ISSN 1648-2824 KALBŲ STUDIJOS. 2014. 25 NR. * STUDIES ABOUT LANGUAGES. 2014. NO. 25

Symbolizing Iconic Indexes: An Intentionality-based Hypothesis on the Emergence of Music

Alessandro Miani

crossref http://dx.doi.org/10.5755/j01.sal.0.25.8535

Abstract. The most recognized feature of music is the capacity to arouse emotions in listeners: are such emotions inherent to music (indexes) or are they evoked by resemblance to an analogous emotional expression (icons)? The claim is that music is a conventionalized imitation of an expression of an emotional state (i.e., a symbolized iconic index), but to reach this form three evolutionary steps in human thinking are required: individual, joint, and collective intentionality, which parallel the emergence — and overlap — of indexes, icons, and symbols. Providing a comparative view of the genetic and neuroanatomical infrastructures required for the emergence of music, it will be hypothesized that: (i) music is rooted in ape vocalizations as an index of emotions performed by individual-intentionality agents; (ii) with the capacity for recursive mindreading of joint-intentionality agents, music evolved as an iconized index of emotions; (iii) as a consequence of demographic changes, collective-intentionality agents created musical instruments in order to reduce the structural complexity of the sign aiming at coordinating in joint musical activities and culturally transmitting symbolized iconic indexes of emotions.

***Key words**: shared intentionality, evolution of music, comparative data, vocal learning, social cognition, music and language, FOXP2.*

It started, perhaps, with the throwing of rocks, followed in succession by axes, spears, boomerangs, bows and arrows, guns, rockets, bombs, nuclear missiles, not to mention insults.

(Corballis, The Recursive Mind)

Introduction

"Let me die" says Arianna (*Lamento d'Arianna*, Monteverdi, 1607): her voice is in a low register, words are spoken slowly interleaved with some breaks. Clearly she is not happy, she is experiencing the emotional state of sadness, musically depicted by low pitch, long notes, and the use of rests. Putting her emotions in the score, Monteverdi musicalized an imitation of an expression of an emotional state. He symbolized an iconic index of emotions.

According to Peirce (1931/ 1958), the relation between a sign and its referent can occur in indexical, iconic, and symbolic ways: an index is a sign that has a causal connection with its referent (smoke is causally connected to the presence of fire), an icon is a sign that maintains a sensorial resemblance (a portrait resembles the depicted person), and in a symbol, such link is arbitrary (the word "dog" has neither resemblance, nor causal connection with the referent "dog"). These sign qualities have been recently used by Koelsch (2011a; 2012) for a taxonomy of musical meanings: an index signals the psychological state of an individual (e.g., emotion, mood, intention), an icon is a musical imitation of an extra-musical sound (e.g., dog, elephant, thunderstorm), and a symbol is conventionally or idiosyncratically associated to an extra-musical event (e.g., national anthem, ringtone, jingle). It has to be stressed, though, that these qualities of the sign do often — if not always, to different extents — overlap (Peirce 1931/ 1958; Garrod et al., 2007). For example, Kivy (2002) noted that the St Bernard's face is expressive of sadness inasmuch as we see it resembling human face when expressing sadness due to some visual aspects of its face that are similar to those of human in expressing sadness (sad eyes, wrinkled brow, drooping mouth, and so on). Similarly, Koelsch stated that indexes emerge from "the imitation of expressions signaling the psychological state of an individual" (2012, p.159). On such overlapping property of signs, there will be proposed here an evolutionary hypothesis for the emergence of music, assuming that from unintentional indexes of emotion, through an intentional iconization, music reached the status of symbolized iconic index of emotions. In doing so, music will be seen as a technology based on the capacity to understand other's mind; an integrative view borrowed from two other evolutionary proposals.

According to Patel (2010), music is a "transformative technology of the mind" which has lasting effects on our brain. It is neither a biological adaptation (see e.g., Miller, 2000; Mithen, 2005; Levitin, 2006), nor a frill (referring to the Pinker's cheesecake; 1997), but rather, like the control of fire, music is an invention that has been intimately integrated with our life to the extent that "there is no a going back, even though we might be able to live without this ability" (Patel, 2008, p.401): neither fire, nor music are crucial for life, clearly, but they are omnipresent. In fact, although having few universals, most of them related to psycho/ biological constraints (e.g., Brown & Jordania, 2011), music is itself a universal activity among humans, even though it has "no readily apparent functional consequence" (Hauser & McDermott, 2003, p.663), and it is also present in the community of Pirahã, of which language does not consider numbers, colors and recursive mechanisms (Everett, 2005). This is because music is tremendously powerful: like food, sex, and drug abuse, it activates the brain regions involved in reward/ motivation, emotions, and arousal (Blood & Zatorre, 2001). Physiologically speaking, it is able to modulate pain (Roy et al., 2012), to affect the consumption of sedative (Koelsch et al., 2011), and to enhance recovery — such as verbal memory and selective attention — after a stroke (Särkämö et al., 2008); furthermore, musicians have a bigger volume of grey matter (Gaser & Schlaug, 2003) and corpus callosum (Schlaug et al., 1995). Patel's approach is particularly useful inasmuch as he claims that music is an invention (resonating with Everett (2012) in regard to language) built on pre-existing cognitive abilities. Being a technology, it is subjected to the ratchet effect, i.e., it is constantly improved through imitative learning and modifications in the context of faithful social transmission (Tomasello, 1999).

Another evolutionary approach considers the involvement of the theory of mind, which is the milestone for human-unique cognition (Tomasello, 1999; 2008; 2014; Corballis, 2011), referring to the ability of humans to understand other beings as having their own mind. According to Livingstone and Thompson (2009), constructing a model of the mental states in order to predict actions (theory theory) and simulating hypothetical states of mind — empathizing (simulation theory) — establish the musical communication as "an affective sandbox for safe emotional exploration" and its use in cultural rituals (Livingstone & Thompson, 2009, p.95). By making a comparison with other arts (painting, sculpture, story-telling, adornments) and speech (prosody), what crucially emerges from their proposal is that music — and arts in general — is (and has been) used for eliciting emotions. So, music is a means for sharing emotions, which is specific to humans (Tomasello, 2008; Corballis, 2011).

1. Is Music Unique to Humans?

It is presumable to think that music and language are two different aspects of the same communicative phenomenon (Cross et al., 2013; Koelsch, 2011b), supposed to be evolved from a common ancestor (e.g., musilanguage; Brown, 2000), and speciated in order to fulfill emotional and referential purposes. As a matter of fact, music and language share many structural similarities: both are based upon a finite set of discrete elements and principles of combination which rule a hierarchical organization; they both make use of symbolic sounds and are structured according to phonology, syntax, and semantics (e.g., Lerdahl & Jackendoff, 1983; Sloboda, 1985; Patel, 2008). More interesting, many studies have demonstrated a neural overlap in processing (e.g., Maess et al., 2001; Patel, 2003) and also that music, like language, is able to convey semantic meaning (Koelsch et al., 2004; Painter & Koelsch, 2011).

However, there are two basic distinctions that must be moved forward: language relies mainly on timbre for phoneme intelligibility, whereas music does on rhythm for synchronization. Also language is rhythmically organized in order to coordinate people in conversational turn talking (Patel, 2010), but music is unique in the fact that two individuals can synchronously communicate at the same time, which is deleterious for language intelligibility. In regard to music, it has been pointed out that the late stages of music perception overlap with the early stages of actions (Koelsch, 2011b). According to Koelsch (2010; 2012), making music together entails a set of social advantages: it prevents social isolations by making contact with other individuals; it engages social cognition; it makes inter-individual emotional states more homogenous; it boosts communication, especially for infants and young children; it triggers pleasure, even in absence of explicit shared goals; it requires cooperation and inter-individual trust; and finally, it increases the social cohesion in a group.

The ease by which humans engage in joint actions through musical activity, suggests that our closest and intelligent living relatives, the great apes, may engage, at least, in rhythm. Merchant and Honing (2014) reviewed such issue pointing out that among primates, humans are the only ones able to entrain rhythm, but evidence leads to a "gradual audiomotor evolution hypothesis": rhesus monkeys, for instance, cannot extrapolate beat from a complex auditory rhythm, yet they can detect the starting point (Honing et al., 2012). However, despite the lack of rhythmic engagement, nonhuman primates seem more sensitive to pitch: Japanese macaques are good to discriminate consonant vs. dissonant intervals (Izumi, 2000), which elicit different neural patterns (Fishman et al., 2001), but they do not show any preferences (McDermott & Hauser, 2004). So, how can it be possible to pinpoint the emergence of music given that our closest relatives are not musically gifted?

Indeed, the present research suggests that music emerged from the vocalizations of non-musically gifted apes. The paper is subdivided in three sections: in a comparative perspective, the first two parts show the biological and psychological — hardware & software — infrastructure. The first part concerns the vocal learning ability presented from a behavioral, neuroanatomical, and genetic perspective, while a brief digression on human and ape emotions is outlined in order to pave the way for the motives underlying music. The second part is narrowed to the foundation for language and music, i.e., shared intentionality, and a partial view on human mindreading and gestures in human children and nonhuman primates will lead to the formulation of the hypothesis. It will be suggested that music started from ape vocalizations, moved towards a separation from emotions, and then conventionalization. Such changes are reflected by a change and overlap in the nature of the sign.

Vocal Learning: Comparative Studies

One might raise a question: does the rhythmic entrainment suggest higher cognitive abilities? Given that other primates cannot afford it, the answer could be "yes". Merchant and Honing (2014) suggest that this is caused because they lack the neural connections between auditory and motor systems; in fact, among primates humans are the only ones that exhibit this complex neural pathway (Egnor & Hauser, 2004). However, two pioneer studies investigated the rhythm issue bringing other results: Snowball — a sulphur-crested cockatoo (*Cacatua galerita eleonora*) — and Alex — an African grey parrot (*Psittacus erithacus*) — are skilled dancers. In the first study, Patel and colleagues (2009) played Snowball's favorite song (Backstreet Boys' *Everybody*) at different tempi (up to ±20 %) showing that the parrot bobbed his head synchronously. Similarly, Schachner and collaborators (2009) mentioned the Alex's ability to dance even to novel musical stimuli; moreover, by creating a database querying *Youtube* with "dance" and the animal name, the authors documented that only 14 species (on *Youtube*) are able to engage in rhythm. Despite a very limited set of other living organisms can synchronize to each other and taught to entrain rhythm (Greenfield & Schul, 2008; Cook et al., 2013), it is plausible the idea that rhythmic synchronization is due to the vocal learning ability (Patel, 2006).

1. Convergent Behaviours

Vocal learning refers to the ability to learn vocalizations that are not innate, but learned. For example, humans do not need to learn to cry (Fitch & Jarvis, 2013), but they need to learn language. Among vocal learner species there are some convergent behaviors. Above all, they clearly rely on the auditory channel as it provides both the source and the feedback to a process that basically entails a motor action. Animal "forbidden experiments" demonstrated that songbirds isolated from others develop aberrant songs and show a lower reproductive success (Marler, 1999; Williams et al., 1993). Learning has to occur in a particular temporal window, namely critical age, after which it is not possible to vocalize correctly. The (in)famous case of Genie (Fromkin et al., 1974), a feral child who has been forced to live the first 13 years of her life in complete social isolation and acquired a dramatically undeveloped language, provides evidence for the critical age in language acquisition. Although not vital, something similar to the critical period has been documented also for music: adults are better in verbal memory tasks if they received musical training before the age of 12 (Chan et al., 1998), and there is a strong correlation between the absolute pitch and the age of onset of musical training (Brown et al., 2002; for a critical view, see Levitin & Zatorre, 2003). The critical period is characterized by the babbling stage. Given the wide range of sounds producible by the human phonatory apparatus, the set of sounds used by a given language, and the ability for a baby to be sensitive to "those" sounds rather than others, babbling stage is essential for learning the relationship between oral movements and auditory outcomes (Patel, 2008), and it occurs — gesturally — also in deaf children exposed to sign language (Petitto & Marentette, 1991) and in songbirds (Doupe & Kuhl, 1999).

There is no need to mention here the syntactic and phonologic aspects of human language, but a syntax-like feature in nonhuman animal's songs has been documented. For example, Alex, the African grey parrot, was not only a good dancer, it also used combinatorial principles in the way that three-year-old human children do (Pepperberg, 1999). The humpback whales' (*Megaptera novaeangeliae*) songs present a hierarchical structure: notes are combined into phrases, which form themes and songs that are cyclically repeated (Payne, 2000). Appealing is the idea that nonhuman animal could handle human-specific recursive mechanisms for their vocalizations: Gentner and colleagues (2006) surprisingly found that European starlings (*Sturnus vulgaris*) use recursive center-embedded grammar; however, Corballis (2007) suggested that more than one level of center-embedded recursion is difficult to parse even for humans, thus the Starling's success in Gentner's study was probably due to discrimination of patterns rather than the use of whatever grammatical rule. Lastly, dialects — i.e., geographically defined and culturally transmitted songs — are very spread among vocal learners (e.g., 6,000 human languages). A well-known case in this field is that of humpback whales (Payne, 2000): their songs change in relation to different populations and time, from month to month, through improvisation and imitation, which suggests a human-like culture enforced by population agreement.

2. Neuroanatomy

Neuroanatomically speaking, a crucial role in the vocal learning ability is played by the basal ganglia, which links the auditory inputs to the motor outputs (Patel, 2006) and it is also important for sequential movements (Doupe & Kuhl, 1999) as well as for processing temporal patterns in both language and music (Grahn, 2012); moreover, the left hemisphere appears to be dominant for both language and songs (Fitch & Jarvis, 2013). Doupe and Kuhl (1999) reviewed that other structures are connected in this circuitry such as cerebellum, involved in rate and rhythm of vocalizations, and thalamus, for maintaining stable vocalizations. The vocal learning path-way, is often referred to as the cortical-basal ganglia-thalamic-cortical loop, emphasizing the circular nature of this path: from auditory to motor, from motor to auditory areas. Fitch and Jarvis (2013) stressed the importance of the connection between the motor cortical regions and the (brainstem) motor neurons involved in the control of the vocal organs (syrinx and larynx in birds and humans) that does not exist (or is very weak; Iriki, unpublished observation cited in Arbib & Iriki, 2013, p.493) in non-vocal learners; that is why other primates cannot finely control their orofacial movements, hence vocalizations.

Now, the question someone may raise is: do all vocal learners have a common ancestor? If it were, we would expect that all great apes share the vocal learning ability with us, but this is not the case. Yet, as Fitch and Jarvis (2013) show, vocal learners are sparsely distributed among species, and such an ability could have either gained or lost. Humans, songbirds and other vocal learners, suggested the authors, are too rare in animal kingdom, and it is unlikely that they all had a vocal learner common ancestor.

3. Genetics

One of the genetic components of the vocal learning ability is the gene called FOXP2. This gene has several downstream effects involving brain, lungs, heart, and gut (Vargha-Khadem et al., 2005), hence it is neither the gene for language, nor the gene for vocal learning ability, yet is still crucial for their appearance (Fitch & Jarvis, 2013). Despite mammals and birds split 300 million years ago, human FOXP2 differs from songbird FoxP2 in only 2 % of the protein. Only three amino acids changed since humans diverged from mice, around 70 million years ago, but two of them are present in only humans and not in other primates, suggesting that this gene has been fixed in humans around 4-6 million years ago, when we split from chimpanzees (Vargha-Khadem et al., 2005). The human version of this gene jumped out at the scientific community from the naturally mutated genome of the KE family, that reported throughout three generations developmental verbal dyspraxia in almost half population (Lai et al., 2001). Affected members were unable to produce fine orofacial movements during speech, they were blind to the inflectional features of English (luckily they did not live in ancient Rome), and displayed more general impairments in language production rather than comprehension (Vargha-Khadem et al., 2005); one might think that such impairments are related to motor areas, but patients shown no deficits in manual praxis. The gene's maximal expression is in the brain development: neural images from affected patients revealed that the volume of the caudate nuclei (a basal ganglia structure) correlated with performances in oral praxis, they had low levels of grey matter in Broca's area, and did not show lateralization when processing language.

Consistently, results from animal experiments show that genetic modification in the Area X (a basal ganglia structure necessary for song learning) of zebra finches resulted in imprecise and more variable (not stable) imitated songs (Haesler et al., 2007). More interesting, is the case of humanized mice that have been implanted with our version of FOXP2 (Enard et al., 2009). Ironically, during a Nobel conference speech, Svante Pääbo remarked that when the researchers were testing the results, they attempted to talk to the mice for six months, but no results jumped out. Again, although it might be funny, this is crucial to understand that FOXP2 is not directly related to language, but rather, along its downstream consequences, the gene regulates several domains: mice reported a qualitatively change in their vocalizations, a great modification in the basal-ganglia, and an induced robust long-term neuronal depression which is directly related to the synaptic plasticity, hence to the capacity for learning (Fisher & Scharff, 2009).

4. Vocalizations and Emotions

Emotions are deeply tied to music, and we all agree that there is no need to read statistics to understand it. However, a meta-analysis conducted by Tirovolas & Levitin (2011) revealed that such a topic has been one of the most investigated in the *Music Perception* journal between 1983 and 2010. The uses of music for "emotional" purposes are diverse: music can be used for changing, releasing, and matching emotions as well as for enjoying, comforting, and relieving stress (Juslin & Västfjäll, 2008), but how does music induce emotions? Juslin and Västfjäll (2008) proposed six psychological mechanisms by which emotions are aroused by music: by acoustical characteristics that signal a potentially important urgent event; by a repeated pairing with positive or negative stimuli; by mimicking an expression of an emotional state; by inducing visual images; by evoking particular memory associated with; and by violating, delaying, and confirming listener's expectations.

Taking the emotion of fear as an example of potentially important event, several studies investigated how this is processed in the brain when elicited by music: amygdala seems to play an important role. The patient S. M., with a bilateral damage restricted to both amygdalae, was selectively impaired to recognize sad and fearful music, she judged the peaceful music as less relaxing, and the scary one as less arousing (Gosselin et al., 2007). Koelsch and colleagues (2006) reported that also the processing of permanently dissonant music — i.e., negative emotional valence — occurs in amygdala, which additionally has a role in recognizing fearful human vocalizations (Dellacherie et al., 2011) and prosody (Scott et al., 1997). Interesting, during fearful musical stimuli, amygdala is functionally connected to the visual cortex suggesting an increase of visual alertness (Koelsch et al., 2013), that is, fight or flight.

In a similar vein, monkey studies (Adolph, 2010 and references therein) reported that amygdala lesions resulted in less caution towards a potential predator and less timidity towards human strangers; they lose their social status and are left dying in the wild. If the lesion occurs in the neonatal period, monkeys show diminished fear towards novel objects and a more prosocial behaviour. In rats, lesions to the amygdala resulted in the inability to acquire Pavlovian conditional vocalization responses (Borszcz & Leaton, 2003). Anatomically speaking, the generation of apes calls involves subcortical structures such as cingulate cortex, diencephalon, and brainstem structures (Corballis, 2010), while, on auditory perception, they show a human-like auditory lateralization (Poremba et al., 2004): species-specific vocalization activate the left superior temporal gyrus (also behaviourally reflected by the tendency to turn the right ear to the sound source) which is not elicited by other sound classes. Moreover, the superior temporal gyrus shows a specialization in regard to calls-type — anterolateral — and sound-source location — caudolateral — (Ghazanfar & Hauser, 2001). In humans a bilateral damage of the facial motor cortex causes a loss of voluntary control over speech, whereas in other primates there are no significant consequences on the production of calls, phenomenon that reflects differences in terms of innateness/ learnedness (Jürgens, 2002). So, brain injuries research suggests that amygdala is involved in processing emotional stimuli in both human and nonhuman primates. While the auditory perception shows similarities, the production shows a different pattern: human speech stems from cortical structures, whereas ape vocalizations stem from the subcortical ones. Paraphrasing these results, intentionality seems to play a crucial role.

Shared Intentionality

Shared intentionality is the rationale for human communication: both communicator and recipient understand others actions as intentionally motivated to communicate. Communicative acts work recursively:

(i) communicator's goal is that recipient knows something,

(ii) recipient understands that communicator wants her to know something,

(iii) communicator recognizes that recipient wants to understand and respond to his desire for her to know something (Tomasello, 2008, pp.204–5).

Generally, there are three basic communicative motives: requesting help, offering help, and sharing emotions (Tomasello, 2008); the latter is paramount for music. If nonhuman primates can, in some extraordinary cases, request and offer help (Warneken & Tomasello, 2006), what is clear is that they cannot align with conspecifics due to social conformity or solidarity. Conversely, humans have the need to be like others and be liked by others (Tomasello, 2008) in order to avoid isolation. One can align with others by sharing emotions and attitudes: for example, showing enthusiasm for painting is not just an informative declaration such as "I want you to know that I like painting", rather it is a means for increasing common ground and psychological closeness in a way that "I want you to feel something so that we can share attitudes/ feeling together" (Tomasello, 2008, p.87). This behavior can be ontogenetically traced when infants point to something displaying their emotions alternating gaze with adults: "Look! The cat is on the mat!". Liszkowski, Carpenter and Tomasello (2007) reported that when an adult expresses disinterest, the infant stops pointing to the object of her interest because she understands that her enthusiasm towards such object is not shared with the adult. In a large scale, sharing attitudes and feelings creates affiliation and social identity, alliances are stipulated on the basis of common interests, and groups become more cohesive.

1. Individual, Joint, and Collective Intentionality

Broadly speaking, according to Tomasello (2014), the evolutionary path of human thinking is characterized by individual, joint, and collective intentionality, roughly corresponding to think about myself, my partner, and the group to which I belong. Apes are individual animals, their thinking is based on individual intentionality that results in a competitive behavior. Although they are able in some cases to join in group activities (choosing the best cooperator, Melis et al., 2006b), apes still think about themselves: if they are given the choice to acquire or eat food alone or in group, they prefer the solitary option, and food conflicts are usually resolved towards the dominant ape.

In a rather different way, humans behave cooperatively. According to the interdependence hypothesis (Tomasello et al., 2012), the first step towards human cognition is represented by joint intentionality, that is, cooperation as a mutualistic collaboration. A change of ecology forced humans to compete in foraging with the expansion of terrestrial monkeys, pushing them in the direction of a dietary niche based on meat. Realizing that solitary hunting brings less chance of success, they started to cooperate in a dyadic form and started to share the spoils becoming more tolerant to each other (Melis et al., 2006a; see also Gibbons, 2014). As a consequence, they understood they were interdependent and began to think about the well-being of the partner for the sake of common — reflecting proper — interests.

According to Tomasello and colleagues (2012), with the growth of population within groups and the beginning of competition between groups, small-scale collaborative foraging did not suffice the new social environment and the need for food. New cognitive demands appeared: coordinating towards joint goals, mastering complex skills and technology, and transmitting them faithfully within and across generations. Large groups required more coordination with strangers, and also a kind of group selection in order to avoid free ridings; reputation played a crucial role. Hence, collective intentionality (Tomasello, 2014) drove humans to create social norms and cultural conventions in order to regulate the diversity of unproductive behaviors. Cultural conventions are needed where complex skills and technologies are required to be learned by a novice: a joint activity is more efficiently performed under conventions for the fact that both agents expect the partner to act in the same manner.

2. Recursive Mindreading

Despite challenged by the "*querelle* of recursion" blown-up from Everett (2005) and the supposed recursive patterns of birds (Gentner, 2006), a well-recognized uniqueness of human beings is the capacity for recursion (Hauser, Chomsky & Fitch, 2002; Corballis, 2011). Recursion refers to the computational mechanism by which, through hierarchical embedding, from a finite set of elements can virtually arise limitless sentences (e.g., "this is the cat that worried the rat that ate the malt that lay in the house that Jack built", Corballis, 2007, p.698). Besides recursive mechanisms, what humans are very good at is exploiting their social brain for understanding others intentions, which appears to be correlated with the social complexity (Dunbar, 1998; Gowlett et al., 2012). That is because the bigger is the group, the more cognitive demand is involved, such as recognizing individual's faces, their behaviors, their relationships, thinking about absent people and their role, and, indeed, understanding others intentions.

Understanding others intentions, also called "mindreading", involves a recursive mechanism: computers are zero-order intentional machines because they do not know they know, someone-order organisms do, and some others know that someone else knows; a third-order intentional organism knows that someone else knows that someone else knows, and so on. Humans think generally up to the fourth grade and rarely exceed the sixth as "Peter knows that Jane believes that Mark thinks that Paula wants Jake to suppose that Amelia intends to do something" (Dunbar, 1998, p.188). Communicative intentions are either third or fourth order because communicator does not want only that the recipient knows something, but she wants he knows that she wants he knows something in such a way as "I want you to know I want you to know [something]" (Tomasello, 2008, p.95). Outside such a recursive framework, even pantomimes could not have been arisen because at the same time the communicator gesticulates, attempting to convey a message, she is also communicating her intention to communicate — I want you to know I am pantomiming [so, I want you to know] X —, thus, without communicative intentions, a pantomime appears a bizarre action without any message conveyed.

3. Indexical and Iconic Gestures in Children and Apes

Although apes are not able to acquire new vocalizations, also through intensive training with humans, they can spontaneously learn to point to unreachable food. Captive monkeys do it only with humans because they are used to give monkeys free food, so they do not point in the wild because they do not expect such a cooperation from conspecifics. Apes, however, cannot pantomime and cannot understand iconic gestures (Tomasello et al., 1997). Because they lack of communicative intentions, they are not able to communicate iconic referents: even if an ape communicator could succeed in pantomiming, the ape recipient will not understand her because he lacks the recursive mindreading.

In stark contrast children are ready to point from the third month of age, but, having no communicative intentions yet, those gestures cannot be considered pointing as those of human adults. However, as soon as they understand intentions of others, at 10-14 months, they start pointing in order to request things and to share emotions, which appears to be a universal way to communicate at that age (Liszkowski et al., 2012). Pantomimes in the form of conventionalized gestures such as “OK” and “bye bye” arise much later in respect to pointing. Creative iconic gestures appear even later because they require the infant to have skills of imitation and simulation in order to represent something that is perceptually absent. Serving basically the same function, pantomimes are in competition with language: around the second year of age, when they start to acquire more language, infants prefer linguistic conventions to the iconic gestural ones — that show a decline in production.

Symbolized Iconic Indexes of Emotion Hypothesis

Summing up, given the similarities between music and language and the fact that great apes cannot entrain rhythm, it may be fair to think that music is a human species-specific activity, accompanied by language, that requires higher cognitive capacities. However, many linguistic and musical behaviors have been documented also in animal realm, showing neuroanatomical and genetic similarities. In terms of vocal expression of emotions, a great difference between humans and apes lays in the fact that humans can control their vocalizations, while apes cannot, suggesting the key role of intentional communication. Along human lineage, it could be hypothesized a change in the nature of intentionality, from individual to collective, wherein the ability for recursive mindreading seems at the basis of human communication. Consistent with comparative data on the emergence of indexical and iconic gestures in children and apes, it can be now put forward the hypothesis on the phylogenetic emergence of music stemmed from ape vocalizations.

The claim here is that music evolved in three main steps, involving an overlap (albeit to different extents) between indexes, icons, and symbols. In a first step, music originated from ape vocalizations as an index of an emotional state performed by an individual-intentionality agent; in a second step, joint-intentionality humans could escape from a mechanistic and unintentional reaction to an emotional state and could imitate expressions of emotions out from the here and now in an iconic form, which can be understood thanks to the recursive mindreading; finally, in big groups, collective-intentionality subjects conventionalized such vocalizations through the discretization of musical instruments in order to musically cooperate, and to transmit the musical heritage.

1. Indexes of Emotions

An index of emotions is not yet music, but is the first step through which it is possible to move towards the emergence of music. Apes vocalizations, a mechanical response to an emotional state, are indexes of emotions. Nonhuman primates are not vocal learners, that is, they have a genetically fixed repertoire of vocalizations that does not vary within the species and cannot be controlled (Tomasello, 2008). In fact, in stark contrast with songbirds and humans, monkeys raised in isolation or fostered by other monkey species still produce their own species vocalizations. What is crucial for the link with music is the fact that their vocalizations are deeply tied to an emotional state, and attempting to separate those vocalizations from a triggering emotion seems an impossible task (Goodall, 1986). Also alarm calls are not intended to alert other individuals about the imminent danger, but rather, again, they are an individualistic expression of emotions (Tomasello, 2008), and even though a macaque mother sees a predator approaching her offspring, she does not vocalize as long as she herself is in danger (Cheney & Seyfarth, 1990). Since their vocalizations cannot be voluntarily controlled (although squirrel monkeys seem to do it; Jürgens, 2002), apes can neither simulate an expression of an absent emotion, nor learn different ways to express an emotion; they simply cannot communicate absent referents (Liszkowski et al., 2009).

In a similar way, also modern humans display vocal indexes of emotions: laughing and crying, for example, are neither culturally learned nor separated from an emotion that is being experienced. Clearly, we express emotions also when we speak, for example, talking with vocal tremor during a public speech, loud voice when feeling gleeful, and high-pitched voice when greeting a sexually desirable person (Bachorowski, 1999). These expressions are not voluntarily controlled, but are the consequence of an emotional state (and physiological constraints) and are separate from language in a way that also nonsense sentences cue an emotional state (Scherer et al., 1991); these cues work also across different languages and cultures, suggesting an almost universal set of inference rules (Scherer et al., 2001). Although to a different extent, (cultural) indexes of emotions are present when using swearwords. Swearing is an expression of negative emotions (Rassin & Muris, 2005) and it has the power to increase pain tolerance and decrease pain perception (Stephens et al., 2009). As a matter of fact, swearing taps into the deep and emotional brain involving the limbic system and the right basal ganglia (Pinker, 2007) activating the fight or flight mechanism; hence it is probably rooted in ancient survival urges.

2. Iconic Indexes of Emotions

If indexicalizing emotions is the not-yet-music starting point, iconizing indexes of emotions is the first step towards the onset of music as we know it today. While indexes are deeply tied to the here and now, icons have the property to escape this condition because they depend on the communicator’s skills to imitate and simulate an entity that is not perceptually present, and, from the recipient side, on skills of imagination (Tomasello, 2014). The implementation of icons in a communicative system is consistent with the Mithen’s (2005) evolutionary hypothesis according to which the mimetic — i.e., iconic — factor attached to the holistic, manipulative, multi-modal, and musical communication system had marked the passage to the early-human musical proto-language ancestor. So, it can be supposed that early humans started to iconize emotions during joint activities (e.g., during foraging): a caveman may have pantomimed to his cave-mate an episode where a third peer had been previously injured by a tiger, pointing at the same time to the place where it happened; in doing so, he could have iconized the expression of emotion of the injured mate, exploiting the ability of the recipient to understand communicative acts recursively.

Communicating iconic vocal indexes of emotions is possible when an agent has communicative intentions in the way that "I musician want you listener to know that I musician am pantomiming an expression of emotion to you listener" (adapted from Tomasello, 2008). Because of this, the inability of apes to produce music is not due their lack of imagining (and to some extents they do; Tomasello, 2014), but rather to the lack of communicative intentions; they do not have the communicative motive as humans do, thus they are not interested in doing so with conspecifics (in particular, sharing emotions). Given the acoustic similarities between vocal expression of emotions and the musical ones (Juslin & Laukka, 2003), these iconic indexes of emotions do work for the reason that both communicators and recipients have the common ground based on production/ perception of emotions, which, also through music, appear to be universally expressed and recognized (Fritz et al., 2009). According to the Juslin-and-Västfjäll (2008) framework, iconic indexes of emotions may elicit an emotional contagion because the perception of voice-like aspects of music are mimicked internally. In this regard, Fritz and Koelsch (commentary in Juslin & Västfjäll, 2008) also noted that such mirroring occurs only during the perception of pleasant stimuli (Koelsch et al., 2006), suggesting that the emotional contagion occurs accompanied by the approach behavior, i.e., directing attention to others. Thus, understanding intentions of others plays a crucial role in music. Steinbeis and Koelsch (2009) presented atonal musical sequences to subjects with the cues that it was either written by a composer or generated by a computer: the cortical network for mental state attribution was activated solely when the music was supposed to be written by a composer. This, from the communicative side, narrows the definition of music to something made by humans with communicative intentions (Koelsch, 2011c).

Still, iconized indexes of emotions are not yet music rather they are vocalizations that, escaped from the here and now, can finally move towards a new form of communication system that is music as we know today.

3. Symbolized Iconic Indexes of Emotions

Symbolizing iconic indexes of emotion is the step required to have music. Once a sign (iconic vocalization) is being extrapolated from its causal source (emotional state), it can move towards new forms. In the field of experimental semiotics, Garrod and collaborators (2007) have shown that iconic signs tend to change their nature in the direction of symbols due to communicative interactions; precisely, as a consequence of the migration of structural complexity from the sign itself to the users' memory, the sign become simpler, more schematic, and converges to each other thanks to a process of grounding. Another study (Verhoef et al., 2014) has shown that, through iterated learning, phonetic and phonological combinatorial structures can also emerge using a slide whistle (a particular flute without holes but with a metal plunger that permits to modulate pitch in a given continuum of sounds like voice and trombone); this is crucial in linking musical instruments to combinatorial mechanisms.

As mentioned above, since at a certain point in history groups became larger (Tomasello et al., 2012), there had been new needs: coordinative and transmitive motives arisen, thus, natural gestures became collective linguistic conventions (Tomasello, 2014). In regard to music, transmitive and coordinative needs can be accomplished by conventionalizing iconic indexes of emotions by the means of musical instruments. With reference to transmission, a way to transform an idiosyncratic pantomime of an expression of emotions into a transmissible information is to decrease the structural complexity of the sign by restricting the infinite possibility of producible sounds that the vocal apparatus can produce; in other words, discretizing the vocal continuum into a finite set of sounds. In this regard, flutes appear to be the first attempt to conventionalize vocalizations. Indeed, the discovery of a flute dating 30,000 years (Conard et al., 2009) not only attests the presumed emergence of musical instruments (excluding that other organic perishable artifacts may be used; see e.g., Cross, (1999) in reference to Dauvois (1989) regarding stalactitic structures), but it proves that such process of conventionalization was already in use. With reference to coordination, lithophones permitted to punctuate and stress rhythm during collective rituals, by discretizing the continuum of time. Especially with the need for coordination with strangers in many activities, which would have required for the novice a certain amount of time and effort to learn arbitrarily conventionalized norms and proto) language, entraining in music activity would have made the process of group integration much easier for them through an alignment of emotional states. In fact, given the music's multiplicity of "aboutness" of meanings (floating intentionality), it was thus possible a consequence-free exploratory behavior (Cross, 1999) as a medium for maintaining social flexibility (Cross, 2004).

Lastly: why "symbolized", rather than "conventionalized", iconic indexes? As in a given language, phonemes constitute a finite set extrapolated from an infinite array of producible sounds, so it is for music: a finite set of sounds (notes) are cut from an infinite array of sounds. Now, each singular musical sound (vs. the holistic vocalization) is an autonomous entity, and as such it can be perceived and handled; the iconic sign (a set of notes) starts thus its separation from its mimicked source by moving towards the symbolic form, of which interpretation relies on cultural common ground (collective memory). For instance, while iconic indexes of emotions are an analogical representation of an emotional state, the symbolized ones, on the other hand, are digitalized representations, which entail a certain degree of arbitrariness and openness; as a matter of fact, the finite set of elements permits in any case a potential infinite amount of symbolic meanings. The openness given by musical instruments permits the expressions of emotions to be exaggerated eliciting even a more emotional responses (consistent with the "peak shift" principle by Ramachandran & Hirstein, 1999): if "angry" is vocally expressed by fast rate, loud intensity, and a harsh timbre, a musical instrument could be even faster, louder, and harsher (Juslin & Västfjäll, 2008). Without the ability to manage symbols, music could not be able to serve a plethora of uses such as group affiliation, ritual accompaniment, intermediary with divinities, associations to — and amplifications of — visual meanings, and so on.

Conclusions

Broadly speaking, music is an expression of an emotional state; narrowing, it is actually an imitation of an expression of an emotional state, but going in depth in the nowadays-form, music is a conventionalized imitation of an expression of an emotional state. That is because music is a cultural artifact — a technology —, and as such, for social needs it has to be constrained for coordinative and transmissive motives, a process wherein musical instruments played a crucial role. The power of music to arouse emotions is due to its iconic nature, which mimicking universally expressed emotions can be universally understood thanks to the human ability to recursively read the minds in a way that human communication involves the desire for communication in addition to the message itself. It has been proposed, taking intentionality as benchmark, that the change in the nature of expression of emotions by music parallels the change in the nature of intentionality: indexes, icons, and symbols are byproducts of individual, joint, and collective intentionality. Yet, this is not the whole story. If generally speaking the phylogenesis of music can be thought stemming from nonhuman primate unintentional vocalizations, it is also true that some remarkable biological necessaries are not shared with them, that is, above all the voluntary control over their vocalizations which is due to a particular neuroanatomical circuitry that enables the vocal learning pathway. In the absence of this ability it would not be possible to learn cultural vocalizations other than the innate ones, a process that exploits the neural link between auditory and motor brain areas, where a subcortical structure (basal ganglia) plays a role in processing sequential movements and temporal patterns. Vocal learners are sparsely distributed among species, suggesting a genetic convergence of the FOXP2 gene of which expression involves mainly the brain and lungs development, which are the biological conditions for language and music appearance — at least in the vocal form.

At this point, some interrogatives still exist. We know that litho-phones appeared in human lineage before flutes (respectively 40 and 30,000 years ago; Cross et al., 2002; Conard et al., 2009), and that apes spontaneously engage in "drumming" in the wild (Fitch, 2006); consistent with the gestural origin of language hypotheses (e.g., Tomasello, 2008; Arbib, 2005; Corballis, 2002), it is not to be excluded a motor — rather than vocal — hypothesis for the emergence of music. In this case, the emotional trigger should be reassessed rising other questions such as "when" and "why" emotions irrupted into music, and "how" and "why" the change to vocal medium took place. However, although apes have a good dexterity (Quallo et al., 2009), and despite rhythm can be considered a motor — rather than vocal — activity, in order to engage in rhythm they need the genetic equipment, therefore, also the motor origin of music hypothesis relies on the appearance of the vocal learning ability.

Finally, assuming that musical instruments triggered — or pushed — the symbolization of music, could it be possible that they also helped to bootstrap combinatorial mechanisms in language? Could it be the reverse (i.e., language compositionality helped music symbolization)? May these have arisen together as a consequence of a brain predisposed for such a mechanism? This line of questioning appears to revoke the egg/ chicken dilemma. However, it is now clear that to some extents symbolic behavior is traced very early in human history (Balter, 2009), much earlier than the appearance of musical instruments: taking apart the surprisingly symmetric hand-axes of the Oldowan industry (1.4 million years ago; Mithen, 2005), ornamental objects such as painted shells dating 82,000 years ago (Bouzouggar et al., 2007) and the geometric abstract representations on a pieces of ochre going back to 70,000 years ago (Henshilwood et al., 2002) attested in some ways the manifestation of symbolic behavior in *Homo sapiens*. More interesting is the discovery of the 35,000-year-old mammoth-ivory Venus of Hohle Fels (Conard, 2009), which appears to be the oldest example of figurative art. Being contemporary to the appearance of flutes and lithophones (±5,000 years) as well as the cave-paintings, this Venus attests the "creative explosion" (Balter, 2009) in humans, providing evidence for an already-modern human artistic and symbolic behavior in action.

Attempting to date the arrival of music is a pretentious endeavor. It has been found that Neanderthals shared with us the two changes that differ from human and chimpanzee in the FOXP2 gene (Krause et al., 2007), thus, it has been suggested that they had the capacity for language and speech as well as many other modern-human-like behaviors such as controlling fire, cooking food, burying their deaths, and so on (Dediu & Levinson, 2013). So, were Neanderthals musical beings, as Mithen (2005) suggested? Did we, both Neanderthals and us, inherit the musical gift from our common ancestor, as Falk (2004) suggested in relation to motherese as a trigger for language in late Australopithecus and early *homo sapiens*? Despite it is almost attested that Neanderthals had not musical instruments (d'Errico et al., 1998 rejected the bone-as-flute hypothesis of Turk 1997), it has been also suggested that they were not technological and cognitively much different from us, therefore the cultural (Villa & Roebroeks, 2014) and genetic exchange scenarios (Prüefer et al., 2014) are not to be rejected (it has to be noted that in outside-Africa humans, the Neanderthal genome varies between 2 % and 7.5 %; Stringer, 2012). Given that Neanderthals and modern humans overlapped for 2,600-5,400 years (Higham et al., 2014), it is not to be excluded also the fact that different musico-linguistic (e.g., musilanguage; Brown, 2000) communicative cultural technologies may have been swapped in that period (culture evolves faster than genes), "ratcheting" more elaborate communicative systems. Thus, it is not a hazard to reject the Prometheus — or Big Bang (see Corballis, 2011) — theory in favor of a more gradual hypothesis where language and music piggybacked each other providing means for implementing technologies that serve the needs for referencing and expressing emotions, from which there is no stepping back.

Concluding, being technologies, music and language could be conceived as "new machine[s] made out of old parts" (Tomasello, 2008, in reference to Bates, 1979) that rely on the capacity to read other's mind (Livingstone & Thompson, 2009). Without a recursive theory of mind, in fact, it would be not possible to manage iconic absent referents and so extrapolate indexes of emotion from the here and now, starting the process of symbolization due to cultural constraints. Moreover, the cultural recycling of cortical maps hypothesized by Dehaene and Cohen (2007) claims that a recent technology, too recent to be involved in evolution, is processed in the brain cortex entailing a structural modification of it invading older brain circuits; this is what Patel (2010) suggested for music too. These old parts constitute the modular processing of music (Peretz & Coltheart, 2003; Koelsch, 2011b) and they can also account for the selective impairments in music and language processing due to brain damages as well as for the fact that language and music — in regard to syntax — share neural resources but have different representations (Patel, 2003). This phenomenon can clarify the different purposes that language and music serve: referring to the outside world and referring to the inside world, i.e., expressing emotions.

References

1. Adolph, R., 2010. What Does the Amygdala Contribute to Social Cognition? Annals of the New York Academy of Sciences, 1191, pp.42–61.
   http://dx.doi.org/10.1111/j.1749-6632.2010.05445.x
2. Arbib, M. & Iriki, A., 2013. Evolving the Language- and Music-ready Brain. *In*: Language, Music, and the Brain, Arbib, M. A., Strüngmann Forum Reports, vol. 10, Lupp, J., (ed.) Cambridge, MA: MIT Press.
3. Arbib, M., 2005. From Monkey-like Action Recognition to a Human Language: An Evolutionary Framework for Neurolinguistics. Behavioral and Brain Sciences, 28(2), pp.105–167.
   http://dx.doi.org/10.1017/S0140525X05000038
4. Bachorowski, J. A., 1999. Vocal Expression and Perception of Emotion. Current Direct. Psychological Science, 8 (2), pp.53–57.
5. Balter, M., 2009. On the Origin of Art and Symbolism. Science, 323, pp.709–711.
   http://dx.doi.org/10.1126/science.323.5915.709
6. Bates, E., 1979. The Emergence of Symbols: Cognition and Communication in Infancy. New York: Academic Press.
7. Blood, A. J. & Zatorre, R. J., 2001. Intensely Pleasurable Responses to Music Correlate with Activity in Brain Regions Implicated in Reward and Emotion. Proceedings of the National Academy of Sciences of the United States of America, 98(20), pp.11818–11823.
   http://dx.doi.org/10.1073/pnas.191355898
8. Borszcz, G. S. & Leaton, R. N., 2003. The Effect of Amygdala Lesions on Conditional and Unconditional Vocalizations in Rats. Neurobiology of Learning and Memory, 79, pp.212–225.
   http://dx.doi.org/10.1016/S1074-7427(03)00002-9
9. Bouzouggar, A., Barton, N., Vanhaeren, M., d'Errico, F. Collcutt, S. et al., 2007. 82,000-year-Old Shell Beads from North Africa and Implications for the Origins of Modern Human Behavior. Proceedings of the National Academy of Sciences, 104(24), pp.9964–9969.
10. Brown, S. W., 2000. The "Musilanguage" Model of Music Evolution. *In*: The Origins of Music, Wallin, N. L., Merker, B. & Brown, S. W. (eds.). MIT Press, pp.271–300.
11. Brown, S. W. & Jordania, J., 2011. Universals in the World's Musics. Psychology of Music, 41(2), pp.229–48.
12. Brown, W. A., Sachs, H., Cammuso, K., & Folstein, S. E., 2002. Early Music Training and Absolute Pitch. Music Perception, 19, pp.595–597.
    http://dx.doi.org/10.1525/mp.2002.19.4.595
13. Chan, A. S., Ho, Y. C., Cheung, M. C., 1998. Music Training Improves Verbal Memory. Nature 396, pp.128.
    http://dx.doi.org/10.1038/24075
14. Cheney, D. L. & Seyfarth, R. M., 1990. Attending to Behaviour Versus Attending to Knowledge: Examining Monkeys' Attribution of Mental States. Animal Behaviour, 40, pp.742–753.
    http://dx.doi.org/10.1016/S0003-3472(05)80703-1
15. Conard, N. J., 2009. A Female Figurine from the Basal Aurignacian of Hohle Fels Cave in Southwestern Germany. Nature, 459, pp.248–252.
    http://dx.doi.org/10.1038/nature07995
16. Conard, N. J., Malina, M. & Munzel, S. C., 2009. New Flutes Document the Earliest Musical Tradition in Southwestern Germany. Nature, 460, pp.737–740.
17. Cook, P., Rouse, A., Wilson, M. & Reichmuth, C., 2013. A California Sea Lion (Zalophus californianus) Can Keep the Beat: Motor Entrainment to Rhythmic Auditory Stimuli in a Nonvocal Mimic. Journal of Comparative Psychology, vol. 127, No. 4, pp.412–427.
    http://dx.doi.org/10.1037/a0032345
18. Corballis, M. C., 2002. From Hand to Mouth: The Origins of Language. Princeton, NJ: Princeton University Press.
19. Corballis, M. C., 2007. Recursion, Language, and Starlings. Cognitive Science, 31, pp.697–704.
    http://dx.doi.org/10.1080/15326900701399947
20. Corballis, M. C., 2010. Mirror Neurons and the Evolution of Language. Brain and Language, 112, pp.25–35.
    http://dx.doi.org/10.1016/j.bandl.2009.02.002
21. Corballis, M. C., 2011. The Recursive Mind: The Origins of Human Language, Thought, and Civilization. Princeton, NJ: Princeton UP.
22. Cross, I., 1999. Is Music the Most Important Thing We Ever Did? Music, Development and Evolution. *In*: Music, Mind, and Science, Won Yi, S. (ed.), Seoul: Seoul National University Press, pp.10–39.
23. Cross, I., 2004. Music and Meaning, Ambiguity and Evolution. *In*: Musical communication, Miell, D., MacDonald, R. & Hargreaves, D. Oxford: Oxford University Press, pp.27–43.
24. Cross, I., Fitch, T., Aboitiz, F., Iriki, A., Jarvis, E. D. et al., 2013. Culture and Evolution. *In*: Arbib, M. A., Language, Music, and the Brain, Strüngmann Forum Reports, Lupp, J. (ed.), vol. 10, Cambridge, MA: MIT Press.
25. Cross, I., Zubrow, E., & Cowan, F., 2002. Musical Behaviours and the Archaeological Record: A Preliminary Study. *In*: Mathieu, J., Experimental Archaeology. British Archaeological Reports International Series, 1035, pp.25–34.
26. d'Errico, F., Villa, P., Pinto Llona, A. C. & Ruiz Idarraga, R., 1998. A Middle Palaeolithic Origin of Music? Using Cave-bear Bone Accumulations to Assess the Divje Babe I Bone "Flute." Antiquity, 72, pp.65–79.
27. Dauvois, M., 1989. Son et musique paléolithiques. Les Dossier d'Archéologie, 142, pp.2–11.
28. Dediu, D. & Levinson, S. C., 2013. On the Antiquity of Language: the Reinterpretation of Neandertal Linguistic Capacities and its Consequences. Frontiers in Psychology, 4, pp.1–17.
    http://dx.doi.org/10.3389/fpsyg.2013.00397
29. Dehaene, S. & Cohen, L., 2007. Cultural Recycling of Cortical Maps. Neuron, 56, pp.384–98.
    http://dx.doi.org/10.1016/j.neuron.2007.10.004
30. Dellacherie, D., Hasboun, D., Baulac, M., Belin, P. & Samson S., 2011. Impaired Recognition of Fear in Voice and Reduced Anxiety After Unilateral Temporal Lobe Resection. Neuropsychologia, 49, pp.618–629.
    http://dx.doi.org/10.1016/j.neuropsychologia.2010.11.008
31. Doupe, A. J., & Kuhl, P. K., 1999. Birdsong and Human Speech: Common Themes and Mechanisms. Annual Review of Neuroscience, 22, pp.567–631.
    http://dx.doi.org/10.1146/annurev.neuro.22.1.567
32. Dunbar, R. I. M., 1998. The Social Brain Hypothesis. Evolutionary Anthropology, 6, pp.178–190.
    http://dx.doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8
33. Egnor, S. E., & Hauser, M. D., 2004. A Paradox in the Evolution of Primate Vocal Learning. Trends in Neurosciences, 27, pp.649–654.
    http://dx.doi.org/10.1016/j.tins.2004.08.009

34. Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T. et al., 2009. A Humanized Version of Foxp2 Affects Cortico-Basal Ganglia Circuits in Mice. Cell, May 29, 137(5), pp.961–71.

35. Everett, D., 2005. Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language. Current Anthropology, 46, pp.621–646. http://dx.doi.org/10.1086/431525

36. Everett, D., 2012. Language: The Cultural Tool. London: Profile Books Ltd.

37. Falk, D., 2004. Prelinguistic Evolution in Early Hominins: Whence Motherese? Behavioral and Brain Sciences, 27, pp.491–503. http://dx.doi.org/10.1017/S0140525X04000111

38. Fisher, S. E. & Scharff, C., 2009. FOXP2 as a Molecular Window into Speech and Language. Trends in Genetics, 25 (4), pp.166–77. http://dx.doi.org/10.1016/j.tig.2009.03.002

39. Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., et al., 2001. Consonance and Dissonance of Musical Chords: Neuronal in Auditory Cortex of Monkeys and Humans. Journal of Neuro-physiology, 86, pp.271–278.

40. Fitch, W. T. & Jarvis, E. D., 2013. Birdsong and Other Animal Models for Human Speech, Song, and Vocal Learning. *In*: Arbib, M. A., Language, Music, and the Brain, Strüngmann Forum Reports, vol. 10, Lupp, J. (ed.) Cambridge, MA: MIT Press.

41. Fitch, W. T., 2006. The Biology and Evolution of Music: A Comparative Perspective. Cognition, 100, pp.173–215. http://dx.doi.org/10.1016/j.cognition.2005.11.009

42. Fritz, T., Jentschke, S., Gosselin, N., Sammler, D. et al., 2009. Universal Recognition of Three Basic Emotions in Music. Current Biology, 19(7), pp.573–576. http://dx.doi.org/10.1016/j.cub.2009.02.058

43. Fromkin, V., Krashen, S., Curtis, S., Rigler, D. & Rigler, M., 1974. The Development of Language in Genie: A Case of Language Acquisition Beyond the "Critical Period". Brain Lang, 1, pp.81–107. http://dx.doi.org/10.1016/0093-934X(74)90027-3

44. Garrod, S., Fay, N., Lee, J., Oberlander, J. & MacLeod, T., 2007. Foundations of Representation: Where Might Graphical Symbol Systems Come From? Cognitive Science, 31(6), pp.961–987. http://dx.doi.org/10.1080/03640210701703659

45. Gaser, C., & Schlaug, G., 2003. Brain Structures Differ Between Musicians and Non-musicians. Journal of Neuroscience, 23, pp.9240–9245.

46. Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C., 2006. Recursive Syntactic Pattern Learning by Songbirds. Nature, 440, pp.1204–1207. http://dx.doi.org/10.1038/nature04675

47. Ghazanfar, A. & Hauser, M. D., 2001. The Auditory Behaviour of Primates: A Neuroethological Perspective. Current Opinion in Neurobiology, 11, pp.712–720. http://dx.doi.org/10.1016/S0959-4388(01)00274-4

48. Gibbons, A., 2014. How We Tamed Ourselves—And Became Modern. Science, vol. 346, no. 6208, pp.405–406. http://dx.doi.org/10.1126/science.346.6208.405

49. Goodall, J., 1986. The Chimpanzees of Gombe. Patterns of Behavior. Cambridge, Mass.: Harvard University Press.

50. Gosselin, N., Peretz, I., Johnsen, E. & Adolphs, R., 2007. Amygdala Damage Impairs Emotion Recognition from Music. Neuropsychologia, vol. 45, pp.236–244. http://dx.doi.org/10.1016/j.neuropsychologia.2006.07.012

51. Gowlett, J., Gamble, C. & Dunbar, R., 2012. Human Evolution and the Archaeology of the Social Brain. Current Anthropology, vol. 53, no. 6, pp.693–722. http://dx.doi.org/10.1086/667994

52. Grahn, J., 2012. Neural Mechanisms of Rhythm Perception: Current Findings and Future Perspectives. Topics in Cognitive Science, pp.1–22.

53. Greenfield, M. D. & Schul, J., 2008. Mechanisms and Evolution of Synchronous Chorusing: Emergent Properties and Adaptive Functions in Neoconocephalus katydids (Orthoptera: Tettigoniidae). Journal of Comparative Psychology, 122, pp.289–297. http://dx.doi.org/10.1037/0735-7036.122.3.289

54. Haesler, S., Rochefort, C., Georgi, B. et al., 2007. Incomplete and Inaccurate Vocal Imitation After Knock-down of FoxP2 in Songbird Basal Ganglia Nucleus Area. X. PLoS Biology 5 (12), pp.2885-2897. http://dx.doi.org/10.1371/journal.pbio.0050321.

55. Hauser, M. D., & McDermott, J., 2003. The Evolution of the Music Faculty: A Comparative Perspective. Nature Neuroscience, 6, pp.663–668. http://dx.doi.org/10.1038/nn1080

56. Hauser, M. D., Chomsky, N., & Fitch, W. T., 2002. The Faculty of Language: What is it, Who Has It, and How Did It Evolve? Science, 298, pp.1569–1579. http://dx.doi.org/10.1126/science.298.5598.1569

57. Henshilwood, C., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., et al., 2002. Emergence of Modern Human Behavior: Middle Stone Age Engravings from South Africa. Science, 295, pp.1278–1280. http://dx.doi.org/10.1126/science.1067575

58. Higham, T., Douka, K., Wood, R., Ramsey, C. B., et al. 2014. The Timing and Spatiotemporal Patterning of Neanderthal Disappearance. Nature, 512, pp.306–309. http://dx.doi.org/10.1038/nature13621

59. Honing, H., Merchant, H., Háden, G. P., Prado, L. & Bartolo, R., 2012. Rhesus Monkeys (Macaca mulatta) Detect Rhythmic Groups in Music, But not the Beat. PLoS ONE 7 (12), pp.1–10. http://dx.doi.org/10.1371/journal.pone.0051369.

60. Izumi, A., 2000. Japanese Monkeys Perceive Sensory Consonance of Chords. Journal of the Acoustical Society of America, 108, pp.3073–3078. http://dx.doi.org/10.1121/1.1323461

61. Jürgens, U., 2002. Neural Pathways Underlying Vocal Control. Neuroscience and Biobehavioral Reviews 26, pp.235–58.

62. Juslin, P. N. & Laukka, P., 2003. Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? Psychological Bulletin 129, pp.770–814. http://dx.doi.org/10.1037/0033-2909.129.5.770

63. Juslin, P. N., & Västfjäll, D., 2008. Emotional Responses to Music: the Need to Consider Underlying Mechanisms. Behavioral and Brain Sciences, 31, pp.559–575. http://dx.doi.org/10.1017/S0140525X08005293

64. Kivy, P., 2002. Introduction to a Philosophy of Music. Oxford: Oxford University Press.

65. Koelsch, S., Kasper, E., Sammler, D., et al., 2004. Music, Language and Meaning: Brain Signatures of Semantic Processing. Nature Neuroscience, 7, pp.302–307. http://dx.doi.org/10.1038/nn1197

66. Koelsch, S., Fritz, T., von Cramon, D. Y., Muller, K., & Friederici, A. D., 2006. Investigating Emotion with Music: An fMRI Study. Human Brain Mapping, 27, pp.239–250. http://dx.doi.org/10.1002/hbm.20180

67. Koelsch, S., 2010. Towards a Neural Basis of Music-evoked Emotions. Trends in Cognitive Sciences, 14, pp.131–137. http://dx.doi.org/10.1016/j.tics.2010.01.002

68. Koelsch, S., 2011a. Towards a Neural Basis of Processing Musical Semantics. Physics of Life Reviews, 8(2), pp.89–105.

69. Koelsch, S., 2011b. Towards a Neural Basis of Music Perception – A Review and Updated Model. Frontiers in Psychology, 2(110), pp.1-15. http://dx.doi.org/10.3389/fpsyg.2011.00110.

70. Koelsch, S. 2011c. Transitional Zones of Meaning and Semantics in Music and Language Reply to Comments on "Towards a Neural Basis of Processing Musical Semantics". Physics of Life Reviews, 8(2), pp.125–128.

71. Koelsch, S., 2012. Brain and Music. Chichester: Wiley-Blackwell.

72. Koelsch, S., Fuermetz, J., Sack, U., Baur, K., Hohenadel, M., et al., 2011. Effects of Music Listening on Cortisol Levels and Propofol Consumption During Spinal Anesthesia. Frontiers in Psychology, 58, pp.1–9.

73. Koelsch, S., Skouras, S., Fritz, T., Herrera, P., Bonhage, C., Kussner, M., & Jacobs, A., 2013. The Roles of Superficial Amygdala and Auditory Cortex in Music-evoked Fear and Joy. NeuroImage, 81, pp.49–60. http://dx.doi.org/10.1016/j.neuroimage.2013.05.008

74. Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., et al., 2007. The Derived FOXP2 Variant of Modern Humans was Shared with Neandertals. Current Biology, 17, pp.1908–1912. http://dx.doi.org/10.1016/j.cub.2007.10.008

75. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P., 2001. A Forkhead-domain Gene is Mutated in a Severe Speech and Language Disorder. Nature, 413, pp.519–523. http://dx.doi.org/10.1038/35097076

76. Lerdahl, F. & Jackendoff, R., 1983. A Generative Theory of Tonal Music. Cambridge, MA: MIT Press.

77. Levitin, D. J. & Zatorre, R. J., 2003. On the Nature of Early Training and Absolute Pitch: A reply to Brown, Sachs, Cammuso and Foldstein. Music Perception 21, pp.105–110. http://dx.doi.org/10.1525/mp.2003.21.1.105

78. Levitin, D. J., 2006. This is Your Brain on Music: The Science of a Human Obsession. New York: Dutton.

79. Liszkowski, U., Carpenter, M., & Tomasello, M., 2007. Reference and Attitude in Infant Pointing. Journal of Child Language, 34, pp.1–20. http://dx.doi.org/10.1017/S0305000906007689

80. Liszkowski, U., Schäfer, M., Carpenter, M. & Tomasello, M., 2009. Prelinguistic Infants, But Not Chimpanzees, Communicate about Absent Entities. Psychological Science, 20, pp.654–660. http://dx.doi.org/10.1111/j.1467-9280.2009.02346.x

81. Liszkowski, U., Brown, P., Callaghan, T., Takada, A. & De Vos, C., 2012. A Prelinguistic Gestural Universal of Human Communication. Cognitive Science, 36, pp.698–713. http://dx.doi.org/10.1111/j.1551-6709.2011.01228.x

82. Livingstone, S. & Thompson, W. F., 2009. The Emergence of Music from the Theory of Mind. Musica Scientiae, Special issue 2009/ 10 "Music and Evolution", pp.83–115.

83. Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D., 2001. Musical Syntax is Processed in Broca's Area: An MEG study. Nature Neuroscience, 4(5), pp.540–545.

84. Marler, P., 1999. On Innateness: Are Sparrow Songs "Learned" or "Innate"? *In*: Hauser, M. D. & Konishi, M. The Design of Animal Communication. Cambridge, MA: MIT Press, pp.293–318.

85. McDermott, J., & Hauser, M., 2004. Are Consonant Intervals Music to Their Ears? Spontaneous Acoustic Preferences in a Nonhuman Primate. Cognition, 94 (2), pp.B11–B21. http://dx.doi.org/10.1016/j.cognition.2004.04.004

86. Melis, A., Hare, B. & Tomasello, M., 2006a. Engineering Cooperation in Chimpanzees: Tolerance Constraints on Cooperation. Animal Behaviour, 72, pp.275–286. http://dx.doi.org/10.1016/j.anbehav.2005.09.018

87. Melis, A., Hare, B. & Tomasello, M., 2006b. Chimpanzees Recruit the Best Collaborators. Science, 31, pp.1297–1300. http://dx.doi.org/10.1126/science.1123007

88. Merchant, H. & Honing, H., 2014. Are Non-human Primates Capable of Rhythmic Entrainment? Evidence for the Gradual Audiomotor Evolution Hypothesis. Frontiers in Psychology, 7 (274), pp.1–8. http://dx.doi.org/10.3389/fnins.2013.00274.

89. Miller, G. F., 2000. Evolution of Music Through Sexual Selection. *In*: Wallin, N. L., Merker, B. & Brown, S., The Origins of Music. Cambridge, Mass: The MIT Press.

90. Mithen, S., 2005. The Singing Neanderthals: The Origins of Music, Language, Mind and Body. London: Weidenfeld and Nicolson.

91. Painter, J. G., & Koelsch, S., 2011. Can Out-of-context Musical Sounds Convey Meaning? An ERP Study on the Processing of Meaning in Music. Psychophysiology, 48, pp.645–655. http://dx.doi.org/10.1111/j.1469-8986.2010.01134.x

92. Patel, A. D., 2003. Language, Music, Syntax and the Brain. Nature Neuroscience, 6, pp.674–681. http://dx.doi.org/10.1038/nn1082

93. Patel, A. D., 2006. Musical rhythm, linguistic rhythm, and human evolution. Music Perception, 24, pp.99–104. http://dx.doi.org/10.1525/mp.2006.24.1.99

94. Patel, A. D., 2008. Music, Language, and the Brain. New York, NY: Oxford University Press.

95. Patel, A. D., Iversen, J. R., Bregman, M. R. & Schulz, I., 2009. Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. Current Biology, 19, pp.827–830. http://dx.doi.org/10.1016/j.cub.2009.03.038

96. Patel, A. D., 2010. Music, Biological Evolution, and the Brain. *In*: Bailar, M., Emerging Disciplines. Houston, TX: Rice University Press, pp.91–144.

97. Payne, K., 2000. The Progressively Changing Songs of Humpback Whales: A Window on the Creative Process in a Wild Animal. *In*: Wallin, N.L., Merker, B. & Brown, S., The Origins of Music. Cambridge, Mass: The MIT Press.

98. Peirce, C. S., 1931/ 1958. Collected Papers of Charles Sanders Peirce. Hathorne, C., Weiss, P. & Burks, A. (eds), vols. 1–8. Cambridge, MA: Harvard University Press.

99. Pepperberg, I. M., 1999. The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots. Cambridge, MA, USA: Harvard Univ Press.

100. Peretz, I. & Coltheart, M., 2003. Modularity of Music Processing. Nature Neuroscience, 6, pp.688–691. http://dx.doi.org/10.1038/nn1083

101. Petitto, L. A., & Marentette, P. F., 1991. Babbling in the Manual Mode: Evidence for the Ontogeny of Language. Science, 251, pp.1493–1496. http://dx.doi.org/10.1126/science.2006424

102. Pinker, S., 1997. How the Mind Works. New York: Norton & Company.

103. Pinker, S., 2007. The Stuff of Thought: Language As a Window into Human Nature. New York: Viking.

104. Poremba, A., Malloy, M., Saunders, R. C., Carson, R. E., Herscovitch, P. & Mishkin, M., 2004. Species-specific Calls Evoke Asymmetric Activity in the Monkey's Temporal Poles. Nature, 427, pp.448–51. http://dx.doi.org/10.1038/nature02268

105. Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman et al., 2014. The Complete Genome Sequence of a Neanderthal from the Altai Mountains. Nature, 505, pp.43–49. http://dx.doi.org/10.1038/nature12886

106. Quallo, M. M., Price, C. J., Ueno, K., Asamizuya, T., Cheng, K. et al., 2009. Gray and White Matter Changes Associated with Tool-use Learning in Macaque Monkeys. Proceedings of the National Academy of Sciences, 106, pp.18379–18384. http://dx.doi.org/10.1073/pnas.0909751106

107. Ramachandran, V. S. & Hirstein, W., 1999. The Science of Art: A Neurological Theory of Aesthetic Experience. Journal of Consciousness Studies, 6, pp.15–51.

108. Rassin, E. & Muris, P., 2005. Why Do Women Swear? An Exploration of Reasons for and Perceived Efficacy of Swearing in Dutch Female Students. Personality and Individual Differences, 38, pp.1669–1674. http://dx.doi.org/10.1016/j.paid.2004.09.022

109. Roy, M., Lebuis, A., Hugueville, L., Peretz, I. & Rainville, P., 2012. Spinal Modulation of Nociception by Music. European Journal of Pain, vol. 16, pp.870–877. http://dx.doi.org/10.1002/j.1532-2149.2011.00030.x

110. Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S. et al., 2008. Music Listening Enhances Cognitive Recovery and Mood after Middle Cerebral Artery Stroke. Brain, vol. 131, pp.866–876. http://dx.doi.org/10.1093/brain/awn013

111. Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D., 2009. Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. Current Biology, 19, pp.831–836. http://dx.doi.org/10.1016/j.cub.2009.03.061

112. Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T., 1991. Vocal Cues in Emotion Encoding and Decoding. Motivation and Emotion, 15, pp.123–148. http://dx.doi.org/10.1007/BF00995674

113. Scherer, K. R., Banse, R., & Wallbott, H. G., 2001. Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. Journal of Cross-Cultural Psychology, 32, pp.76–92. http://dx.doi.org/10.1177/0022022101032001009

114. Schlaug, G., Jancke, L., Huang, Y., Staiger, J. E., & Steinmetz, H., 1995. Increased Corpus Callosum Size in Musicians. Neuropsychologia, 33, pp.1047–1055. http://dx.doi.org/10.1016/0028-3932(95)00045-5

115. Scott, S. K., Young, A. W., Calder, A. J., Hellawell, D. J., Aggleton, J. P., & Johnson, M., 1997. Impaired Auditory Recognition of Fear and Anger Following Bilateral Amygdala Lesions. Nature, 385(6613), pp.254–257. http://dx.doi.org/10.1038/385254a0

116. Sloboda, J., 1985. The Musical Mind: The Cognitive Psychology of Music. New York: Oxford University Press.

117. Steinbeis, N. & Koelsch, S., 2009. Understanding the Intentions Behind Man-made Products Elicits Neural Activity in Areas Dedicated to Mental State Attribution. Cerebral Cortex, 19(3), pp.619–623. http://dx.doi.org/10.1093/cercor/bhn110

118. Stephens, R., Atkins, J., & Kingston, A., 2009. Swearing as a Response to Pain. NeuroReport: For Rapid Communication of Neuroscience Research, 20, pp.1056–1060.

119. Stringer, C. B., 2012. What Makes a Modern Human. Nature, 485, pp.33–35.

120. Tirovolas, A. K., & Levitin, D. J., 2011. Music Perception and Cognition Research from 1983 to 2010: A Categorical and Bibliometric Analysis of Empirical Articles in Music Perception. Music Perception, 29(1), pp.23–36. http://dx.doi.org/10.1525/mp.2011.29.1.23

121. Tomasello, M., Call, J., & Gluckman, A., 1997. The Comprehension of Novel Communicative Signs by Apes and Human Children. Child Development, 68, pp.1067–1081. http://dx.doi.org/10.2307/1132292

122. Tomasello, M., 1999. The Cultural Origins of Human Cognition. Harvard University Press.

123. Tomasello, M., 2008. Origins of Human Communication. Cambridge, Mass: MIT Press.

124. Tomasello, M., Melis, A., Tennie, C., Wyman, E. & Herrmann, E., 2012. Two Key Steps in the Evolution of Human Cooperation: the Interdependence Hypothesis. Current Anthropology. 53(6), pp.673–692. http://dx.doi.org/10.1086/668207

125. Tomasello, M., 2014. A Natural History of Human Thinking. Cambridge, Mass: Harvard University Press.

126. Turk, I., 1997. Mousterion Boneflute and Other finds from Divje babe I cave site in Slovenia. Ljubljana: Institut za Arhaeologijo.

127. Vargha-Khadem, F., Gadian, D. G., Copp, A. & Mishkin, M., 2005. FOXP2 and the Neuroanatomy of Speech and Language. Nature Reviews Neuroscience, 6, pp.131–138. http://dx.doi.org/10.1038/nrn1605

128. Verhoef, T., Kirby, S. & de Boer, B., 2014. Emergence of Combinatorial Structure and Economy Through Iterated Learning with Continuous Acoustic Signals, Journal of Phonetics, 43, pp.57–68. http://dx.doi.org/10.1016/j.wocn.2014.02.005

129. Villa, P. & Roebroeks, W., 2014. Neandertal Demise: an Archaeological Analysis of the Modern Human Superiority Complex. PLoS ONE 9(4): e96424. http://dx.doi.org/10. 1371/journal.pone.0096424.

130. Warneken, F. & Tomasello, M., 2006. Altruistic Helping in Human Infants and Young Chimpanzees. Science, 31, pp.1301–1303. http://dx.doi.org/10.1126/science.1121448

131. Williams, H., Kilander, K. & Sotanski, M. L., 1993. Untutored Song, Reproductive Success and Song Learning. Animal Behavior, 45, pp.695–705. http://dx.doi.org/10.1006/anbe.1993.1084

Alessandro Miani

Portretinių (ikoninių) indeksų simbolizavimas: intencionalumu paremta hipotezė apie muzikos buvimą

Santrauka

Labiausiai pripažintas muzikos bruožas yra jos gebėjimas sužadinti klausytojų emocijas: ar tokios emocijos būdingos muzikai (indeksams), ar jos sužadinamos dėl muzikos panašumo į analogišką emocinę raišką (portretai, ikonos)? Teigiame, kad muzika yra konvencionalizuotas emocinio būvio raiškos imitavimas (t. y. simbolizuotas portretinis (ikoninis) indeksas), bet norint pasiekti tokią formą, reikia trijų evoliucinių žmogaus mąstymo žingsnių: individualaus, jungtinio ir kolektyvinio intencionalumo, kurie sugretina arba beveik sutampa su indeksais, ikonomis ir simboliais. Pateikdami palyginamąjį genetinių ir neuroanatominių infrastruktūrų, reikalingų muzikos buvimui, vaizdą, pateikiama hipotezė, kad (i) muzikos yra ir bezdžionės vokalinėje raiškoje, kaip rodiklis emocijų, kurias sukelia individo intencionalumo veiksniai; (ii) gebėdama rekursyviai skaityti bendro intencionalumo veikėjų mintis, muzika evoliucionavo kaip ikonizuotas emocijų rodiklis; (iii) kaip demografinių pokyčių pasekmė, kolektyvinio intencionalumo veiksniai sukūrė muzikos instrumentus, siekdami sumažinti struktūrinį ženklo sudėtingumą, turėdami tikslą koordinuoti bendras muzikines veiklas bei kultūriškai perduoti simbolizuotus ikoninius emocijų indeksus.

Straipsnis įteiktas 2014 11
Parengtas spaudai 2014 12

About the author

Alessandro Miani, a Master student in Cognitive Semiotics.
Primary research interest: structural relationship between syntax and semantics in music and language from a cognitive point of view; pragmatic differences between language and music; evolution of music; animal cognition; biology of language.
Address: Institut for Æstetik og Kommunikation - Center for Semiotik, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark.
E-Mail: ale.miani@yahoo.it