The Role of Music in Evolution of the Mind and Consciousness

Leonid Perlovsky
Harvard University

Abstract

Based on cognitive science and mathematical models of the mind, the paper describes a fundamental role music plays in the functioning and evolution of the mind, consciousness, and cultures. The paper considers a split in the vocalizations of proto-humans into two types: one less emotional and more concretely-semantic, evolving into language, and the other preserving emotional connections along with semantic ambiguity, evolving into music. The proposed hypothesis departs from other authors in considering specific mechanisms of the mind-brain, which required evolution of music parallel with evolution of cultures and languages. I argue that evolution of language toward the semantically powerful tool of today required emancipation from emotional encumbrances. I discuss opposite, no less powerful mechanisms that required a compensatory evolution of music toward more differentiated and refined emotionality. The need for refined music in the process of cultural evolution is grounded in fundamental mechanisms of the mind. This is why today's human mind and cultures cannot exist without today's music. The presented hypothesis gives a basis for future analysis of why different evolutionary paths of languages were paralleled by different evolutionary paths of music. I suggest experimental verifications of this hypothesis in psychological and neuroimaging research.

"A poet's duty is to try to mend
The edges split between the soul and body
The talent's needle. And only voice is thread."
Joseph Brodsky

An Unsolved Mystery

Music is a mystery; according to Darwin (1871), it "must be ranked amongst the most mysterious with which (man) is endowed." "Music is a human cultural universal that serves no obvious adaptive purpose, making its evolution a puzzle for evolutionary biologists" (Masataka, 2008). Kant (1790), who so brilliantly explained the epistemology of the beautiful and the sublime, could not explain music: "(As for) the expansion of the faculties… in the judgment for cognition, music will have the lowest place among (the beautiful arts)… because it merely plays with senses". Pinker (1997) follows Kant, suggesting that music is an "auditory cheesecake," a byproduct of natural selection that just happened to "tickle the sensitive spots." In 2008, Nature published a series of essays on Music; authors agreed that music is a cross-cultural universal, still "none… has yet been able to answer the fundamental question: why does music have such power over us?" (Editorial, 2008). "We might start by accepting that it is fruitless to try to define 'music'." (Ball, 2008). These are just a sampling of quotes from accomplished scientists.

In this paper we use arguments based on cognitive science and mathematical models of the mind to suggest that music serves a most important and concrete function in evolution of the mind and cultures. We elucidate this function and suggest experimental verification of this hypothesis.

Theories of Music Origins

The origins of music have challenged philosophical thought for thousands of years. Aristotle listed the power of music among the unsolved problems (Aristotle, IV BCE/1995, p.1434). During the last two decades, the powers of music that previously seemed mysterious are receiving scientific foundations due to the research of scientists in several fields. Integration of this research in recent years provides evidence for the evolutionary origins and roles of music. This section attempts to provide a glimpse into contemporary research on origins of music by summarizing views of select authors.

Justus and Hustler (2003) and McDermott and Houser (2003) review evidence for evolutionary origins of music. They emphasize that an unambiguous identification of genetic evolution as a source of music origins requires innateness, domain specificity for music, and uniqueness to humans (since no other animals make music in the sense humans do). The conclusions of both reviews are similar, i.e., "humans have an innate drive to make and

enjoy music." There is much suggestive evidence supporting a biological predisposition for music. Certain basic abilities for music are guided by innate constraints.

Still, it is unclear that these constraints are uniquely human since they "show parallels in other domains." It is likely that many musical abilities are not adaptations for music, but are based on more general-purpose mechanisms. There are "some intriguing clues about innate perceptual biases related to music, but probably not enough to seriously constrain evolutionary hypothesis." "Available evidence suggests that the innate constraints in music are not specific to that domain, making it unclear, which domain(s) provided the relevant selection pressures." "There is no compelling reason to argue categorically that music is a cognitive domain that has been shaped by natural selection." In Nature's series of essays on music McDermott (2008) writes: "Music is universal, a significant feature of every known culture, and yet does not serve an obvious, uncontroversial function".

In commentaries to these reviews published a year later, Trainor (2004) argues that for higher cognitive functions, such as music, it is difficult to differentiate between adaptation and exaptation (structures originally evolved for other purposes and used today for music), since most such functions involve both "genes and experience." Therefore the verdict on whether music is an evolutionary adaptation should be decided based on advantages for survival. Fitch (2004) comments that biological and cultural aspects in music are hopelessly entangled, and "the greatest value of an evolutionary perspective may be to provide a theoretical framework." Livingstone and Thompson (2004) emphasize a multimodal nature of musical experience engaging affect and exploring theories based on exaptations of "an earlier system of affective communication." It is therefore interesting, they suggest, exploring correlations between musicality and emotional intelligence. They emphasize human symbolic ability leading to art, including music and our capacity for "symbolic hierarchical systems."

Before reviewing other select authors, we would comment that the hypothesis in this paper corresponds to many of the suggestions and ideas in this section. We discuss a fundamental function of music in the evolution of language, mind, and culture, which has not been previously discussed and which may provide additional directions to search for evolutionary mechanisms of music. In this way the paper relates to biological roots of music, to its origins in "an earlier system of affective communication," it bears on discussions of evolution vs. exaptation, and human symbolic ability. Although we do not discuss musicality vs. emotional intelligence, this idea is conceptually close to the hypothesis proposed in this paper.

Huron (1999) emphasizes that in the search for evolutionary origins of music it is necessary to look for complex multistage adaptations, built on prior adaptations, which might have evolved for several reasons. He discusses social reasons for music origins and lists several possible evolutionary advantages of music: mate selection, social cohesion, the coordination of group work,

auditory development, developing auditory skills, refined motor coordination, conflict reduction, preserving stories of tribal origins. However, the list of possible uses of music by itself does not explain musical power over human psyche; does not explain why music and not some other, nonmusical activities have been used for these purposes.

Cross (2008a, 2008b, 2008c) concentrates on evolutionary arguments specific to music. He integrates neuroscientific, cognitive, and ethnomusicological evidence and emphasizes that it is inadequate to consider music as "patterns of sounds" used by individuals for hedonic purposes. Music should be considered in the context of its uses in pre-cultural societies for social structuring, forming bonds, and group identities. A strong argument for evolutionary origins of music is its universality; music exists in all scientifically documented societies around the globe. Cross emphasizes that music possesses common attributes across cultures: it exploits the human capacity to entrain to social stimuli. He argues that music is necessary for the very development of culture. Cultural evolution is based on ability to create and perceive socio-intentional aspect of meaning. This is unique to human and it is created by music. Cross presents a three-dimension account of meaning in music, combining "biologically generic, humanly specific, and culturally enactive dimensions." Thus evolution of music was based on already existing in animal world biological and genetic mechanisms.

The capacity for culture (Cross, 2008b) requires transmission of information, but also the context of communication. Therefore "music and language constitute complementary components of the human communicative toolkit." The power of language is in "its ability to present semantically decomposable propositions." Language, because of its concreteness, on one hand enabled exchange of specific and complicated knowledge, but on the other hand could exacerbate oppositions between individual goals and transform an uncertain encounter into a conflict.

Music is a communicative tool with opposite properties. It is semantic, but in a different way than language. Music is directed at increasing a sense of 'shared intentionality.' Music major role is social, it serves as an 'honest signal' (that is it "reveals qualities of a signaler to a receiver") with nonspecific goals. This property of music, "indeterminacy of meaning or floating intentionality," allows for individual interactions while maintaining different "goals and meanings" that may conflict. Thus music "promotes the alignment of participants' sense of goals." Therefore Cross hypothesized that successful living in societies promoted evolution of such communication system.

Cross suggests that music evolved together with language rather than as its precursor. Evolution of language required a re-wiring of neural control over the vocal tract, and this control had to become more voluntary for language. At the same time a less voluntary control, originating in ancient emotional brain regions, had to be maintained for music to continue playing the role of 'honest signal.' Related differences in neural controls over the vocal tract

between primates and humans were reviewed in Perlovsky (2005, 2006b, 2006e, 2007).

As juvenile periods in hominid lineages lengthened (altricialization), music took a more important role in social life (Cross, 2008c). The reason is that juvenile animals, especially social primates, engage in play, which prepares them to adult lives. Play involves musical-like features, thus proto-musical activity has ancient genetic roots. Lengthening of juvenile periods was identified as possibly a fundamental for proto-musical activity and for origin of music. Infant directed speech (IDS) has special musical (or proto-musical) qualities that are universal around the globe. This research was overviewed in Trehub (2003). She has demonstrated that IDS exhibit many similar features across different cultures. Young infants are sensitive to musical structures in human voice. Several researchers related this sensitivity to the "coregulation of affect by parent and child" (Dissanayake, 2000), and consider IDS an important evolutionary mechanism of music origin.

Dissanayake (2008) considers music primarily as a behavioral and motivational capacity. Naturally evolving processes led to ritualization of music through formalization, repetition, exaggeration, and elaboration. Ritualization led to arousal and emotion shaping. This occurred naturally in IDS, in process of mother-infant interaction, which in addition to specially altered voice involved exaggerated facial expressions and body movements in intimate one-to-one interaction. Infants 8 weeks old already are sensitive to this type of behavior, which reinforces emotional bonding. This type of behavior and infant sensitivity to it are universal throughout societies, which suggests evolved inborn predisposition. Dissanayake further emphasizes that such proto-musical behavior has served as a basis for cultural-specific inventions of ritual ceremonies for uniting groups as they united mother-infant pairs. The origins of music, she emphasizes, are multi-modal, involving aural, visual, and kinesic activity, which has occurred in social rather than solitary settings. She describes structural and functional resemblances between mother-infant interactions, ceremonial rituals, and adult courtship, and relates these to properties of music. All these, she proposes, suggest an evolved "amodal neural propensity in human species to respond—cognitively and emotionally—to dynamic temporal patterns produced by other humans in context of affiliation."

This combination of related adaptations was biologically motivated by co-occurrence of bipedalism, expanding brain size, and altricialization (Cross, 2008c; Dissanayake, 2008) and was fundamental to human survival. This is why, according to Dissanayake, proto-musical behavior produces so strong emotions, and activates brain areas involved in ancient mechanisms of reward and motivation, same areas that are involved in satisfaction of most powerful instincts of hunger and sex.

Mithen (2007) presents an impressive array of evidence that Neanderthals possibly have proto-musical ability. He argues that music and language have evolved by differentiation of early proto-human voice sounds "Hmmmm"

undifferentiated proto-music-language. The development was facilitated by vertical posture and walking, which required sophisticated sensorimotor control, a sense of rhythm, and possibly ability for dancing.

The differentiation of Hmmmm, he dates to after 50,000 BP. Further evolution toward music occurred for religious purposes, which he identifies with supernatural beings. Currently music is not needed, it has been replaced by language, it only exists as inertia, as a difficult to get rid off remnant of the primordial Hmmmm. An exception could be religious practice, where music is needed since we do not know how to communicate with gods.

Mithen explains why music is often perceived as a conversation, and why we feel it as having a meaning, both of these are remnants of Hmmmm. Onomatopoeia is also a survival of Hmmmm. Among a number of properties of music explained by Mithen, I would emphasize relation of music to emotions, this was present in original Hmmmm. Songs recombine language and music into original Hmmmm, however Mithen gives no fundamental reason or need for this recombination.

Mithen summarizes the state of knowledge about vocalization by apes and monkeys. Unlike older views, calls could be deliberate, however their emotional-behavioral meanings are probably not differentiated; this is why primates cannot use vocalization separately from emotional-behavioral situations (and therefore cannot develop language - LP), this area is still poorly understood. While addressing language in details, Mithen (and other scientists as well) give no explanation for why human learn language by about age of five, but the corresponding mastery of cognition takes the rest of lifetime; steps toward explaining this are taken in Perlovsky (2006c, 2006d) and in this paper.

Mithen's view on religion contradicts (Jaynes, 1976) documented evidence for relatively late proliferation of supernatural beings in religious practice, and to mathematical and cognitive explanations for the role of religiously sublime in workings of the mind (Perlovsky, 2001; Levine & Perlovsky, 2008).

Juslin and Västfjäll (2008) analyze mechanisms of musical emotions. They emphasize that in the multiplicity of papers considering music and emotions, the very use of the word 'emotion' is not well defined. They discuss a number of neural mechanisms involved with emotions and different meanings implied for the word 'emotion'. I would mention here just two of these. First, consider the so called basic emotions, which are most often discussed; we have specific words for these emotions: fear, sexual-love, jealousy, thirst… Mechanisms of these emotions are related to satisfaction or dissatisfaction of basic instinctual bodily needs such as survival, procreation, a need for water balance in the body… An ability of music to express basic emotions unambiguously is a separate field of study. Second, consider the complex or 'musical' emotions (sometimes called 'continuous'), which we 'hear' in music and for which we do not necessarily have special words. Mechanisms and role of these emotions in the mind and cultural evolution are subjects of this paper.

Levitin (2008) classified music in six different types, fulfilling six

fundamental needs, and (as far as I understood him) eliciting six basic emotions. He suggests that music has originated from animal cries and it functions today essentially in the same way, communicating emotions. An ability to communicate emotions with voice and to correctly perceive emotions in voice has given and continues to give evolutionary advantage and is the basis for emotional intelligence. Emotions motivate us to act and neural connections facilitating this are bidirectional, action and movement may elicit emotions: "emotions and motivation are two sides of the same evolutionary coin." It is more difficult, he writes, "to fake sincerity in music than in spoken language." The reason that music evolved this way as an 'honest signal' because it "simply" co-evolved with brains "precisely to preserve this property." (Given the fact that even as simple animals as birds can fake their cries (Lorenz, 1981) I have my doubts about this "simply;" further doubts arise as soon as we think about actors, singers, and poets, not only contemporary professionals, but also those existing in traditional societies (Meyer, Palmer, & Mazo, 1998) since time immemorial.)

This paper discusses mechanisms of music evolution from differentiation of original proto-music-language to its contemporary refined states. Discussions of mechanisms that evolved music from IDS to Bach and Beatles in previously proposed theories are lacking or unconvincing. Why do we need the virtual infinity of "musical emotions" that we hear in music (e.g. in classical Western music)? Is it an aberration or do they address potentially universal human needs? Dissanayake (2008) suggests that this path went through ceremonial ritualization, due to "a basic motivation to achieve some level of control over events…" If "for five or even ten centuries… music has been emancipated from its two-million year history and its adaptive roots says more about the recency and aberrance of modernity…" Cross (2008c) argues against this conclusion: "…it would be impossible to remove music without removing many of the abilities of social cognition that are fundamental to being human." He concludes that "there are further facets to the evolutionary story (of the origins of music) requiring consideration. Investigation of the origins, emergence and nature of musical behaviors in humans is in its early stages, and has plenty more to reveal." This paper addresses a novel hypothesis that might clarify some of these remaining "further facets," and provides bases for further research in several directions.

Fundamental Mechanisms of the Mind

This section summarizes fundamental mechanisms of the mind: concepts, instincts, emotions, and behavior, which will serve as a first step toward more complicated mechanisms essential for understanding the role of music. The content of this section summarizes neuro-cognitive and mathematical arguments considered, in detail, in Perlovsky (2000, 2006d) and in references therein.

The most accessible to our consciousness is a mechanism of the mind,

which operates with concepts. The mind understands the world in terms of concepts. *Concepts* serve as internal models of objects and situations. This analogy is quite literal, e.g., during visual perception of an object, a concept-model in the mind (memory, representation) projects an image onto the visual cortex, which is matched there to an image projected from the retina (this simplified description is discussed in more details in the above references; experimental neuro-imaging proof of this mechanism with detailed description of the brain regions involved is given in (Bar, Kassam, Ghuman, Boshyan, Schmid, Dale, et al, 2006). Perception occurs when the two images are successfully matched.

The "mechanism of concepts" evolved for instinct satisfaction. Instincts are mechanisms of survival that are much more ancient than mechanisms of concepts. Psychological literature actively discusses mechanisms of instincts and these discussions can be followed in the given references. Here we follow these references in considering the mechanism of *instincts* as similar to internal sensors that measure vital organism parameters, important for normal functioning and survival. For example, a low sugar level in blood indicates an instinctual need for food. This sensor measurement and the requirement to maintain it within certain limits we consider to be an "instinct." The remaining function of satisfying this instinct we consider the appropriate level of analysis in this paper. (Biologists and neuro-cognitive scientists may consider these mechanisms in much more detail, however our level of analysis is determined by our aim: to understand the fundamental mechanisms of music).

Emotions designate a number of various mechanisms which are surveyed, for example, in Juslin, & Västfjäll (2008). Here we consider emotions as neural signals connecting instinctual and conceptual brain regions. Emotions (or emotional neural signals) communicate instinctual needs to conceptual recognition-understanding mechanisms of the brain, so that concept-models corresponding to objects or situations that can potentially satisfy instinctual needs receive preferential attention and processing resources in the brain (Grossberg, & Levine, 1987; Perlovsky 2000, 2006d). Thus emotional signals evaluate concepts for the purpose of instinct satisfaction. This evaluation is not according to rules or concepts (like in rule-systems of artificial intelligence), but according to a different instinctual-emotional mechanism described in the given references.

Conceptual-emotional understanding of the world results in actions in the outside world or within the mind. We only touch on the *behavior* of improving understanding and knowledge, the behavior inside the mind directed at improving concepts. Let us mention that there are "lower-level" autonomous behavioral responses, which humans share with animals and which do not involve mechanisms of concepts. We would not need to consider them here for understanding the role of music.

The above theory describing conceptual-emotional recognition and understanding encompasses the mechanisms of intuition, imagination,

planning, conscious, unconscious, and many others, including aesthetic emotions. Here we would touch on mechanisms that will be referred to later. For example, visual imagination occurs when one contemplates objects or situations with closed eyes. Contemplated concept-models project images on visual cortex causing visual imagination. Most of the brain operations are unconscious, for example, individual neuronal firings. A significant part of conceptual perception is an unconscious process; for example, visual perception takes about 150 ms, which is a long time when measured in neuronal firings (about 10 ms per neuron). Initial concept-model projections on the visual cortex are vague and the human mind is not conscious of them. Only when concept-model projections match objects and become crisp do conscious perceptions occur. It is possible to make the vague concept-model projections conscious: close your eyes and imagine an object in front of you; this imagination is usually vague, not as crisp as perception of an object with open eyes. Let us now move to mechanisms of aesthetic emotions.

The Knowledge Instinct

To satisfy instinctual needs, e.g. eating or procreation, the mind first of all should perceive objects around and understand situations. As discussed, this task requires matching concept-models to the surroundings. But objects around would never exactly match old concept-model-memories. Angles, lightings, and positions are always different which, in fact, has presented difficulties to artificial intelligence and pattern recognition since the 1950s until recently (e.g. see Perlovsky, 2000, 2006d). For this reason initial projections of concept-models are vague and they approximately match many different objects. To actually perceive specific objects, the mind has to modify concepts so that they "fit" concrete objects and situations present in the ever-changing world. This mechanism operates independently of human desire "to perceive," it is an inborn autonomous mechanism, more fundamental than eating or procreation. It is aimed at satisfying a basic need, to understand the world around by making concept-models "similar" to surroundings. The mind has an inborn instinct that "senses" this similarity and maximizes it. This mechanism is called the knowledge instinct (Perlovsky, 2006d) since knowledge is the measure of correspondence between concepts and the world.

Emotions that evaluate satisfaction or dissatisfaction of this instinct are felt as harmony or disharmony between the knowledge and the world. They are not related directly to "lower" bodily needs, but only to "higher" need for knowledge. In this sense they are "higher," "spiritual," aesthetic emotions. We would like to emphasize that aesthetic emotions are not peculiar to perception of art; they are inseparable from every act of perception and cognition. Relation of these emotions to the beautiful and to musical emotions will be considered later. During perception of everyday objects these emotions usually are below a threshold of conscious registration. We do not feel emotionally elated when correctly understand a simple everyday object in front of our

eyes. But, due to scientific knowledge of cognitive neural mechanisms and their mathematical models we know that these emotional neural signals are there. And it is easy to prove experimentally. As soon as perception and understanding of the surrounding world does not work we feel disharmonious, disturbed, or even threatened – this is the routine matter of thriller movies, which show us situations that do not fit our concept-models. At the level of simple objects this perception mechanism is mostly autonomous, like workings of our stomach. As long as a stomach works perfectly, we do not notice its existence emotionally. But as soon as it fails, we feel it emotionally right away.

The Hierarchy of the Mind

The mind is not a strict hierarchy and many cognitive neuroscientists refer to the heterarchy of the mind. For simplicity we will use the word hierarchy in this paper. At every level of the hierarchy, top-down signals generated by concept-models at this level are matched to bottom-up signals coming from concept-models recognized and understood at lower levels. The mind involves a hierarchy of multiple levels of concept-models, from simple perceptual elements (like edges, or moving dots), to concept-models of objects, to complex scenes, and up the hierarchy… toward the highest concept-models. These highest concept-models near the top of the hierarchy are essential for understanding the nature of the beautiful and spiritually sublime (Perlovsky, 2002, 2006a, 2006d).

To understand this let us first attend to the perception-cognition of a simple situation-scene, say an office of a professor. It is not sufficient for our knowledge instinct to understand individual objects in the office such as books, shelves, chairs, desk, computer… we can sit in a chair or read a book, but this understanding will only take us so far (animals also understand objects and what they can do with some of them). The knowledge instinct drives us to understand "the office" in its unity of constituent objects. For this purpose we have a higher-level concept-model of an "office." Similarly, we understand a concert hall, and any other situation by using appropriate-level concepts that we have for this purpose. Let me repeat this word: purpose; every higher-level concept has a purpose to make a unified sense out of individual lower-level concepts. In this process lower-level concepts acquire higher-level "sense" or meaning of making up something "bigger", something more meaningful, than their lower-level meanings. In this way our understanding of the world can move from a "book" to "office," to "university", to "educational system," and so on… to concepts near the top of our minds. These concepts "attempt" to make sense, to understand the meaning of our entire experience. We understand-perceive-feel them as related to the meaning and purpose of our lives.

This last sentence requires several clarifications. First, let me remind that even a simple object, when imagined with closed eyes is vaguer and less

conscious than when perceived with open eyes. But abstract concepts at higher levels of the mind hierarchy can not be "perceived with open eyes." Correspondingly they are forever vaguer and less accessible to our consciousness than simple objects. This statement may sound startling; we do not necessarily experience all abstract concepts this way. The reason abstract concepts may sometimes seem crisp, clear, and conscious will be addressed in the next section. Second, vaguer and less conscious concepts may also be mixed up with emotional contents. For example, talking about your favorite political party may require special efforts to separate conceptual understanding from emotional involvement. This is why concepts at the top of our mind could be at once less conscious and emotionally charged. This combination makes it difficult for us to discuss these concepts. Many of my friends (scientists) when asked: "Does your life have a meaning and purpose?" will reply with great doubts. However, as soon as the question is reformulated: "So your life does not have any more meaning and purpose than that piece of rock at the side of the road?" At this point most of people agree that the idea of the meaning and purpose of life might be vague and barely conscious, but it is so important that we cannot live without it. In fact reading this paper would be a very boring exercise, if you do not believe that your life has a purpose. It would be more fun to get drunk or high on drugs.

Life experience does not convince us that our lives have meaning and purpose; random deaths and destructions abound. But believing in one's purpose is tremendously important for survival; it is necessary for concentrating will and power on achieving higher goals in life. This is why even partial understanding of contents of the highest concept-models is so important. When we feel that indeed our lives have meaning, in these rare fleeting moments we feel the knowledge instinct satisfaction at the highest level as an aesthetic emotion of the beautiful.

This discussion is not necessarily novel in the history of philosophy. Aristotle and Kant discussed similar ideas. Aristotle wrote (IV BCE/1995) that the beautiful is a "unity in manifold." The only way to understand the world in its unity, he wrote, is as if it had a purpose. Kant (1790) understood the beautiful as "aimless purposiveness" of the faculty of judgment; Kantian judgment corresponds to mechanisms of aesthetic emotions as discussed in Perlovsky (2006d). Also, "aimless" in Kant means that it is not aimed at satisfying lower bodily needs. Kant did not appreciate a need for adaptation of concept-models and could not formulate the idea of the knowledge instinct. This caused him great difficulty, he goes around "aimless purposiveness" emphasizing that it is not aimless, that it is highly spiritual, but without the knowledge instinct idea he could not give a positive definition of the beautiful. I added a contemporary scientific context. My formulations might be crisper, because they are based on a mathematical theory (Perlovsky, 2006d).

The Dual Hierarchy of Cognition and Language

The mind hierarchy as discussed above tacitly assumed a single hierarchy of cognitive models. To resolve certain difficulties glossed over in the above, and to get closer to understanding musical emotions, we now consider the dual hierarchy of cognition and language; we summarize arguments discussed in details in Perlovsky (2006d, 2006c, 2007).

Recognizing that cognition and language are not the same, that these abilities are served by different mechanisms of the mind, started a revolution in the 20th century linguistics initiated by Chomsky (1957). Many psycholinguists and evolutionary linguists today disagree with Chomsky's complete separation of language from cognition and denial of evolutionary origin of language. Detailed discussions can be found in the given references and further references therein. Here we summarize conclusions important for understanding the role of music.

We start with few difficulties encountered when attempting to understand the mechanisms of interaction of cognition and language. Language plays such an important role in thinking that it is difficult to comprehend what cognition is without language, what are the mechanisms? Normal children acquire language by about the age of five, by seven they can talk about much content of the entire culture. If mother is looking for a job, a kid it seems knows everything about this process. But of course the kid cannot go on the street and find job. What exactly is missing in terms of neural mechanisms? How kids learn which words and sentences correspond to which objects and situations? Many psychologists still adhere to "associationism," an idea due to Locke that kids just learn associations between words and objects, etc. But as discussed in the given references, this is mathematically impossible. Some people master language very well, while inept with other people; opposite examples also abound. So, what are the mechanisms that make language and cognition so interdependent, and at the same time so separate? And what exactly animals are missing that they cannot learn language?

According to given references, the main mechanism of interaction between cognition and language is a dual concept-model. Each concept has two parts, language part (a word or phrase) and cognitive part (an object or situation). When a child is born these are vague placeholders that later will acquire a concrete content. By the age of five much of language models are crisp, clear, and conscious, but the corresponding cognitive models may remain vague and unconscious. By the age of four, everyone knows e.g. about good and bad guys, but who can claim at 30 or 40 or 70, that he or she can use these concepts in real life without errors? Philosophers argue about the meanings of good and evil for millennia. Even for everyday concepts, which linguistic parts are crisp and conscious in every child's mind, it will take the rest of life to acquire equally crisp and conscious cognitive models. Likely most of cognitive concept-models never attain equally conscious and crisp states. This is why most people most of the time speak with words without being fully conscious about cognitive contents of what they say. These properties of

language-cognition interaction are explained by the mechanism of dual models.

The mechanism of dual models is fundamentally important for emergence of the hierarchy of the mind for the following reason. Learning concept-models of cognition is grounded in experience only at the lower levels of concrete objects, at this level human abilities are no different from animals' (Spelke, & Kinzler, 2007). But understanding situations and abstract concepts can not be based on experience. The referenced publications discuss in details why this is mathematically impossible: there are too many combinations of objects and events (more then all elementary events in the life of the Universe). Life experience would never be sufficient to learn which combinations are meaningful to form abstract concepts.

Possibly the mechanism of dual models is sufficient to make the difference between humans and animals. It is possible to teach a dog to bring shoes on verbal command. Does it mean that a dog possesses a mechanism of dual models? No, a dog can learn concepts of object-shoes and object-sound-"shoes" based on direct visual and hearing perception, and learns to associate these two concept-objects. But nobody would even attempt to teach a dog abstract concepts such as rationality, abstractness, or law. Learning these abstract concepts requires the mechanism of dual models: the knowledge instinct drives the mind to find combinations of objects and events, which correspond to language models.

Learning cognitive models at higher levels is based on language models. Language hierarchy is learned "ready-made" from the surrounding language. The knowledge instinct drives the mind to learn cognitive hierarchy corresponding to language hierarchy. Cognitive models are grounded in language.

Differentiation and Synthesis

The knowledge instinct operates in the dual hierarchy of the mind with two main mechanisms, differentiation and synthesis Perlovsky (2006d, 2007, 2008). At every level of the hierarchy it drives the mind to achieve detailed understanding by creating more specific, diverse and detailed concepts—this is the mechanism of differentiation. At the same time (as we discussed), the knowledge instinct drives us to understand various situations and abstract concepts as a unity of constituent notions. This mechanism of the knowledge instinct operating across hierarchical levels creates higher meanings and purposes—this is a mechanism of synthesis.

The main "tool" of differentiation is language. Language gives our mind a culturally evolved means to differentiate reality in great detail. The evolution of language required neural rewiring of circuits controlling vocalization. Vocal tract muscles in animals are controlled from an old emotional center and voluntary control over vocalization is limited (Deacon, 1989; Schulz, Varga,

Jeffires, Ludlow,& Braun, 2005; Davis, Zhang, Winkworth, & Bandler, 1996; Larson, 1991). Humans, in contrast, possess a remarkable degree of voluntary control over voice, which is necessary for language. In addition to the old mostly involuntary control over vocal tract human have conscious voluntary control originating in cortex.

Correspondingly, conceptual and emotional systems (understanding and evaluation) in animals are less differentiated than in humans. Sounds of animal cries engage the entire psyche, rather than concepts and emotions separately. A well-known example is differentiated calls of vervet monkeys (e.g. see a review in Seyfarth & Cheney, 2003). The calls convey information about different types of predators nearby; however understanding of a situation (concept of danger), evaluation (emotion of fear), and behavior (cry and jump on a tree) are not differentiated, each call is a part of a single concept-emotion-behavior-vocalization psychic state with very little differentiated voluntary control (if any).

Emotions-evaluations in humans have separated from concepts-representations and from behavior (For example, when sitting around the table and discussing snakes, humans do not jump on the table uncontrollably in fear, every time "snakes" are mentioned). We hypothesize that gradual differentiation of psychic states with a significant degree of voluntary control over each part gradually evolved along with language and the brain rewiring.

Therefore, language contributed not only to differentiation of conceptual ability, but also to differentiation of psychic functions of concepts, emotions, and behavior. This differentiation destroyed the primordial synthesis of psyche. With evolution of language human psyche started loosing synthesis, wholeness. Whereas for animals every piece of "conceptual knowledge" is inextricably connected to emotional evaluation of a situation, and to appropriate behavior, satisfying instinctual needs, this is not so for humans. Most of knowledge existing in culture and expressed in language is not connected emotionally to human instinctual needs. This is tremendously advantageous for development of conceptual culture, for science, and technology. Humans can engage in deliberate conversations, and if disagree, do not have to come to blows. But there is a heavy price that humans pay for this freedom of conceptual thinking: human psyche is not automatically whole. Human knowledge accumulated in language is not automatically connected to instinctual needs; sometimes culturally developed conceptual knowledge contradicts instinctual needs inherited from the animal past. Moreover, various parts of knowledge may contradict each other. As discussed, synthesis, the feel of being whole is closely related to successful functioning of the highest models at the top of the hierarchy of the mind, which are perceived as the meaning and purpose of life. Therefore contradictions in the system of knowledge, a disconnect between knowledge and instincts, the lost synthesis, lead to internal crises and may cause clinical depressions. When psychic states missing synthesis preoccupy majority of population, knowledge loses its value, including knowledge of social

organization, cultural calamities occur, wars and destructions (Perlovsky, 2006b, 2006e, 2007, 2008; Diamond, 1997). Evolution of culture requires a balance between differentiation and synthesis. Differentiation is the very essence of cultural evolution. But it may lead to emotional disconnect between conceptual knowledge and instinctual needs, to the lost feeling of the meaning and purpose, including the purpose of any cultural knowledge, and to cultural destruction. Theoretical and experimental evidence suggest that different languages maintain different balances between the emotional and conceptual (Perlovsky, 2007; Harris, Ayçiçegi, & Gleason, 2003; Buchanan, Lutz, Mirzazade, Specht, Shah, Zilles, et al, 2000). Below we examine mechanisms by which music influences this balance.

Differentiated Knowledge Instinct and Musical Emotions

As we discussed, differentiation is the very essence of cultural evolution, but it threatens synthesis and may destroy the entire purpose of culture, and culture itself.**Error! Bookmark not defined.** This instability is entirely human, it does not threaten the animal kingdom because the pace of evolution and differentiation of knowledge from ameba to primates was very slow, and instinctual mechanisms of synthesis apparently evolved along with the brain capacity. This situation drastically changed with the origin of language; accumulation of differentiated knowledge vastly exceeded biological evolutionary capacity to maintain synthesis. Along with the origin of language another uniquely human ability evolved, ability for music. We propose here a scientific hypothesis that music evolved for maintaining the balance between differentiation and synthesis. After formulating arguments, we discuss experimental means by which this hypothesis can be verified.

Many scientists studying evolution of language came to a conclusion that originally language and music were one (Darwin, 1871; Cross, 2008a; Masataka, 2008). In this original state the fused language-music did not threaten synthesis. Not unlike animal vocalizations, sounds of voice directly affected ancient emotional centers and connected semantic contents of vocalizations to instinctual needs. This synthesis was a direct inheritance from animal voicing mechanisms, and to this very day voice affects us emotionally directly through ancient emotional brain centers (Panksepp, & Bernatzky, 2002; Trainor, 2008).

We would like to emphasize the already discussed fact that since its origin language evolved in the direction of enhancing conceptual differentiation ability by separating it from ancient emotional and instinctual influences (here we mean “bodily” instincts, not instincts for knowledge and language). While language was evolving in this more conceptual and less emotional direction, we suggest that ‘another part’ of human vocalization evolved toward less semantic and more emotional direction by enhancing already existing mechanisms of voice-emotion-instinct connection. As language was enhancing

differentiation and destroyed the primordial unity of psyche, music was reconnecting differentiated psyche, restoring the meaning and purpose of knowledge and making cultural evolution possible. Was this process equally successful in every culture? Probably not, but this is a separate field of study for future research.

This was the origin and evolutionary direction of music. Its fundamental role in cultural evolution was maintaining synthesis in the face of increasing differentiation due to language. We now return to the basic mechanisms of the mind, including the knowledge instinct and analyze them in more details in view of this hypothesis.

Discussing the knowledge instinct in previous sections we described the mathematical model of its mechanism, an internal mind's "sensor" measuring similarity between concept-models and the world and related mechanisms of maximizing this similarity. But clearly it is a great simplification. It is not sufficient for the human mind to maximize an average value of the similarity between all concept-models and all experiences. Adequate functioning requires constant resolution of contradictions between multiple mutually contradicting concepts and between individual concepts quickly created in culture and slowly evolving primordial animal instincts. Human psyche is not as harmonious as psyche of animals. Humans are contradictory beings; as Nietzsche (1995/1876) put it, "human is a dissonance." Those of our ancestors who were able to acquire differentiated contradictory knowledge and still maintain wholeness of psyche necessary for concentration of will and purposeful actions had tremendous advantage for survival.

Therefore, we suggest that the knowledge instinct itself became differentiated. It was directed not only at maximizing the overall harmony, but also at reconciling constantly evolving contradictions. This is a hypothesis that requires theoretical elaboration and experimental confirmation. As discussed, emotions related to the knowledge instinct are aesthetic emotions subjectively felt as harmony or disharmony. These emotions had to be differentiated along with the knowledge instinct. Consider high value concepts such as one's family, religion, or political preferences. These concepts 'color' with emotional values many other concepts; and every contradictory conceptual relation requires a different emotion for reconciliation, a different dimension of an emotional space. In other words, a high value concept attaches aesthetic emotions to other concepts. This is a function of the knowledge instinct; concepts act as separate parts of the knowledge instinct; this explains the notion of the differentiated knowledge instinct. Virtually every combination of concepts has some degree of contradictions. The number of combinations is practically infinite (Perlovsky, 2006d). Therefore aesthetic emotions are not just several feelings for which we can assign specific words. There is an infinity, continuum of aesthetic emotions, and most likely the dimensionality of this continuum is huge. We feel this continuum of emotions when listening to music. We feel this continuum in Palestrina, Bach, Beethoven, Mozart, Chaikovsky, Shostakovich, Beatles, and Eminem (and certainly this

mechanism is not limited to western cultures).

I would mention that Spinoza (2005/1677) was the first philosopher to discuss the multiplicity of emotions related to knowledge. Each emotion, he wrote, is different depending on which object it is applied to. There is a principled difference between multiplicity of aesthetic emotions and 'lower' emotions corresponding to bodily instincts. Those emotions, as discussed, are referred to as 'basic' emotions in psychological literature (e.g. see Juslin, & Västfjäll, 2008). As discussed, psychologists identify them; they all have special words, such as 'rage' or 'sadness.' Levitin (2008) argues that there are just six basic types of songs, basic emotions, all related to basic instinctual needs. But Huron (1999) has already argued that this use of music for basic needs is just that, a utilitarian use of music, which evolved for a much more important purpose that cognitive musicologists had not yet been able to identify.

We suggest in this paper that music has two interrelated purposes fundamental to the functioning of individual minds and to evolution of the mind and culture. The first purpose is to differentiate aesthetic emotions. Music creates differentiated emotions required to reconcile conceptual contradictions. The second purpose is to connect concepts to instinctual needs (including the knowledge instinct). Whereas language separates conceptual knowledge from instincts and emotions, music reconnects these ties. Both musical functions suggested here are scientific hypotheses that could and should be further explored theoretically and verified experimentally.

Further Directions

One direction for further research is to relate changes in musical styles to changes in cultures and consciousness. This will connect evolution of music, consciousness, and cultures. A step in this direction was made in Perlovsky (2006b, 2006e, 2008). It was suggested for example that antiphonal music appeared about 2500 years ago along with contemporary consciousness, when fundamental contradictions in human psyche started penetrating into consciousness and created psychic tensions. Tonality was developed beginning in Renaissance, when instinctual and emotional human nature was consciously accepted, creating tensions in psyche with received ideas of spiritually 'high.' Buxtehude and Bach were developing music that could reconcile new contradictions brought in consciousness by the Reformation. Popular songs restore synthesis by connecting conceptual contents of lyrics with emotional contents of music. And contemporary rap music was suggested to have a similar style and function to Ancient Greek dithyrambs, namely to reconcile instinctual needs with (at least some) basic concepts in culture and language.

The complexity of emotional spaces (structure, dimensions) can be studied for various types of music by using existing mathematical techniques. It would be interesting to compare emotional spaces of Eminem and Beethoven to

confirm or disprove various expectations. Experimental techniques of measuring musical emotions would likely have to be perfected first.

The role of timbre in music and language might be related to the discussion in this paper. Levitin (2006) writes that timbre characterizes individual performers more than any other aspect of music. Patel (2008) suggests that language uses timbre systematically more than music does. Is timbre in itself "semantic," whereas melody "emotional"? Is harmony related to the mind hierarchy? Are these intuitions just shallow metaphors or meaningful, experimentally testable hypotheses related to the initial separation of voice into language and music, and to further evolution of cultures and consciousness?

Here we would like to emphasize possible directions for experimental verifications of the suggested mechanisms of the knowledge instinct, dual models, musical emotions and their role in the mind functioning. Neuroimaging techniques can be used for identifying brain modules and neural connections involved in the dual models and knowledge instinct (Levine & Perlovsky, 2008). Experimental techniques used to study cognitive dissonance (Festinger, 1957) can be used to study musical emotions and their role in reconciling contradictions in consciousness. Various types of music can be associated with reconciling specific types of conceptual dissonances (we would expect that the results would depend on psychological types of listeners, people who's feelings are less differentiated might be more affected by tonal music, whereas people consciously differentiating many emotions might be more susceptible to atonal music – this comment, however, is secondary to the main ideas discussed here); neuroimaging techniques can be used in parallel to identify the brain regions involved.

References

Aristotle. (1995). *The complete works.* (The revised Oxford translation, ed. J. Barnes), Princeton, NJ: Princeton Univ. Press. (Original work VI BCE)

Ball, P. (2008). Facing the music. *Nature, 453,* 160-162.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, et al. (2006). Top-down facilitation of visual recognition. USA: *Proceedings of the National Academy of Sciences, 103,* 449-54.

Buchanan, T. W., Lutz, K., Mirzazade, S. Specht, K., Shah, N.J., Zilles, K. et al. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Cognitive Brain Research 9,* 227-238.

Chomsky, N. (1957). *Syntactic Structures.* Haag, Netherlands: Mouton.

Cross, I. (2008a). The evolutionary nature of musical meaning. To appear in *Musicae Scientiae, Special issue on music and evolution.* Preprint <http://www.mus.cam.ac.uk/%7Eic108/PDF/IRMC_MS07_1.pdf>

Cross, I. (2008b). Musicality and the human capacity for culture. *Musicae Scientiae, Special issue*, 147-167.

Cross, I., & Morley, I. (2008c). The evolution of music: theories, definitions and the nature of the evidence. In S. Malloch, & C. Trevarthen (Eds.), *Communicative musicality* (pp. 61-82). Oxford: Oxford University Press.

Darwin, C.R. (1871). *The descent of man, and selection in relation to sex.* London, GB: John Murray.

Davis, P. J., Zhang, S. P., Winkworth A., & Bandler, R. (1996). Neural control of vocalization: respiratory and emotional influences. *J Voice,10,* 23–38.

Deacon, T. (1989). The neural circuitry underlying primate calls and human language. *Human Evolution Journal, 4(5),* 367-401.

Diamond, J. (1997). *Guns, germs, and steel: The fates of human societies.* New York, NY: W.W. Norton, & Co.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interactions. In N. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 389-407). Cambridge, MA: MIT Press.

Dissanayake, E. (2008). If music is the food of love, what about survival and reproductive success? *Musicae Scientiae Special Issue,* 169-195.

Editorial. (2008). Bountiful noise. *Nature, 453,* 134.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Fitch, W. T. (2004). On the biology and evolution of music. *Music Perception, 24,* 85-88.

Grossberg, S., & Levine, D. (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: Blocking, interstimulus interval, and secondary reinforcement. *Applied Optics 26,* 5015-5030.

Harris, C. L., Ayçiçegi, A., & Gleason, J.B. (2003). Taboo words and

reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics, 24,* 561-579.

Huron, D. (1999). *Ernest Bloch Lectures.* Berkeley, CA: University of California Press.

Jaynes, J. (1976). *The origin of consciousness in the breakdown of the bicameral mind*. Boston: Houghton Mifflin Co.

Juslin, P.N., & Västfjäll, D. (2008) Emotional responses to music: The Need to consider underlying mechanisms. *Behavioral and Brain Sciences,* in print.

Justus, T., & Hustler J. J. (2003). Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain specificity. *Music Perception, 23,* 1-27.

Kant, I. (1790). *The critique of judgment.* J.H. Bernard (translator). Amherst, NY: Prometheus Books.

Larson, C.R. (1991). Activity of PAG neurons during conditioned vocalization in the macaque monkey. In A. Depaulis, & R. Bandler (Eds). *The midbrain periaqueductal gray matter* (pp. 23–40). New York, NY: Plenum Press.

Levine, D. S., & Perlovsky, L. I. (2008). Neuroscientific insights on Biblical myths. Simplifying heuristics versus careful thinking: Scientific analysis of millennial spiritual issues. *Zygon, Journal of Science and Religion, 43(4),* 797-821.

Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession.* London: Dutton.

Levitin, D. J. (2008). *The world in six songs.* London: Dutton.

Lorenz, K. (1981). *The foundations of ethology*. New York: Springer Verlag.

Masataka, N. (2008). The origins of language and the evolution of music. *Physics of Life Reviews,* in print.

Meyer, R. K., Palmer, C., & Mazo, M. (1998). Affective and coherence responses to Russian laments. *Music Perception, 16(1),* 135-150.

McDermott, J., & Houser, M. (2003). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23,* 29-59.

McDermott, J. (2008). The evolution of music. *Nature, 453,* 287-288.

Mithen, S. (2007). *The singing Neanderthals: The origins of music, language, mind, and body.* Cambridge MA: Harvard University Press.

Nietzsche, F. (1997). *Untimely Meditations.* R. J. Hollingdale (Translator). Cambridge, England: Cambridge University Press (original publication in 1876).

Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: The neuro-affective foundations of musical appreciation. *Behavioural Processes, 60,* 133-55.

Patel, A. D. (2008). *Music, language, and the brain.* New York, NY: Oxford Univ. Press.

Perlovsky, L. I. (2000). *Neural networks and intellect.* New York, NY: Oxford Univ. Press.

Perlovsky, L. I. (2001). Mystery of sublime and mathematics of intelligence. *Zvezda, 2001(8),* 174-190, St. Petersburg (Russian).

Perlovsky, L.I. (2002) Aesthetics and Mathematical Theory of Intellect. Russian Academy of Sciences, Moscow, Russia: *Iskusstvoznanie, Journal of History and Theory of Art, 2,* 558-594 (Russian).

Perlovsky, L. I. (2005). Evolution of consciousness and music. *Zvezda, 2005(8),* 192-223, St. Petersburg (Russian).

Perlovsky, L.I. (2006a). Modeling field theory of higher cognitive functions. In A. Loula, R. Gudwin, J. Queiroz (Eds.) *Artificial cognition systems.* Hershey, PA: Idea Group (pp. 64-105).

Perlovsky, L.I. (2006b). Music–the first principles. *Musical Theater,* <http://www.ceo.spb.ru/libretto/kon_lan/ogl.shtml>

Perlovsky, L.I. (2006c). Symbols: Integrated cognition and language. In R. Gudwin, J. Queiroz (Eds.). *Semiotics and intelligent systems development.* Hershey, PA: Idea Group (pp.121-151).

Perlovsky, L.I. (2006d). Toward physics of the mind: Concepts, emotions, consciousness, and symbols. *Physics of Life Reviews, 3(1),* 22-55.

Perlovsky, L.I. (2006e). Joint evolution of cognition, consciousness, and music. *Lectures in Musicology, School of Music,* Columbus, OH: University of Ohio.

Perlovsky, L.I. (2007). Evolution of languages, consciousness, and cultures. *IEEE Computational Intelligence Magazine, 2(3),* 25-39.

Perlovsky, L.I. (2008). Music and consciousness. *Leonardo, Journal of Arts, Sciences and Technology, 41(4),* 420-421.

Pinker, S. (1997). *How the mind works.* New York, NY: Norton.

Schulz, G. M., Varga, M., Jeffires, K., Ludlow, C. L., & Braun, A. R. (2005). Functional neuroanatomy of human vocalization: an H215O PET study. *Cerebral Cortex, 15(12),* 1835-1847.

Seyfarth, R. M., & Cheney, D.L. (2003). Meaning and emotion in animal vocalizations. *Ann NY Academy Sci, Dec.,* 32-55

Spelke, E. S., & Kinzler, K. D (2007). Core knowledge. *Developmental Science, 10,* 89-96.

Spinoza, B. (2005). *Ethics.* (E .Curley, Translator). New York, NY: Penguin. (Originally published in 1677).

Trainor, L. (2004). Innateness, learning, and the difficulty of determining whether music is an evolutionary adaptation. *Music Perception, 24,* 105-110.

Trainor, L. (2008). The neural roots of music. *Nature, 453(29),* 598-599.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience, 6(7),* 669-673.