for memory storage, and perhaps for memory retrieval. The amygdala is highly activated by any experiences or memories with a strong emotional component. It becomes active to music patterns, but not to a random collection of sounds or musical tones. If listening to music becomes an emotionally satisfying and pleasurable experience (a reward); humans naturally seek a repetition of it.

## **2.4. Timbre and Rhythm**

From an aesthetic standpoint, timbre is arguably as important as melody as a perceptual feature of music. (Imagine an expressive Paganini violin piece played on a real violin and a cheap Casio synthesizer.) [7] From a cognitive standpoint however, timbre differs sharply from melody in that the former is rarely the basis for organized sound contrasts produced by the individual instruments. However, significantly based on timbre, humans can recognize and differentiate among a myriad of diverse speaking voices. In most contexts rhythm denotes some kind of periodicity like a pattern repeating regularly in time. However, periodicity is only one way of organizing rhythmical structure, which may be also organized around timing, accent and grouping. Thus, cognitively when it comes to recognition, changes in rhythm are much less destructive than the changes in melody. The importance of rhythm lies in the fact that it is closely related to coordination of a synchronized movement like tapping a foot and dancing. Humans are the only species who can link their movement in synchronicity with the beat. This is not surprising since the motor cortex and auditory cortex lie close to each other in the brain and probably share some common “circuitry”. The phenomenon of a regular beat is widespread in musical cultures across the world. Thus, it comes as no surprise to observe a very young child, less than one year old, trying to dance to a regular beat.

## **2.5. Music Origins**

We can also learn a great deal by comparing how the human brain processes music and language. It is inevitable to conclude that music and language share some mutual neural regions, but they both have autonomous circuits as well. From the evolutionary point of view, looking at the close proximity of these areas, we can conclude that regions of the brain which process music and language probably begin undifferentiated and later on gain their specialization in one or the other. For example, when it comes to structural processing—musical syntax—it is localized in regions adjacent to those, which process speech syntax, such as Broca’s area of the frontal lobes in both hemispheres. Research shows that professional musicians, when listening to music, engage their left side of the brain much more than their right side. However, when dealing with basic musical syntax, both hemispheres get engaged, regardless of whether listeners have musical training or not. Furthermore, the regions which process musical semantics—

correlating “the music” with meaning—are located on both sides of the back portions of the temporal lobe, close to Wernicke’s area. The stage for language development is best described here:

Just as fine-tuned hearing evolved from chewing in the reptilian jawbone structure (an “excaption” in the jargon of evolutionists)—as bones selected for biting became co-opted in the small bones of the ear—so human language grew from prelinguistic structures and capacities, building upon traits selected for other reasons. The jump to speech was therefore mediated, not abrupt [8].

Steven Pinker claims that “music is auditory cheesecake, and exquisite confection crafted to tickle the sensitive spots” [9] alluding that the main reason why humans like cheesecake is because it contains fat and sugar which were the nutritional essentials throughout the evolutionary history. So, the early humans developed reward centers in their brain every time they got hold of these precious survival resources such as fat and sugar. For the most economically developed nations, fat and sugar are not any more precious, thus eating too much cheesecake may be considered rather unhealthy. However, the reward mechanisms built over the thousands of years of human evolution still “make us feel good” while eating cheesecake in spite of its potential negative effects. Thus, the obesity in America persists.

Similarly, Antonio Damasio speculates that music dance, painting, and sculpting emerged first, then later on came the ability to improve communication and organize social life. These were two strong influences, which gave the arts additional power of sustainability. He believes that socio-cultural homeostasis of the human species has been shaped by collective workings of many human minds which were able to reshape their own environment in such a way that the human genome got modified to accommodate that human-made change. For example, the farming of milk producing animals modified our genes and made us lactose tolerant [10].

### 3. Conclusion

We can say that music is the intended organization of sounds for particular uses in social and cultural contexts, then the most logical question to ask would be for what uses and what contexts? It is essential to look into the elements that are essential for the understanding of music, such as the existence of musical patterns and their repetition, and the usage of pitch (melody) in creating expectations as understood within a specific cultural music scheme. Recognizing how to deal with musical emotions, which, as stated before, play a very important role in our “understanding” and reaction to music, becomes one of the key elements.

The multiple reinforcing cues of a good song—rhythm, melody, contour—cause music to stick in our heads. That is the reason that many ancient myths, epics, and even the Old Testament were set to music in preparation for being passed down by oral tradition across the generations. As a tool for activation of specific thoughts, music is not as goo’s as language. As a tool for arousing feelings and emotions, music is better than language. The combination of the two— as best exemplified in a love song—is the best courtship of all [11].

### 4. References

- [1] S. Pinker, *How the Mind Works*. W. W. Norton & Company, 2009
- [2] R. Jourdain, *Music, the Brain, and Ecstasy*. William Morrow and Company, Inc., New York, 1997.
- [3] R. Ornstein, *The Evolution of Consciousness*. New York: Prentice Hall Press, 1991.
- [4] R. Ornstein, *The Evolution of Consciousness*. New York: Prentice Hall Press, 1991.
- [5] R. Ornstein, *Multimind*. Boston: Houghton Mifflin Company, 1986.
- [6] D. Huron, *Sweet Anticipation: Music and the Psychology of Expectation*. The MIT Press, 2006
- [7] A. D. Patel, *Music, Language, and the Brain*. Oxford University Press, USA; 2010
- [8] V. S. Ramachandran, *The Tell-Tale Brain: A Neuroscientist’s Quest for What Makes Us Human*. W. W. Norton & Company, 2011
- [9] S. Pinker, *How the Mind Works*. W. W. Norton & Company, 2009
- [10] A. R. Damasio, *Self Comes to Mind: Constructing the Conscious Brain*. Pantheon; 2010
- [11] D. J. Levitin, *This Is Your Brain on Music: The Science of a Human Obsession*. Plume/Penguin, 2007