



Music Evolution in the Laboratory: Cultural Transmission Meets Neurophysiology

Massimo Lumaca^{1*}, Andrea Ravignani^{2,3,4} and Giosuè Baggio⁵

¹ Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Aarhus, Denmark, ² Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium, ³ Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands, ⁴ Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, ⁵ Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology, Trondheim, Norway

OPEN ACCESS

Edited by:

Aleksey Nikolsky,
Independent Researcher, Los
Angeles, CA, United States

Reviewed by:

McNeel Gordon Jantzen,
Western Washington University,
United States
Laura Verga,
Maastricht University, Netherlands
Vera Kempe,
Abertay University, United Kingdom
Seana Coulson,
University of California, San Diego,
United States

*Correspondence:

Massimo Lumaca
massimo.lumaca@gmail.com

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 21 September 2017

Accepted: 29 March 2018

Published: 16 April 2018

Citation:

Lumaca M, Ravignani A and Baggio G
(2018) Music Evolution in the
Laboratory: Cultural Transmission
Meets Neurophysiology.
Front. Neurosci. 12:246.
doi: 10.3389/fnins.2018.00246

In recent years, there has been renewed interest in the biological and cultural evolution of music, and specifically in the role played by perceptual and cognitive factors in shaping core features of musical systems, such as melody, harmony, and rhythm. One proposal originates in the language sciences. It holds that aspects of musical systems evolve by adapting gradually, in the course of successive generations, to the structural and functional characteristics of the sensory and memory systems of learners and “users” of music. This hypothesis has found initial support in laboratory experiments on music transmission. In this article, we first review some of the most important theoretical and empirical contributions to the field of music evolution. Next, we identify a major current limitation of these studies, i.e., the lack of direct *neural* support for the hypothesis of cognitive adaptation. Finally, we discuss a recent experiment in which this issue was addressed by using event-related potentials (ERPs). We suggest that the introduction of neurophysiology in cultural transmission research may provide novel insights on the micro-evolutionary origins of forms of variation observed in cultural systems.

Keywords: cultural transmission, diffusion chains, signaling games, iterated learning, music universals, music diversity, neural predictors, mismatch negativity (MMN)

INTRODUCTION

There has recently been a surge of interest in the biological and cultural origins, and evolution of music (Wallin et al., 2001; McDermott and Hauser, 2005; Patel, 2010). Music is prominent in virtually all human societies, and in its most sophisticated versions it is only attested in humans. This fact raises two important questions: how did music originate? And how did it evolve in its current forms? One intriguing issue here, especially in relation to the cognitive and neural bases of music evolution (Honing et al., 2015), is that of the *evolution of musical structure*. Musical systems are structured at several levels, from melody and harmony to rhythm and composition, in ways that may resemble the organization of other human generative systems, such as language (Jackendoff and Lerdahl, 1983; Jackendoff, 2009). The analogy between language and music may be pushed further, if one considers aspects of music that may be understood “semantically.” Listening to music can evoke a wide range of extra-musical experiences, from emotional feelings (e.g., the sadness suggested by Albinoni’s Adagio in G minor) to the mental imagery of specific referents

(e.g., characters or ideas in Wagnerian *Leitmotifs*) (Patel, 2010). Musical structures can and often do relate to a world of possible experiences and non-musical phenomena (Lerdahl, 2003) *expressively* (by being associated to internal affective states, e.g., emotional qualities), if not *representationally* (via relations of reference and truth, as language does) (Patel, 2010).

In this work, we focus on the cultural origins of musical syntax: the set of principles governing the combination of melody and rhythm into “well-formed” sequences (for a discussion on the evolution of semantic structures see Lumaca and Baggio, 2017, 2018; Ravignani and Verhoef, 2018). Some aspects of musical syntax, such as the organization of temporal structure and pitch intervals, display widespread distribution and striking cross-cultural similarities. For example, the tendency to use small intervals in non-polyphonic melodic phrases, or “proximity,” has been observed across several musical traditions of the world, including indigenous tunes from North America, Europe, and Asia (Dowling, 1968; Von Hippel, 2000). Despite some exceptions, such as Scandinavian and Swiss yodeling music, proximity is a prominent feature of melodic structure. These shared attributes are known as “musical universals.” Nevertheless, their form and frequency differ across and within different musical traditions of the world (Lomax, 1977; Rzeszutek et al., 2012; Savage et al., 2015). How can we explain both the invariance and the variation of structure in music? Which processes underlie the cross-cultural convergence toward common music traits or their diversification? In this paper, we suggest that neuroscience can provide critical methodological and theoretical tools for testing and generating hypotheses on this complex matter.

This article is organized as follows. We start by presenting a recent theoretical perspective in which music is understood as an evolving cultural system, adapting to the human brain [sections Linking Biological and Cultural Levels of Analysis and From Cultural Transmission to Neurophysiology (and Back)]. In section The Cognitive Level: Diffusion Chains and the Evolution of Musical Regularities in the Lab, we describe studies that support this view using data from behavioral experiments. In section The Neural Level: Constraints Imposed by a Neuronal Niche Drive the Emergence of Regularities, we transpose our analysis of cultural adaptation to the neural level. Partly using the “neuronal recycling hypothesis” as a theoretical framework (Dehaene and Cohen, 2007), we argue that music can adapt to a “neuronal niche” defined by the specific information processing constraints imposed by neural circuits originally evolved for auditory streaming.

To our knowledge, no one until recently has investigated this hypothesis by means of brain imaging or neurophysiology. In section Neural Predictors in Cultural Evolution Research, we describe a recent experiment in which this hypothesis was tested combining behavioral and neurophysiological methods. Finally (section The Neural Origins of Cultural Variation), we suggest that the introduction of concepts and methods from neuroscience in music evolution, and cultural evolution in general, can provide new insights on the process of cultural variation.

Linking Biological and Cultural Levels of Analysis

Music may be seen as a complex adaptive system, shaped by various biological, environmental, and cultural factors. This has made it difficult for musicologists and cognitive scientists to analyze the evolutionary origins of musical structure. The predominant view during the last century was the *cultural account*, where music was seen as an entirely socio-cultural construct, free to vary with virtually no biological and environmental constraints on its structure and content (Nettl, 1983; Repp, 1991; Blacking et al., 1995). The striking diversity of musical forms, as attested across and within cultures, and over human history, seems to support this notion (Lomax, 1968; Henry, 1976). Yet, this account has been challenged by experiments in psychology and neuroscience, together supporting a broadly *biological account* of the origins of music. Several studies point to the existence of perceptuo-cognitive biases and constraints in music processing and production (e.g., Trehub, 2000; Drake and Bertrand, 2001; Zatorre, 2001; Peretz and Zatorre, 2005; Deutsch, 2012) with some parallels in other species (Fitch, 2015). On this view, prototypical properties of music, such as a relatively steady beat, smooth melodic contours, tonality, and a narrow distance between adjacent tones (or “pitch proximity”), derive from built-in functional properties of the brain (McDermott and Hauser, 2005), which tend to manifest themselves in most human cultures (Lerdahl, 1992; Savage et al., 2015).

A recent view is that neither the “cultural account” nor the “biological account” can independently provide a satisfactory theory of the origins and evolution of musical structure (Trainor, 2015). Cultural accounts typically focus on the evolution of *musical systems*, while biological accounts investigate the evolution of the *human capacity to perceive, appreciate, and produce music* (also including musicality; Honing et al., 2015). These different accounts, however, may be connected within a more complete explanatory framework, if one accepts that music is neither an entirely arbitrary cultural construct nor strictly a biological product. Much like natural language, music is a cultural construct, which nonetheless rests upon, and is partly shaped by, human neurobiology. Our neurobiological makeup determines the scope and constraints of human auditory memory capacity, hierarchical sequence processing, attention, perceptual hearing threshold, and auditory scene analysis (Snyder, 2008; Deutsch, 2012). This is now a central tenet in the field of music cognition, and it is becoming increasingly accepted in cultural analyses of music, too. The open question is *how* neurobiological capacities, biases and constraints manifest themselves in actual musical systems (Trainor, 2015).

From Cultural Transmission to Neurophysiology (and Back)

Answering this question requires theories, models, and empirical data that can effectively bridge the gap between the classical chasms of (cultural) evolutionary science: between individual-level and population-level processes, micro-evolutionary and macro-evolutionary processes (Mesoudi, 2011). Specifically, one

important question is how the individual's neurobiological endowment manifests itself in music at the population level. This issue was already known in linguistics as the "problem of linkage" (Kirby, 1999). A possible answer is "through cultural transmission." Music, much like language, is not only a richly structured symbolic system, but also a set of behaviors that is maintained over time by intergenerational transmission (Morley, 2013; Le Bomin et al., 2016).

During intergenerational transmission, cultural information must survive a "memory bottleneck" (Deacon, 1997): the set of all neurobiological biases or constraints that bind our capacity to infer (and store) the "rules" that govern a system of information¹. The properties of the cultural system that fit best the human neurobiological filter—e.g., those that make information easier to process, encode, and recall—will have greater likelihood of being passed on to the next generation. If this view is correct, in the long run the neurobiological endowment of individuals should be reflected in the musical corpus at the population-level.

This view of transmission, emphasizing *adaptation of fast-changing cultural systems to a largely stable neurocognitive architecture*, was developed in evolutionary linguistics to account for the emergence of structure in human languages, including putative linguistic universals (Christiansen and Chater, 2008). Recent methodological advances (Mesoudi, 2015; Edmiston et al., 2018) have provided support for this view in controlled laboratory conditions. In most experiments, groups of individuals engage in simple, controlled forms of knowledge transmission, for example from a participant (a sender) to another (a receiver), along a diffusion chain. Each participant represents a "generation," and each interaction between participants allows for the passage of information across generations (Esper, 1925; Bartlett, 1932). The set of items transmitted along a diffusion chain (e.g., linguistic or musical phrases) is a finite sample drawn from the (infinite) set of items that learners have to generalize from. A challenge for research on cultural transmission is to show that core properties of the artificial systems being transmitted are also properties of the actual cultural systems being modeled and that the mechanisms at work in artificial conditions are also at work in real cultural evolution. In a landmark study, Kirby et al. (2008) showed how miniature "languages" emerge in the course of transmission from initial random associations of signals and meanings. When these pairings are transmitted across "generations" of participants, some regularities emerge, including *compositionality* (Hockett, 1960), as observed in human language. This result supports the view that core properties of language can be explained by the interplay of individual cognitive biases (sensu Brighton et al., 2005) and iterated cultural learning and transmission. Recent studies on animal models of cultural learning further support

¹Our definition of "memory bottleneck" includes constraints on perceptual grouping; capacity and temporal limits of auditory memory, serial processing, and attention; constraints on the neurodynamics of the auditory system; perceptual hearing thresholds. We limited this list to constraints "directly" related to basic aspects of perception and cognition. We acknowledge that constraints of a different nature might have a formative power over musical structures (e.g., motoric, motoric-expressive, physiological, cross-modal, and semantics).

this conclusion (e.g., for non-human primates see Claidière et al., 2014; for a seminal study on zebra finches see Fehér et al., 2009).

One way to start bridging this gap in the musical domain, is to assume that music, like language, is a complex adaptive cultural system, shaped for thousands of years by cycles of transmission, acquisition, and use (Morley, 2013). Following this view, neurobiological biases and constraints, as discussed above, brought out through cultural transmission, would exert effects on the form and structure of music (Merker et al., 2015; Trainor, 2015; Mehr et al., 2018). This mechanism could explain some properties of temporal (rhythm, meter) and spectral (melody, harmony) dimensions of musical structure, which are likely to be the result of *adaptations* to the combined pressures of neural constraints and various socio-cultural forces (Merker, 2006; Merker et al., 2015; Trainor, 2015). This would in principle apply to both *invariants*—putative cultural *universals* shared by musical systems or traditions (Savage et al., 2015)—and *variation* among individuals, generations, and traditions.

This point is not new. Lévi-Strauss (1960) had already observed that some structural regularities observed across cultures (e.g., the fact that symbolic material tends to be organized in binary oppositions) are reflections of principles of brain organization. Therefore, neuroscience is expected to contribute to explanations of the emergence and evolution of structural regularities, including their convergence and diversity. However, to date this issue has been addressed only by behavioral studies, and only to explain some invariant aspects of musical structure. In the next section, we summarize three of these lines of experimental work in the field of music evolution.

THE COGNITIVE LEVEL: DIFFUSION CHAINS AND THE EVOLUTION OF MUSICAL REGULARITIES IN THE LAB

In recent experiments, a diffusion chain method was used to study how music evolves in the lab (Ravignani et al., 2016). This study aimed to test whether human psychological biases, amplified by cultural transmission, can explain the emergence of rhythmic universals (Trehub, 2015). In this experiment, participants were given a drumstick and an electronic drum pad. Participants in the first generation listened to 32 randomly generated, hence a-rhythmic, patterns of beats (the input), and were asked to reproduce each of them to the best of their abilities (the output). The "imperfect" output produced by this first generation of participants became the input for the next generation, whose task was to perform the rhythm they heard, and so on, along a diffusion chain. This paradigm is known as "iterated learning" (IL) (Kirby et al., 2008). Given the difficulty to memorize these patterns, errors were introduced in the emerging system of drumming sequences, slightly modifying the original patterns at each generation. Across generations, patterns became increasingly structured and easier to learn. After 8 generations, at the end of each diffusion chain, patterns showed regularities similar to those found across musical traditions of the world. These universal rhythmic regularities included a tendency toward small integer ratios (e.g., 1:1 and 2:1) of intervals between beat

durations, and a relatively steady beat, also termed “isochrony” (Savage et al., 2015). This study represents the very first attempt to “grow” musical universals in the lab (Fitch, 2017), and sheds light on the cognitive and cultural mechanisms underlying the creation and vertical transmission of music (Le Bomin et al., 2016).

An IL study by Verhoef (2012) investigated the cultural evolution of combinatorial structures in musical systems. Participants were first exposed to a set of 12 whistles that they had to imitate immediately after listening by using a slide whistle (training phase). Next, they were asked to reproduce the whole set of signals as they remembered it (recall phase). The sequences generated by a participant were used to train the next one in the diffusion chain, and so on, until the end of the chain. In the course of transmission, structural regularities emerged, as predicted by previous computer simulations (de Boer, 2000). In the last generations, fewer discrete units were reused by individuals in concatenations, repetitions, or mirror forms to produce the entire vocabulary of whistles. Combinatoriality is a “design feature” of human language (Hockett, 1960) and it applies to musical structure, too. For instance, the authors observed that two distinct whistles were often combined into a single pattern by the next generation of individuals. Also, participants tended to produce mirror forms out of single patterns, so that more elements were shared between signals of the same set. With fewer units to memorize, organized in this manner, the set of signals was more structured, more compressed, and easier to learn and reproduce.

A more recent attempt to study music evolution in the lab is the work by Lumaca and Baggio (2017). The authors used a different model of cultural transmission than IL: multi-generational signaling games (MGSGs) (Moreno and Baggio, 2015; Nowak and Baggio, 2016). MGSGs are in essence an iterated variant of signaling games (Lewis, 1969; Skyrms, 2010) that combine basic aspects of semiotic models of coordination and communication (e.g., horizontal transmission; Galantucci and Garrod, 2011) with the intergenerational transmission of IL (Kirby et al., 2008). Two-person signaling games were organized in diffusion chains of 8 generations each. In each game, the sender and receiver were expected to converge, through repeated interactions, on a common code: a signaling system where 5 isochronous melodic riffs were associated to basic or compound emotions. This design can contribute to model different aspects of music transmission: first, a degree of alignment of internal states between musical senders (e.g. composers) and receivers (e.g., an audience) at two main levels, the *structural* and *affective* (Temperley, 2004; Bharucha et al., 2011); second, a partial asymmetry in information flow from senders to receivers, which is present in language and music transmission (e.g., from composers to listeners, from teachers to pupils, etc.). In each signaling trial, the sender was presented on the screen with one of the 5 equiprobable emotions (visualized as human facial expressions) and was asked to compose a 5-note isochronous riff on the computer keyboard. The receiver, after he listened to the riff via headphones, was asked to choose one of the 5 expressive faces displayed on the screen (i.e., the one possibly seen by the sender). A feedback was then presented simultaneously to both

participants’ screens, showing the expressive face seen by the sender and the one chosen by the receiver for the same melodic signal. This procedure was repeated at each successive trial. At the end of the game, the receiver (generation n) became the sender in the next game, with the same structure and a new participant as a receiver (generation $n + 1$), and so on, until the chain was completed. Senders were always asked to transmit the code they had learned in the previous game. Therefore, recall errors in the melodic signals (possibly “innovations”) were introduced throughout the experiment. The authors observed the gradual evolution over generations of several structural features of musical phrases: pitch proximity and continuity, symmetry, and motivic structure.

Despite differences in their assumptions and methods, those three experiments have reached similar conclusions: the immediate effects of psychological constraints on the musical systems may be *weak*, but they are *amplified* in the course of inter-generational transmission (Boyd and Richerson, 1988; Kalish et al., 2007; Kirby et al., 2007; Thompson et al., 2016) or iterated reproduction (Jacoby and McDermott, 2017), leading the evolution of musical structures along non-random paths. If principles of auditory organization and memory constraints operate in similar ways also in the production and perception of actual music, they could similarly shape the evolution of historical systems in the course of iterated transmission. Convergence toward some of the musical structures found across populations (Savage et al., 2015) could be then explained, to some extent, by adaptation to a special niche, constituted by a restricted set of low-level perceptual and memory processes. In the rest of the paper we will refer to this special niche as “neuronal niche” (Dehaene and Cohen, 2007).

THE NEURAL LEVEL: CONSTRAINTS IMPOSED BY A NEURONAL NICHE DRIVE THE EMERGENCE OF REGULARITIES

In recent years, there has been an increasing interest in how the brain accommodates and shapes novel cultural symbolic systems (Dehaene and Cohen, 2007). A leading hypothesis is that some cortical circuits, initially evolved as a result of specific selective pressures, are later “recycled” to accommodate novel cultural functions (Dehaene and Cohen, 2007; Simon et al., 2013; Dehaene et al., 2015; Skeide et al., 2017). Therefore, the acquisition of novel functions is constrained, however weakly, by prior human evolution. Once “culturally recycled,” pre-existing systems and mechanisms maintain some of their original capacities and limitations, providing a neuronal niche within which culture may adapt and evolve. This also means that the variability observed in cultural systems is limited by brain structure and function across individuals and groups.

If this hypothesis is correct, near-universal characteristics of music (Savage et al., 2015) may be traced back to the computational infrastructure of human auditory cortex and other (e.g., motor, attentional etc.) areas of the brain. Trainor (2015) related the origins of certain invariant musical features as adaptations to bottom-up neural mechanisms of auditory

scene analysis (ASA), such as the sequential sound segregation and integration of within-stream elements (Bregman, 1994). These specific mechanisms have evolved specifically to detect and localize multiple sources of auditory objects and to extract regularities from the acoustic environment. They often involve the perceptual grouping of single-event auditory stimuli into auditory streams and operate following Gestalt principles of proximity, similarity, and continuity (Deutsch, 1999). They are automatic (pre-attentive), they emerge early in human development (Demany, 1982; Winkler et al., 2003), and they are widely conserved across species (Fay, 2008). This point shows that the ASA neural circuitry is likely phylogenetically older than human music. Thus, the exaptation (or evolutionary re-use) (Gould and Vrba, 1982) of this more ancient biological mechanism by music should impose constraints on the way music is stored and organized in the brain, and accordingly, on the way it is recalled during transmission. In this regard, perceptual and memory recall advantages have been reported for tone streams that conform to Gestalt principles of organization (Bendixen et al., 2010; Loui, 2012; Rohrmeier and Cross, 2013). The cross-cultural tendency to organize music following these principles (Huron, 2001), in addition to the findings reported by cultural transmission research (Verhoef, 2012; Ravignani et al., 2016; Lumaca and Baggio, 2017), may support the idea that the neurocomputational constraints of the human auditory system constitute a filter through which musical material must pass, adapt, and eventually evolve.

It is surprising that up until recently, no one has attempted to find (counter-) evidence of cultural adaptation using neural measures. Research has shown that even recently-encoded information is shaped by perceptual or memory constraints into more compressed and abstract forms (Tamariz, 2017). Yet, the neural mechanisms underlying this phenomenon remain unknown. One reason is arguably our limited understanding of how information is represented in the brain (Mesoudi et al., 2006). Current whole-brain methods, such as functional magnetic resonance (fMRI), are not well-suited to investigate the precise basis of mental representations (but see Haynes and Rees, 2006; Johnson and Johnson, 2014; Zadbood et al., 2017). Another issue is to establish a link between neural constraints on learning—neural activity underlying specific, fast, and accurate encoding processes (Sadler et al., 2014)—and cultural adaptation. Electrophysiological methods, such as multi-unit recordings, seem ideal for this purpose, but they are too invasive to be performed on healthy individuals. Various animal models of social learning—in songbirds, primates, and other species—have provided useful information in this respect (Araki et al., 2016; Gadagkar et al., 2016; Tchernichovski and Lipkind, 2016). None of these species possesses cultural behaviors as rich and complex as human music. However, some of their behaviors exhibit structured patterns, which are maintained over time through inter-generational transmission. Cultural transmission, in turn, can shape animal vocal behavior so as to fit species-specific learning constraints (Fehér et al., 2009; Fitch, 2009).

The application of techniques and models used in language evolution allow researchers of animal behavior to explore the biology of culturally transmitted systems in simpler and more

controlled conditions, and to answer questions about cultural adaptation that cannot be *directly* answered in humans using current methods (but see next section for *indirect* answers). For example, Araki et al. (2016) used cellular recordings to demonstrate the existence in zebra finches of constraints on neuronal temporal coding that limit song acquisition to certain species-specific temporal features. Juvenile birds acquire their songs by imitating adult tutors. Although zebra finches are not bound to learn only specific sequences, they do show significant consistencies in their vocal repertoires (Lachlan et al., 2016). Do these consistencies result from adaptation of song material to the zebra finch neural constraints on learning? Araki et al. (2016) found that a subset of neurons in the zebra finch auditory cortex responds synchronously and selectively to patterns of inter-syllable silent gap durations, which are typical of their songs. The same cell population was unresponsive to other species' songs. Temporal coding mechanism like this are thought to preserve the species-specific song identity from any random drifts that may be introduced during cultural transmission.

Critically, the same mechanisms might underlie learning behaviors that resemble cultural adaptation in humans. When presented with the songs of other species, zebra finches tend to gradually adjust the duration of inter-syllable intervals toward their own (species-specific) songs' temporal structures, in a way similar to the human adjustment of random auditory stimuli toward Gestalt features. To our knowledge, this work provides the first cellular-level support of the idea of a neurobiological basis of cultural adaptation. It remains to be determined to what extent their findings can be generalized to other species. Would similar neuronal constraints operate in humans? Could they explain perceptual predispositions for some musical features (e.g., for small intervals and isochronous beat)? Are those neuronal constraints species-specific or, instead, are they shared with other species (Nicolai et al., 2014)? Another critical question is whether inter-individual variability in the neural filter is reflected in forms of cultural variation, for example in participant behavior during transmission, or in the shape taken by cultural systems as a result of it. Cross-individual variability is typically regarded as a source of noise in cultural transmission research, and is often removed by means of various procedures. The idea of linking individual neural variability with cultural variation may lend itself well to investigations using brain imaging and electrophysiology, but no one until recently has adopted this approach in cultural transmission research.

Neural Predictors in Cultural Evolution Research

In a recent experiment, Lumaca and Baggio (2016) addressed some of these issues using a neural predictors approach (Berkman and Falk, 2013). This entails use of neuroimaging (fMRI, PET) or electrophysiological methods (EEG/ERPs, MEG) to identify neural predictors of behavior (for examples in the music domain, see Golestani et al., 2002; Zatorre et al., 2012; Zatorre, 2013). Lumaca and Baggio (2016) used neural predictors of signaling behavior as a first approach to examine whether and how symbolic systems adapt to human neural information

processing systems, and to assess the effects of inter-individual variation in neural information processing on three core cultural behaviors: social learning, transmission, and regularization of signal sequences. To this purpose, the authors used one of the best-investigated brain signatures of auditory processing, the mismatch negativity (MMN) (Näätänen et al., 1978).

The MMN is a fronto-central negative wave, evoked by violations of some perceptual regularity (Paavilainen, 2013) which is picked up by the brain in a visual or auditory stimulus stream. The limited influence of attentive processes on the MMN (Paavilainen, 2013) and its onset (~ 200 ms from the relevant stimulus) suggest that the MMN is a *low-level* marker of auditory processing. The encoding of regularities from an auditory input, possibly through the same ASA mechanisms reported above, is an antecedent condition for the elicitation of the MMN (Näätänen et al., 2001). The efficiency of these mechanisms is revealed by the MMN latency and amplitude (Näätänen et al., 1993; Tervaniemi et al., 2001). Larger amplitudes or shorter latencies are typically associated to more accurate representations of the input material and, thus, they are taken as proxies of more efficient encoding mechanisms. The MMN has been used to study how efficiently an individual's auditory system extracts and encodes regularities from acoustic inputs, and how this process may affect linguistic and musical behaviors. For example, differences in ERP responses in infants have been successfully used in various studies to predict cognitive and linguistic development (Molfese and Molfese, 1997; Choudhury and Benasich, 2011). Overall, these studies open up the possibility of using low-level neural markers to predict individual behavior during transmission and acquisition of language, music, and cultural material more generally. Structural properties of symbolic systems may thus be understood as adaptations to information processing bottlenecks during cultural transmission (Kirby, 2001; Tamariz and Kirby, 2015). It should then be possible, for example, to find a relationship between individual brain processing capabilities or limitations, and the degree of regularization imposed by each individual on the cultural material that is being transmitted and acquired.

Neurophysiological (ERP) evidence for this type of effect was provided by Lumaca and Baggio (2016) in the domain of melodic structure. The authors combined ERPs with diffusion chains on two successive days. On day 1, they identified a neural correlate of extracting regularities from 5-tone sequences in musically naïve individuals in a classical auditory oddball paradigm. ERPs were recorded while participants were presented with randomly interleaved standard (80%) and deviant (20%) stimuli: there was no task for the participants, who were watching a silent movie throughout the session. On day 2, participants played a reduced version of MGSGs, with melodic systems of the same kind used by Lumaca and Baggio (2017). Each participant played the first signaling game as receiver (learner) and the second as sender (transmitter)². The main question addressed by the authors was whether constraints and biases on auditory processing could

drive the melodic material toward known Gestalt principles of perceptual organization (Lumaca and Baggio, 2017). The results showed that inter-individual variation in neural information processing, as revealed by the latency of the MMN on day 1, predicted learning and transmission of melodic signaling systems in the MGSGs on day 2. Specifically, individuals with longer MMN latencies performed “worse” in the MGSGs, showing lower coordination, transmission, and accuracy. Yet, these participants introduced more innovations than participants with shorter MMN latencies. Inter-individual variation in neural auditory processing (or regularity encoding) may be sufficient to discriminate “better” from “worse” transmitters, as observed in the cultural transmission of music (Sawa, 2002). However, perhaps the most interesting finding was that participants with longer MMN latencies introduced more regularities in the artificial tone system, reproducing more often melodic structures that were more compressed (signals from the same set became more similar), more proximal (temporally adjacent elements in the signals were closer in pitch), and smoother (the sequences showed a coherent melodic direction) than the sequences they originally received. To our knowledge, this study is the first demonstration that three essential processes underlying cultural evolution (i.e., social learning, transmission, and innovation), and three near-universal properties of melodic structure (i.e., proximity, continuity, and compression) are constrained by the organization of sensory and memory systems in the brain. The MMN is only “the tip of the iceberg” here. The MMN is likely to reflect auditory scene analysis and encoding mechanisms. Constraints on these mechanisms, as revealed (among others) by MMN latencies, may represent a “neuronal niche” through which cultural material must pass, adapt, and evolve (see below). In a cultural evolutionary context, this finding may provide clues to the origins of forms of variation observed in cultural symbolic systems. We discuss this point in the next paragraph.

The Neural Origins of Cultural Variation

Human cultural traits show a myriad different forms across world cultures. Music, like language, provides an excellent example of this diversity, within and between populations (Lomax, 1959; Rzeszutek et al., 2012). For instance, the tendency toward the use of intervals of small size or the division of the octave (2:1) into a limited number of tones (or “discreteness”) as observed in several cultures (Merriam et al., 1956; Dowling, 1968) is counterbalanced by significant diversity, within and between those cultures, in the relative frequency of such traits (Savage et al., 2015). The frequency distribution of proximal intervals (< 700 cents; Savage et al., 2015) differs across musical traditions, with variation being mostly confined to the interval range 0 (unison) to 6 semitones (Huron, 2001). A similar diversity was found in the “tonal material” of musical cultures (i.e., the total set of discrete pitches within an octave), which spans from the 12 semitones of the Western musical scale to the 22–24 microtonal steps of North Indian and Arabic scales (Malm, 1967; Ayari and McAdams, 2003).

coordination work (most code changes) falling to the latter (Nowak and Baggio, 2016).

²In signaling games with fixed roles, including all MGSGs, the receiver tends to learn the code transmitted by the sender. In other words, there is asymmetry in the division of coordination labor between the sender and the receiver, with most

The evolutionary mechanisms that affect the relative frequency of musical characters, such as random cultural drifts and biased selection, have been extensively studied in recent years (Mesoudi, 2015). For example, MacCallum et al. (2012) used a biologically-inspired evolutionary system to explore the effects of “aesthetic” selection on the frequency distribution of musical characters. A population of listeners was asked to rate the pleasantness of randomly generated tunes. The top-rated tunes recombined or mutated into novel variants that were in turn evaluated by a new generation of consumers. The authors reported an over-time increase of characters classically regarded as “musical,” such as isochrony and chordal clarity. This work was the first of its kind to show that consumers’ preferences can deeply shape the evolution of music in the near absence of learning and memory pressures. It is still controversial whether aesthetic preferences are just a social construct, changing over time, or if instead they are themselves stable information processing biases (for an in-depth discussion on this topic see Hodges, 2009; Huron, 2009). In a recent model, Reber et al. (2004) combined the two proposals. Specifically, the authors put forward the hypothesis that aesthetic preferences result from an interaction between knowledge-dependent stylistic rules and information processing fluency for certain stimulus properties (e.g., symmetry, clarity, and the amount of information content) (Nieminen et al., 2011). This may explain the evolution of music toward specific features, such as symmetry and chordal clarity (MacCallum et al., 2012; Verhoef, 2012; Lumaca and Baggio, 2017). A similar proposal was made by Haiman (2011) to explain the emergence of symmetric compounds in language. These arguments are still hypothetical, but we are now starting to understand the *effects of these biases* on the cultural evolution of music (Savage and Brown, 2007). Specifically, we know that these processes can enhance the diversity of musical behaviors and forms, but they can also produce local homogeneity³. While those mechanisms can explain how musical variants spread over time in a population, the sources of variability remain to a large extent elusive.

Up until now, only four main mechanisms of variation have been considered in music: *creative innovation* (e.g., via original musical composition), *borrowing* (through blending or syncretism), *translation* (from one tonal system to another; Alekseyev, 1986), and *random mutation* (errors in music copying or performance) (Savage and Brown, 2007). Lumaca and Baggio (2016) provided evidence for an additional mechanism: *individual neural variability*. One could argue that every individual in a population represents a distinct and unique “neuronal niche” (Dehaene and Cohen, 2007), through which cultural material is filtered and to which it may eventually adapt. Minor inter-individual differences in neural information processing can manifest themselves in differences in musical behavior. Moreover, they can be amplified and spread via different cultural evolutionary mechanisms. Small differences in

learning or information processing can have large system-level effects, if they are amplified by cultural transmission.

One tenet of cultural transmission research is that cultural systems evolve toward certain prior distributions, known as “cognitive attractors” or “inductive biases” (Sperber, 1996; Griffiths et al., 2008). Strong versions of this account have been challenged by recent modeling work (Navarro et al., 2017). The convergence toward priors holds in the (implausible) scenario where all learners are endowed with the same identical prior. However, when learners instantiate (slightly) different constraints, the emerging cultural systems may reflect the more idiosyncratic biases of *some* individuals. In light of our findings, one could suggest that individuals with “tighter bottlenecks” exert a disproportionately large effect on the evolution of musical structures (see Ravignani et al., 2018 for some issues concerning this view). Similarly, differences between populations in brain function and anatomy may, at least in part, be reflected in differences in the structure of the symbolic systems in use. This account has recently found some support in language evolution research. Dediu and Ladd (2007) have shown that the population-level frequencies of two human genes involved in brain growth, Microcephalin and ASPM, are reliably associated with the presence or absence of linguistic tones in that population. The authors’ proposal is that variants of these genes may determine small biases at the individual level in the processing and acquisition of linguistic tones, which may in turn give rise to distinct language variants. Those variants are hardly detectable in individual subjects, because tonal and non-tonal languages can be acquired by any individual, independently of genetic variants (Ladd et al., 2008). But when their effects are amplified by inter-generational transmission (Kirby et al., 2008), these variants may give rise to measurable, large-scale population differences.

Dediu and Ladd (2007) is the first study suggesting that variation, as observed in cultural traits and in their distribution, may originate in interindividual neurogenetic variability. Lumaca and Baggio (2016) provide converging neurophysiological evidence in support of this view (for the genetic bases of inter-individual variation in musicality, see Gingras et al., 2015). Genetic and neural variability are not the only source of cultural variation, but they are likely to play a prominent role in any future theory of the biological roots of culture. For example, Brown et al. (2014) have shown that musical and genetic diversity may correlate to some degree. After sampling a set of traditional songs from 9 indigenous populations in Taiwan, they measured the relative distance for 41 properties of song structure and performance-style. Music and genetic distance among the populations were significantly correlated. A similar relation was found in Eurasian populations (Pamjav et al., 2012). The study of genetic and neural variability may help address questions that were considered taboo in ethnomusicology since fairly recently: for example, whether a causal relationship exists between the distribution of some gene variants and aspects of musical systems and behaviors (Jordania, 2006, p. 101; Nikolsky, 2015). Such a theory requires the synergic and coordinated effort of genetics, neuroscience, and research on cultural evolution. The recent drive toward a “grand synthesis” of the latter discipline (Brewer et al., 2017) makes this possibility somewhat more likely.

³The re-use of Wagner’s musical ideas by other composers during Nazi Germany and the emergence and maintenance of stylistic clusters in contemporary pop music are clear examples of biased selection.

CONCLUSIONS

In this paper, we have argued that some of the most fundamental (and still unresolved) issues in music evolution can be addressed using the methods of cognitive neuroscience. This approach so far suggests a novel hypothesis on the mechanisms behind forms of cultural variation in musical systems. This line of work can also shed light on the “problem of linkage” (Kirby, 1999). Up until recently, this problem has been framed at only two levels of explanation. At the *behavioral level*, individual behaviors (e.g., code changes) that serve coordination and communication are linked to population-level patterns. At the *cognitive level*, sensory or memory constraints in individuals are identified in order to account for properties (e.g., structural features) of cultural systems. We suggest that a third level, the *neural level*, should be taken into consideration when developing accounts of the origins and evolution of structure in cultural systems, as is the case for accounts of the organization and function of information processing systems (Marr, 1982; Baggio et al., 2012, 2014, 2016). Thus, we can address questions in the cultural domain such as: which sources produce cultural diversity (computational level); through which mechanisms it may arise (e.g., inter-individual variation; algorithmic level); and which physical substrates, if any, those mechanisms exploit (i.e., the human brain; implementational level). We believe that explanations at all three levels are necessary to understand human cultural transmission. This requires (1) analyzing the structural and dynamic properties of the cultural systems (or codes) themselves, (2) determining how those are shaped by perceptual and cognitive biases and constraints, and (3) identifying the biological roots of such biases and constraints using neural and genetic data. This proposal generates several new questions, such as: to what extent do neural processes drive cultural evolution? How does inter-individual

variation in brain function and structure affect variation in cultural behaviors? How does the distribution of neural traits in a population affect the structure of the symbolic system itself? How do these traits interact with aesthetic processing biases and the environment at large in the cultural evolution of music? How specific and accurate can neuroprediction get in the context of cultural evolution? Here, we hope to have shown that these questions are worth asking, and are largely amenable to scientific inquiry.

AUTHOR CONTRIBUTIONS

ML wrote the article. AR and GB made additional contributions and edited the manuscript. All authors approved the manuscript for publication.

FUNDING

AR was supported by funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501 with the research Foundation Flanders (FWO) (Pegasus² Marie Curie fellowship 12N5517N awarded to AR), and a visiting fellowship in Language Evolution from the Max Planck Society (awarded to AR).

ACKNOWLEDGMENTS

We are grateful to Monica Tamariz, Bruno Gingras, and Aleksey Nikolsky for their helpful comments during the revision of the manuscript. Center for Music in the Brain is funded by the Danish National Research Foundation (DNRF117).

REFERENCES

- Alekseyev, E. (1986). *Early Folkloric Intonation. Pitch Aspect [Раннефольклорное Интонирование: Звуковысотный Аспект]*. Moscow: Sovetskii Kompozitor.
- Araki, M., Bandi, M. M., and Yazaki-Sugiyama, Y. (2016). Mind the gap: Neural coding of species identity in birdsong prosody. *Science* 354, 1282–1287. doi: 10.1126/science.aah6799
- Ayari, M., and McAdams, S. (2003). Aural analysis of Arabic improvised instrumental music (taqsim). *Music Percept.* 21, 159–216. doi: 10.1525/mp.2003.21.2.159
- Baggio, G., Stenning, K., and van Lambalgen, M. (2016). “Semantics and Cognition,” in *The Cambridge Handbook of Formal Semantics*, eds M. Aloni and P. Dekker (Cambridge: Cambridge University Press), 756–774.
- Baggio, G., van Lambalgen, M., and Hagoort, P. (2012). “Language, Linguistics and Cognition,” in *Handbook of the Philosophy of Linguistics*, eds R. Kempson, T. Fernando, and N. Asher (Amsterdam, NL: Elsevier), 325–355.
- Baggio, G., van Lambalgen, M., and Hagoort, P. (2014). Logic as Marr’s computational level: four case studies. *Top. Cogn. Sci.* 7, 287–298. doi: 10.1111/tops.12125
- Bartlett, F. C. (1932). *Remembering: An Experimental and Social Study*. Cambridge, MA: Cambridge University Press.
- Bendixen, A., Denham, S. L., Gyimesi, K., and Winkler, I. (2010). Regular patterns stabilize auditory streams. *J. Acoust. Soc. Am.* 128, 3658–3666. doi: 10.1121/1.3500695
- Berkman, E. T., and Falk, E. B. (2013). Beyond brain mapping: using neural measures to predict real-world outcomes. *Curr. Dir. Psychol. Sci.* 22, 45–50. doi: 10.1177/0963721412469394
- Bharucha, J. J., Paroo, K., and Curtis, M. (2011). “Alignment of brain states: response to commentaries,” in *Language and Music as Cognitive Systems*, eds P. Rebuschat, M. Rohrmeier, J. A. Hawkins, and I. Cross (Oxford: Oxford University Press), 195–198.
- Blacking, J., Byron, R., and Nettl, B. (1995). *Music, Culture, and Experience: Selected Papers of John Blacking*. Chicago, IL: University of Chicago Press.
- Boyd, R., and Richerson, P. J. (1988). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.
- Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- Brewer, J., Gelfand, M., Jackson, J. C., MacDonald, I. F., Peregrine, P. N., Richerson, P. J., et al. (2017). Grand challenges for the study of cultural evolution. *Nat. Ecol. Evol.* 1:70. doi: 10.1038/s41559-017-0070
- Brighton, H., Smith, K., and Kirby, S. (2005). Language as an evolutionary system. *Phys. Life Rev.* 2, 177–226. doi: 10.1016/j.plev.2005.06.001
- Brown, S., Savage, P. E., Ko, A. M. S., Stoneking, M., Ko, Y. C., Loo, J. H., et al. (2014). Correlations in the population structure of music, genes and language. *Proc. R. Soc. B.* 281:20132072. doi: 10.1098/rspb.2013.2072
- Choudhury, N., and Benasich, A. A. (2011). Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year

- language and cognitive abilities. *Clin. Neurophysiol.* 122, 320–338. doi: 10.1016/j.clinph.2010.05.035
- Christiansen, M. H., and Chater, N. (2008). Language as shaped by the brain. *Behav. Brain Sci.* 31, 489–489. doi: 10.1017/S0140525X08004998
- Claidière, N., Smith, K., Kirby, S., and Fagot, J. (2014). Cultural evolution of systematically structured behaviour in a non-human primate. *Proc. Biol. Sci.* 281:20141541. doi: 10.1098/rspb.2014.1541
- Deacon, T. (1997). *The Symbolic Species*. New York, NY: W.W.Norton.
- de Boer, B. (2000). Self-organization in vowel systems. *J. Phon.* 28, 441–465. doi: 10.1006/jpho.2000.0125
- Dediu, D., and Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10944–10949. doi: 10.1073/pnas.0610848104
- Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398. doi: 10.1016/j.neuron.2007.10.004
- Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234–244. doi: 10.1038/nrn3924
- Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behav. Dev.* 5, 261–276. doi: 10.1016/S0163-6383(82)80036-2
- Deutsch, D. (1999). “Grouping mechanisms in music,” in *The Psychology of Music 2nd Edn*, ed D. Deutsch (San Diego, CA: Academy Press), 99–134.
- Deutsch, D. (2012). *The Psychology of Music 3rd edn*. San Diego, CA: Academy Press.
- Dowling, W. J. (1968). Rhythmic fission and perceptual organization. *J. Acoust. Soc. Am.* 44:369. doi: 10.1121/1.1970461
- Drake, C., and Bertrand, D. (2001). The quest for universals in temporal processing in music. *Ann. N. Y. Acad. Sci.* 930, 17–27. doi: 10.1111/j.1749-6632.2001.tb05722.x
- Edmiston, P., Perlman, M., and Lupyan, G. (2018). Repeated imitation makes human vocalizations more word-like. *Proc. R. Soc. B* 285:20172709. doi: 10.1098/rspb.2017.2709
- Esper, E. A. (1925). A technique for the experiment investigation of associative interference in artificial linguistic material. *Lang. Monographs* 1, 1–47.
- Fay, R. R. (2008). “Sound source perception and stream segregation in nonhuman vertebrate animals,” in *Auditory Perception and Sound Sources*, eds W. A. Yost, A. N. Popper, and R. R. Fay (New York, NY: Springer), 307–323.
- Fehér, O., Wang, H., Saar, S., Mitra, P. P., and Tchernichovski, O. (2009). *De novo* establishment of wild-type song culture in the zebra finch. *Nature* 459, 564–568. doi: 10.1038/nature07994
- Fitch, W. T. (2009). Animal behaviour: birdsong normalized by culture. *Nature* 459, 519–520. doi: 10.1038/459519a
- Fitch, W. T. (2015). Four principles of bio-musicology. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20140091. doi: 10.1098/rstb.2014.0091
- Fitch, W. T. (2017). Cultural evolution: lab-cultured musical universals. *Nat. Hum. Behav.* 1:0018. doi: 10.1038/s41562-016-0018-0
- Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., and Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science* 354, 1278–1282. doi: 10.1126/science.aah6837
- Galantucci, B., and Garrod, S. (2011). Experimental semiotics: a review. *Front. Hum. Neurosci.* 5:11. doi: 10.3389/fnhum.2011.00011
- Gingras, B., Honing, H., Peretz, I., Trainor, L. J., and Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20140092. doi: 10.1098/rstb.2014.0092
- Golestani, N., Paus, T., and Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron* 35, 997–1010. doi: 10.1016/S0896-6273(02)00862-0
- Gould, S. J., and Vrba, E. S. (1982). Exaptation—a missing term in the science of form. *Paleobiology* 8, 4–15. doi: 10.1017/S0094837300004310
- Griffiths, T. L., Kalish, M. L., and Lewandowsky, S. (2008). Review. Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3503–3514. doi: 10.1098/rstb.2008.0146
- Haiman, J. (2011). “Competing motivations,” in *The Oxford handbook of Linguistic Typology*, ed J. J. Song (Oxford, UK: Oxford University Press), 148–165.
- Haynes, J.-D., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534. doi: 10.1038/nrn1931
- Henry, E. O. (1976). The variety of music in a north Indian village: reassessing cantometrics. *Ethnomusicology* 20:49. doi: 10.2307/850820
- Hockett, C. F. (1960). The origin of speech. *Sci. Am.* 203, 89–96. doi: 10.1038/scientificamerican0960-88
- Hodges, D. A. (2009). “The neuroaesthetics of music,” in *The Oxford Handbook of Music Psychology*, eds S. Hallam, I. Cross, and M. Thaut (Oxford, UK: Oxford University Press), 247–262.
- Honing, H., ten Cate, C., Peretz, I., and Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20140088. doi: 10.1098/rstb.2014.0088
- Huron, D. (2001). Tone and voice: a derivation of the rules of voice-leading from perceptual principles. *Music Percept.* 19, 1–64. doi: 10.1525/mp.2001.19.1.1
- Huron, D. (2009). Aesthetics, in *The Oxford Handbook of Music Psychology*, eds S. Hallam, I. Cross, and M. Thaut (Oxford, UK: Oxford University Press), 151–159.
- Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Percept.* 26, 195–204. doi: 10.1525/mp.2009.26.3.195
- Jackendoff, R. S., and Lerdahl, F. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- Jacoby, N., and McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* 27, 359–370. doi: 10.1016/j.cub.2016.12.031
- Johnson, M. R., and Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. *Front. Hum. Neurosci.* 8:59. doi: 10.3389/fnhum.2014.00059
- Jordania, J. (2006). *Who Asked the First Question? The Origins of Human Choral Singing, Intelligence, Language and Speech*. Tbilisi: Logos.
- Kalish, M. L., Griffiths, T. L., and Lewandowsky, S. (2007). Iterated learning: intergenerational knowledge transmission reveals inductive biases. *Psychon. Bull. Rev.* 14, 288–294. doi: 10.3758/BF03194066
- Kirby, S. (1999). *Function, Selection, and Innateness: The Emergence of Language Universals*. Oxford, UK: Oxford University Press.
- Kirby, S. (2001). Spontaneous evolution of linguistic structure—an iterated learning model of the emergence of regularity and irregularity. *IEEE Trans. Evol. Comput.* 5, 102–110. doi: 10.1109/4235.918430
- Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10681–10686. doi: 10.1073/pnas.0707835105
- Kirby, S., Dowman, M., and Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5241–5245. doi: 10.1073/pnas.0608222104
- Lachlan, R. F., van Heijningen, C. A. A., Ter Haar, S. M., and Ten Cate, C. (2016). Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. *Front. Psychol.* 7:980. doi: 10.3389/fpsyg.2016.00980
- Ladd, D. R., Dediu, D., and Kinsella, A. R. (2008). Languages and genes: reflections on biolinguistics and the nature-nurture question. *Biolinguistics* 2, 114–126.
- Le Bomin, S., Lecointre, G., and Heyer, E. (2016). The evolution of musical diversity: the key role of vertical transmission. *PLoS ONE* 11:e0151570. doi: 10.1371/journal.pone.0151570
- Lerdahl, F. (1992). Cognitive constraints on compositional systems. *Contemp. Music Rev.* 6, 97–121. doi: 10.1080/07494469200640161
- Lerdahl, F. (2003). Two Ways in which music relates to the world. *Music Theor. Spectr.* 25, 367–373. doi: 10.1525/mts.2003.25.2.367
- Lévi-Strauss, C. (1960). Éloge de l’anthropologie. Inaugural Lecture at Collège de France.
- Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Lomax, A. (1959). Folk song style. *Am. Anthropol.* 61, 927–954. doi: 10.1525/aa.1959.61.6.02a00030
- Lomax, A. (1968). *Folk Song Style and Culture*. Washington, DC: American Association for the Advancement of Science.
- Lomax, A. (1977). Universals in song. *World Music* 19, 117–129.
- Loui, P. (2012). Learning and liking of melody and harmony: further studies in artificial grammar learning. *Top. Cogn. Sci.* 4, 554–567. doi: 10.1111/j.1756-8765.2012.01208.x

- Lumaca, M., and Baggio, G. (2016). Brain potentials predict learning, transmission and modification of an artificial symbolic system. *Soc. Cogn. Affect. Neurosci.* 11, 1970–1979. doi: 10.1093/scan/nsw112
- Lumaca, M., and Baggio, G. (2017). Cultural transmission and evolution of melodic structures in multi-generational signaling games. *Artif. Life* 23, 406–423. doi: 10.1162/ARTL_a_00238
- Lumaca, M., and Baggio, G. (2018). Signaling games and the evolution of structure in language and music: a reply to Ravnani and Verhoef (2018). *Artif. Life*.
- MacCallum, R. M., Mauch, M., Burt, A., and Leroi, A. M. (2012). Evolution of music by public choice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12081–12086. doi: 10.1073/pnas.1203182109
- Malm, W. P. (1967). *Music Cultures of the Pacific, the Near East, and Asia*. Englewood Cliffs, NJ: Prentice Hall.
- Marr, D. (1982). “Visual information processing: the structure and creation of visual representations,” in *Recognition of Pattern and Form*, ed D. Albrecht (Berlin: Springer), 59–87.
- McDermott, J., and Hauser, M. (2005). The origins of music: innateness, uniqueness, and evolution. *Music Percept.* 23, 29–59. doi: 10.1525/mp.2005.23.1.29
- Mehr, S. A., Singh, M., York, H., Glowacki, L., and Krasnow, M. M. (2018). Form and function in human song. *Curr. Biol.* 28, 356–368. doi: 10.1016/j.cub.2017.12.042
- Merker, B. (2006). The uneven interface between culture and biology in human music (commentary). *Music Percept.* 24, 95–98. doi: 10.1525/mp.2006.24.1.95
- Merker, B., Morley, I., and Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20140095. doi: 10.1098/rstb.2014.0095
- Merriam, A. P., Whinery, S., and Fred, B. G. (1956). Songs of a rada community in trinidad. *Antropos* 51, 157–174.
- Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. Chicago, IL: University of Chicago Press.
- Mesoudi, A. (2015). Cultural evolution: a review of theory, findings and controversies. *Evol. Biol.* 43, 481–497. doi: 10.1007/s11692-015-9320-0
- Mesoudi, A., Whiten, A., and Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behav. Brain Sci.* 29, 329–347. doi: 10.1017/S0140525X06009083
- Molfese, D. L., and Molfese, V. J. (1997). Discrimination of language skills at five years of age using event-related potentials recorded at birth. *Dev. Neuropsychol.* 13, 135–156. doi: 10.1080/87565649709540674
- Moreno, M., and Baggio, G. (2015). Role asymmetry and code transmission in signaling games: an experimental and computational investigation. *Cogn. Sci.* 39, 918–943. doi: 10.1111/cogs.12191
- Morley, I. (2013). *The Prehistory of Music: Human Evolution, Archaeology, and the Origins of Musicality*. Oxford, UK: Oxford University Press.
- Näätänen, R., Gaillard, A. W., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42, 313–329. doi: 10.1016/0001-6918(78)90006-9
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., and Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *Neuroreport* 4, 503–506. doi: 10.1097/00001756-199305000-00010
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., and Winkler, I. (2001). “Primitive intelligence” in the auditory cortex. *Trends Neurosci.* 24, 283–288. doi: 10.1016/S0166-2236(00)01790-2
- Navarro, D. J. A., Perfors, A., Kary, A., Brown, S., and Donkin, C. (2017). When extremists win: on the behavior of iterated learning chains when priors are heterogeneous. *CogSci.* 847–852.
- Nettl, B. (1983). *The Study of Ethnomusicology: Twenty-nine Issues and Concepts*. Chicago, IL: University of Illinois Press.
- Nicolai, J., Gundacker, C., Teeselinck, K., and Güttinger, H. R. (2014). Human melody singing by bullfinches (*Pyrrhula pyrrula*) gives hints about a cognitive note sequence processing. *Anim. Cogn.* 17, 143–155. doi: 10.1007/s10071-013-0647-6
- Nieminen, S., Istók, E., Brattico, E., Tervaniemi, M., and Huottilainen, M. (2011). The development of aesthetic responses to music and their underlying neural and psychological mechanisms. *Cortex* 47, 1138–1146. doi: 10.1016/j.cortex.2011.05.008
- Nikolsky, A. (2015). Evolution of tonal organization in music mirrors symbolic representation of perceptual reality. Part-1: Prehistoric. *Front. Psychol.* 6:1405. doi: 10.3389/fpsyg.2015.01405
- Nowak, I., and Baggio, G. (2016). The emergence of word order and morphology in compositional languages via multigenerational signaling games. *J. Lang. Evol.* 1, 137–150. doi: 10.1093/jole/lzw007
- Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. *Int. J. Psychophysiol.* 88, 109–123. doi: 10.1016/j.ijpsycho.2013.03.015
- Pamjav, H., Juhász, Z., Zalan, A., Nemeth, E., and Damdin, B. (2012). A comparative phylogenetic study of genetics and folk music. *Mol. Genet. Genomics.* 287, 337–349. doi: 10.1007/s00438-012-0683-y
- Patel, A. D. (2010). *Music, Language, and the Brain*. Oxford, UK: Oxford University Press.
- Peretz, I., and Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114. doi: 10.1146/annurev.psych.56.091103.070225
- Ravnani, A., Delgado, T., and Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nat. Hum. Behav.* 1:0007. doi: 10.1038/s41562-016-0007
- Ravnani, A., Thompson, B., Grossi, T., Delgado, T., and Kirby, S. (2018). Evolving building blocks of rhythm: how human cognition creates music via cultural transmission. *Ann. N.Y. Acad. Sci.* doi: 10.1111/nyas.13610
- Ravnani, A., and Verhoef, T. (2018). Which melodic universals emerge from repeated signaling games? *Artif. Life*. 24.
- Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver’s processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3
- Repp, B. H. (1991). “Some cognitive and perceptual aspects of speech and music,” in *Music, Language, Speech and Brain*, eds J. Sundberg, L. Nord, and R. Carlson (Stockholm: MacMillan Press), 257–268.
- Rohrmeier, M., and Cross, I. (2013). Artificial grammar learning of melody is constrained by melodic inconsistency: Narmour’s principles affect melodic learning. *PLoS ONE* 8:e66174. doi: 10.1371/journal.pone.0066174
- Rzeszutek, T., Savage, P. E., and Brown, S. (2012). The structure of cross-cultural musical diversity. *Proc. Biol. Sci.* 279, 1606–1612. doi: 10.1098/rspb.2011.1750
- Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., et al. (2014). Neural constraints on learning. *Nature* 512, 423–426. doi: 10.1038/nature13665
- Savage, P. E., and Brown, S. (2007). Toward a new comparative musicology. *Anal. Approach. World Music* 2, 148–197.
- Savage, P. E., Brown, S., Sakai, E., and Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8987–8992. doi: 10.1073/pnas.1414495112
- Sawa, G. D. (2002). Oral transmission in Arabic music, past and present. *Oral Trad.* 4, 254–265.
- Simon, G., Lanoë, C., Poirel, N., Rossi, S., Lubin, A., Pineau, A., et al. (2013). Dynamics of the anatomical changes that occur in the brains of schoolchildren as they learn to read. *PLoS ONE* 8:e81789. doi: 10.1371/journal.pone.0081789
- Skeide, M. A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., et al. (2017). Learning to read alters cortico-subcortical cross-talk in the visual system of illiterates. *Sci. Adv.* 3:e1602612. doi: 10.1126/sciadv.1602612
- Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford, UK: Oxford University Press.
- Snyder, B. (2008). “Memory for music,” in *Oxford Handbook of Music Psychology*, eds S. Hallam, I. Cross, and M. Thaut (Oxford, UK: Oxford University Press), 107–117.
- Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Oxford: Oxford University Press.
- Tamariz, M. (2017). Experimental studies on the cultural evolution of language. *Annu. Rev. Linguist.* 3, 389–407. doi: 10.1146/annurev-linguistics-011516-033807
- Tamariz, M., and Kirby, S. (2015). Culture: copying, compression, and conventionality. *Cogn. Sci.* 39, 171–183. doi: 10.1111/cogs.12144
- Tchernichovski, O., and Lipkind, D. (2016). Encoding vocal culture. *Science* 354, 1234–1235. doi: 10.1126/science.aal3205
- Temperley, D. (2004). Communicative pressure and the evolution of musical styles. *Music Percept.* 21, 313–337. doi: 10.1525/mp.2004.21.3.313

- Tervaniemi, M., Rytönen, M., Schröger, E., Ilmoniemi, R. J., and Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learn. Mem.* 8, 295–300. doi: 10.1101/lm.39501
- Thompson, B., Kirby, S., and Smith, K. (2016). Culture shapes the evolution of cognition. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4530–4535. doi: 10.1073/pnas.1523631113
- Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20140089. doi: 10.1098/rstb.2014.0089
- Trehub, S. E. (2000). “Human processing predispositions and musical universals,” in *The Origins of Music*, eds N. L. Wallin, B. Merker, and S. Brown (Cambridge, MA: MIT Press), 427–448.
- Trehub, S. E. (2015). Cross-cultural convergence of musical features. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8809–8810. doi: 10.1073/pnas.1510724112
- Verhoef, T. (2012). The origins of duality of patterning in artificial whistled languages. *Lang. Cogn.* 4, 357–380. doi: 10.1515/langcog-2012-0019
- Von Hippel, P. (2000). Redefining pitch proximity: tessitura and mobility as constraints on melodic universals. *Music Percept.* 17, 315–327. doi: 10.2307/40285820
- Wallin, N. L., Merker, B., and Brown, S. (2001). *The Origins of Music*. Cambridge, MA: MIT Press.
- Winkler, I., Kushnerenko, E., Horváth, J., Ceponiene, R., Fellman, V., Huottilainen, M., et al. (2003). Newborn infants can organize the auditory world. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11812–11815. doi: 10.1073/pnas.2031891100
- Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., and Hasson, U. (2017). How we transmit memories to other brains: constructing shared neural representations via communication. *Cereb. Cortex* 27, 4988–5000. doi: 10.1093/cercor/bhx202
- Zatorre, R. J. (2001). Neural specializations for tonal processing. *Ann. N. Y. Acad. Sci.* 930, 193–210. doi: 10.1111/j.1749-6632.2001.tb05734.x
- Zatorre, R. J. (2013). Predispositions and plasticity in music and speech learning: neural correlates and implications. *Science* 342, 585–589. doi: 10.1126/science.1238414
- Zatorre, R. J., Delhommeau, K., and Zarate, J. M. (2012). Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Front. Psychol.* 3:544. doi: 10.3389/fpsyg.2012.00544

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lumaca, Ravignani and Baggio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.