UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN FILOSOFÍA DE LA CIENCIA

INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

# THE EVOLUTIONARY ROOTS OF MUSIC: MUSIC AS AN ASSEMBLAGE OF COGNITIVE CAPACITIES FOR SOCIAL INTERACTION

**TESIS**
QUE PARA OPTAR POR EL GRADO DE
DOCTOR EN FILOSOFÍA DE LA CIENCIA


**PRESENTA:**
LUIS ALEJANDRO VILLANUEVA HERNÁNDEZ

**DIRECTOR DE TESIS:**
DR. SERGIO F. MARTÍNEZ MUÑOZ
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS, UNAM

**COMITÉ TUTOR:**
DR. FRANCISCO VERGARA SILVA
INSTITUTO DE INVESTIGACIONES BIOLÓGICAS, UNAM
DR. JOSÉ LUIS DIAZ GOMEZ
FACULTAD DE MEDICINA, UNAM
DR. IAN CROSS
FACULTY OF MUSIC, UNIVERSITY OF CAMBRIDGE
DR. GERD B. MÜLLER
KONRAD LORENZ INSTITUTE FOR EVOLUTION AND COGNITION RESEARCH (KLI)



CIUDAD UNIVERSITARIA, CD. DE MÉXICO.
MARZO DE 2020

## *Agradecimientos*

La confianza, apoyo y gran afecto que mi familia me ha brindado ha sido una plataforma sólida sobre la cual se han construido los cimientos de mi andar. Ahí he aprendido que caminar por la vida no es fácil, pero que el trabajo constante tiende a dar frutos. Esta tesis es producto de ese andar, y es una manera humilde de agradecer todo el cariño y confianza que mi familia y muchos seres queridos han depositado en mí, desde siempre.

La fortuna de haber podido adentrarme a las prácticas musicales en poblaciones totonacas de la Sierra Norte de Puebla, México, me permitió ampliar mi concepto de música, lo cual, jugó un papel central en el desarrollo de esta tesis. Mi acercamiento con el pueblo totonaco también me dejó muy en claro que el camino para comprender a cabalidad lo valioso de ese legado musical sigue siendo largo de recorrer. Expreso toda mi admiración, agradecimiento y respeto a esos grandes músicos. Con ellos aprendí que la música es un elemento esencial para generar lazos comunitarios de apoyo mutuo. Por ello, agradezco también a todos con quienes haciendo música, en distintos momentos de mi vida, y en diferentes ubicaciones geográficas, hemos podido construir comunidad. De manera especial, a todos con quienes he aprendido, tocado y compartido, el gusto y respeto por el *son* de México.

Mi interés por comprender aspectos relevantes de nuestra capacidad musical, desde una perspectiva evolutiva, cognitiva, filosófica y transcultural, me condujo al Doctorado en Filosofía de la Ciencia de la UNAM. Agradezco al Dr. Sergio Martínez por su interés y apoyo en las primeras formulaciones de este proyecto de investigación. Su gran calidad humana y su sabia orientación durante mis estudios de doctorado fueron fundamentales para que esta tesis pudiera realizarse. Siempre me resultó admirable su capacidad para identificar caminos comunes donde se entrecruzan distintos saberes científicos y filosóficos, así como su manera asertiva de visibilizar esas rutas de encuentro a partir de la elaboración cuidadosa de nuevas herramientas conceptuales. Haber contado con el Dr. Martínez como director de tesis benefició enormemente la consecución de este trabajo, y me deja una gran motivación para buscar renovadas maneras de pensar y hacer filosofía de la ciencia.

En relación a los miebros de mi comté tutor, al Dr. Francisco Vergara, le agradezco el gran impulso que me transmitió por explorar exuberantes y dinámicos mundos de constantes intersecciones biológicas, culturales, cognitivas, artísticas… y un largo etc. Su visión amplia sobre los procesos evolutivos y su brillante manera de abordar estos temas motivó mi empeño en tratar de comprender, de manera más integral, fenómenos tan complejos y fascinantes como es el hecho de hacer música. Por otro lado, al Dr. José Luis Díaz, agradezco su permanente respaldo académico durante todo este proceso. Su interés y amplia experiencia en asuntos neurocognitivos de la música, se vieron reflejados en sus puntuales y valiosas observaciones vertidas en mi tesis. Sus comentarios concernientes al contenido y a la forma de presentar mi trabajo me aportaron grandes enseñanzas que, sin duda, me seguirán siendo de mucha ayuda en investigaciones posteriores.

Al Dr. Ian Cross, le agradezco mucho haberme introducido a un amplio espectro de temas sobre cognición musical. Su crítica inteligente y afilada en contra de maneras eurocéntricas -y por ende, sesgadas- de entender la cognición musical, han dejado una honda huella en esta tesis. Al día de hoy, me queda claro que he podido vislumbrar solo una parte del gran alcance y profundidad de su trabajo académico, pero me motiva saber que con el tiempo iré explorando, y descubriendo, distintas facetas de sus ideas. Le estoy agradecido por mostrar mucho interés en mi trabajo etnomusicológico con poblaciones indígenas de México, y haberme recomendado introducir parte de esos resultados en esta tesis.

Al Dr. Gerd B. Müller, le agradezco su permanente interés en darle seguimiento a mi investigación. Su mirada de amplio alcance al abordar los procesos evolutivos y su atinada crítica hacia los modelos reduccionistas de evolución biológica, me brindaron herramientas conceptuales muy valiosas para desarrollar mi trabajo. Las charlas que hemos podido sostener dentro y fuera del KLI, me aportaron grandes enseñanzas, las cuales, siempre fueron enmarcadas con esa gran sonrisa que caracteriza su sencillez como persona y su gran calidad humana. Le agradezco mucho su lectura cuidadosa de la tesis, sus observaciones y comentarios puntuales, los cuales, beneficiaron enormemente el resultado final de este manuscrito.

La escritura de esta tesis también se vio beneficiada de apoyos institucionales y de espacios académicos en donde las ideas aquí contenidas fueron desarrollándose. La beca Conacyt nacional para estudios de doctorado me permitió realizar mis estudios en la UNAM. Ello me abrió el acceso al variado y dinámico entorno académico del Instituto de Investigaciones Filosóficas, siendo el lugar idóneo para que las ideas contenidas en este proyecto germinaran y fueran teniendo una mejor forma. Los seminarios de “Racionalidad, Razonamiento y Cognición” y “Artefactualidad y Cognición”, a cargo del Dr. Martínez, fueron ámbitos en los que aprendí mucho de mis compañeros. La presentación de los textos y el intercambio de ideas ahí vertidas, dejaron una impronta muy valiosa en esta etapa de mi formación.

A través del programa Beca-Mixta Conacyt, tuve la oportunidad de realizar una estancia de investigación de enero a junio de 2016, en la Facultad de Música de la Universidad de Cambridge, bajo la supervisión del Dr. Ian Cross. Dicha estancia amplió significativamente mis horizontes de investigación. Mi participación en las actividades del *Centre for Music and Science*, dirigido por el Dr. Ian Cross, me sumergió en un río caudaloso de recursos conceptuales que me permitieron entender cada vez con mayor claridad que la música no se reduce a la producción sonora, sino que se trata de una manera fundamental de interacción social humana. Las discusiones en seminarios, cursos y charlas entre profesores y compañeros con los que conviví durante ese tiempo en Cambridge, aportaron mucho a este trabajo y enriquecieron significativamente mi experiencia de vida.

La conclusión de esta tesis en gran medida ha sido resultado de mi estancia de investigación en el Konrad Lorenz Institute for Cognition and Evolution Research, en Klosterneuburg, Austria (KLI). Primero con la beca *Writing-Up* que me fue otorgado por dicho Instituto de marzo de 2018 a septiembre del mismo año. Posteriormente, a partir de un *Hosting Agreement* que el KLI me extendió para permanecer como investigador invitado de septiembre de 2018 a la fecha. No podría enlistar el sin número de experiencias acumuladas durante mi estancia en este lugar tan especial. A cada uno de los miembros que, de manera permanente, conforman la gran familia del KLI, y a cada uno de los fellows que he tenido el privilegio de conocer durante este lapso de tiempo, les expreso mi gratitud por todo lo aprendido y por tantas sonrisas compartidas. Quiero resaltar, además, que no me habría sido posible concluir esta segunda etapa de investigación en Austria, sin el respaldo incondicional y la mano solidaria que me han brindado seres humanos excepcionales cuya amistad, que aquí se ha forjado, trasciende fronteras y tiempos. A todos ellos, mi profundo agradecimiento y mi cariño sincero.

Febrero de 2020
Klosterneuburg, Austria

***Dedicatoria:***

*Esta tesis la dedico a dos estrellas -Alejandra y José Luis- que partieron en momentos tempranos de mi vida; y a dos luceros -Mamá Ufro y Mamá Cata- que que desde mi infancia, han alumbrando mi camino.*

## TABLE OF CONTENTS

Chapter 3. Musical instruments and cognitive evolution



Chapter 4. Niche Construction Theory and music making

**Chapter 5. Music making and social interaction: a co-evolutionary approach**

# INTRODUCTION

Consider the following indigenous celebration that takes places in Northern Puebla, Mexico; the festivity of San Salvador, the Saint Patron of the Totonac Municipality of Huehuetla. From very early in the morning, dancers and musicians gather in the atrium of the church, while inside the priest initiates the first mass of the celebration. In a few hours, the atrium becomes an open-doors public scenario where ritual dances and music are simultaneously performed. At around noon, a couple of men come out from the church carrying on their shoulders the sculpture of the Saint Patron. The priest walks next to them sounding a bell. Many people start walking behind the priest by holding colorful and well decorated candles. The religious procession begins, people will cross over the main streets of the town followed by dancers and musicians. Violins, jaranas, quintas huapangueras, flutes, drums, dancers' zapateados and other acoustic artifacts are remarkably well-coordinated. The emotional atmosphere is intense…The festivity will last several days.[1]

Note that musical practices are deeply attached to the social life of this indigenous community. This festivity also reveals that music involves an ample range of cognitive capacities and skills co-dependently supported by, and distributed among, several cultural practices of this celebration. For instance, the emergence of a mutual alignment between the motor capacities required for both making music and dancing is worthy of note. As a personal anecdote, when I started playing the Danza de Negritos in the violin, Salvador Aquino, a great huastec violinist of this region, instructed me not to focus solely on the music notes, but also to have always in mind, while playing, the bodily movements of the dancers. This means that every music note has to serve as an acoustic platform allowing dancers to master their movements in a comfortable and enjoyable way. Dancers often mention that when music is well played, it makes them dance well too. The same principle works in the other direction. Bodily movements of dancers have an impact on the way music is played. I suggest that

---

[1] See the images in the Appendix section

this mechanism of mutual influence results from the activation of a basic set of cognitive capacities and skills (e.g. motor control, bodily synchronization, and spatial-temporal orientation) involved in both cultural practices. Note that this cluster of capacities and skills also underlie many other social activities going on during this celebration, such as: carrying up the sculpture of the Saint Patron, sounding bells, and walking together in a synchronized way during the procession. This leads us to think that the human capacity for making music does not rely on a specific set of capacities linked exclusively to music. It rather suggests that our capacity for making music is supported by a flexible assemblage of cognitive capacities and skills involved in different social practices.

Surprisingly, this ethnographic observation of a sharing set of capacities and skills involved in distinct cultural practices – supported by the local judgments of these indigenous practitioners – has been misperceived in many evolutionary models of music committed with the task of identifying a music-specific cognitive domain. They have taken for granted that music arose as a whole package, by its own, and at once. These models mainly aim at responding to two issues: the exact moment when music may have first appeared, and its precise adaptive function.

These concerns still prevail in several academic circles. In some seminars and conferences where I have presented my research, I have often dealt with colleagues demanding answers for these kind of queries. My research is not part of this debate, at least not as it has been traditionally formulated. My work is not committed with the task of identifying one evolutionary cause that presumably may explain the emergence of music in a specific period of time. I think that this monolithic and mono-causal explanatory way of addressing the origins of music misperceives the complexity of its target. Furthermore, I will argue that looking for the evolutionary path of certain cognitive capacities exclusively linked to music is a misleading enterprise. The aim of this dissertation is to develop an alternative theoretical model that would allow us to give an evolutionary account of the integration of a basic set of cognitive capacities and skills that may have scaffolded the human capacity for making music. In general terms, I will argue that most likely our ability for making music evolved alongside our basic capacities for social interaction.

In what follows, I will summarize the main ideas developed across the chapters. This overview will provide a general picture of the discussion on music evolution, and will briefly show the theoretical and empirical resources that support my evolutionary account.

The first chapter begins with the discussion about the origins of music from the social-Darwinism perspective of the 19th century. According to that model, it was thought that the study of non-Western music, mostly perceived as primitive, would help us to understand how Western classical music emerged to dominate the world. Later, multiple reactions against this perspective enhanced the development of new disciplinary orientations. After the first half of the twentieth century, ethnomusicology arose as a new field of research. This anthropological orientation stood up against any hierarchical comparison of musics, and it rather advocated for the study of music in all its diversity and as a fundamental aspect of human behavior (Merriam, 1964). In the following decades, however, a comparative agenda in musicology reappeared. By the 70s and 80s, several ethnomusicological studies aimed to compare music from different regions of the world, particularly from regions with similar biological environments and based on similar social structures. This also revived the interest in investigating the possible origins of music. By the end of the twentieth century, ethnomusicologists, musicologists, psychologists, biologists and neurologists, among others, aimed to determine whether the emergence of music was grounded in biological and universal features or whether it was conditioned by culture. In 1997, Nils Wallin, Biörn Merker, and Steven Brown organized a conference in Florence on the topic "The Origins of Music". The talks were published in 2000 under the same title. This compendium aimed to develop a renewed interdisciplinary field of studies under the label of evolutionary musicology. In the introduction, the editors pointed out that all the contributions were theoretically based on the adaptationist perspective of evolutionary psychology.

To have a better understanding of the impact of this evolutionary approach on several models of music evolution, in the next sections of this chapter, I present a general survey of the main tenets of Darwinian adaptationism and the central role it played in the structuring of the Standard Evolutionary Theory (SET), also known as the Modern Synthesis (MS). Starting from Darwin (1859), adaptation was understood as the result of the evolutionary process through which natural selection gradually molded organisms to fit to their specific environment. During the 1920's, the development of population genetics led scholars to consider that natural selection is the sole directional factor in evolution enhancing the adaptation of organisms by choosing the genes that potentiate the expression of specific traits that confer adaptive advantage. In the 1940's, the unification of these ideas became the core of the MS. It is worth noting that this evolutionary approach privileges an externalist and gene-centric explanation of evolution. Accordingly, the shape and structure of organisms results from the external action that natural selection exerts on them and the inheritance of selected genes. Genes were conceived as the units of information responsible for the transmission of biological and cultural traits across generations (for a critical review of this approach see Pigliucci and Müller, 2010).

Next, it will be shown that the consideration of natural selection as the sole evolutionary force generating biological adaptations was also extended to modeling processes of cultural transmission and cultural evolution (e.g., Dawkins, 1976; Boyd and Richerson, 2005; Barkow, Cosmides & Tooby 1992). One of these research programs has been known as evolutionary psychology; its promoters claim that natural selection is the unique mechanism capable of explaining the origins of human cognitive and cultural traits (Barkow, Cosmides & Tooby, 1992). On this view, the complexity of the human mind, conceived of as a set of cognitive specialized systems, was naturally selected and designed during the Pleistocene in order to solve specific survival problems of that time, such as mating, escaping from predators, child rearing, decision making, and etc. In this chapter, I will show that this claim is in line with the idea of an alleged cultural explosion that occurred in Europe during the Upper Paleolithic, which has been characterized by the emergence, as a package, of all the cognitive, biological, and cultural features currently attributed to modern humans, including speech, symbolic thinking, and our musical and artistic capacities (Andersson, 1994; Brown, 2000a, 2000b; Dissanayake, 2000; G. Miller, 2000; Mithen, 2006, 2010; Morley, 2013; Renfrew & Morley, 2009; Trehub, 2000; Wallin, Merker, & Brown, 2000). After this general overview, in the last part of this chapter, I will outline several models of music evolution to show that its main arguments are based on the evolutionary psychology perspective. These evolutionary models of music assume that music arose as a naturally selected package, a package designed to deal with several specific survival pressures. Thus, each model emphasizes a specific survival strategy presumably potentiated by the emergence of music, such as sexual selection, social cohesion, collective coordination, caregiving, and communication.

Given that these models of music evolution are grounded in the evolutionary psychology perspective, which in turn is based upon the main tenets of the Modern Synthesis, the second chapter begins by disclosing the problematic aspects of the MS and the way this evolutionary approach served as the basis for the evolutionary psychology approach. In general terms, I will show that the main weakness of evolutionary psychology is its basic assumption that the structure of human cognition was naturally designed to solve the specific survival problems that our ancestors faced during the Pleistocene. In contrast, I will rather argue that most likely the complex structure of human cognition evolved in distinct periods of time and space through a wide range of evolutionary processes, as suggested by recent archaeological, geological, and physiological findings. This means that the models of music evolution based on the evolutionary psychology framework suffer from the same explanatory limitations. In general terms, these models of music are surrounded by three problematic assumptions: 1) the idea that natural selection is the sole evolutionary force that amalgamated the cognitive components required for making music in order to solve a specific survival issue; 2) the assumption that music arose as a whole package; and 3) the idea that music arose alongside modern humans.

After this critical review, I will present the main tenets of an alternative approach on evolutionary biology: the exaptationist paradigm (Gould & Lewontin, 1979). According to this view, an exaptation refers to a structure or attribute whose current use differs from its originally evolved function. In other words, an exaptation provides new uses to ancient traits originally shaped by natural selection to face different adaptive issues (Gould and Vrba, 1982). Based on a general survey of this evolutionary approach, it will be shown that several models of music evolution have also been erected on the basis of this theoretical perspective (Pinker, 1997, Patel 2008, 2010; Livingstone & Thompson, 2009; Panksepp, 2009). In general terms, we will see that these models defend the idea that music arose by using a set of capacities previously selected for other purposes. Apart from the distinct particularities endorsed by these models, it will be shown that they all disagree in considering the origins of music as an indivisible emerging package. Instead, these models recognize that music is constituted by diverse cognitive abilities and suggest that these cognitive components of music evolved in different periods of time. This is a considerable theoretical improvement. However, I will show that the debate between the adaptationist and exaptationist models of music still leaves unquestioned the narrow idea that natural selection is the only evolutionary mechanism that explains the emergence of any biological and cultural trait, including music (Tomlinson, 2015). Thus, at the end of this chapter, I will show that the core ideas of these models deal with fundamental explanatory limitations of the neo-Darwinist perspective.

At this point, one might think that an alternative explanation of the origins of music should rather take into consideration the development and use of artifacts archaeologically characterized as musical instruments. In order to assess the scope of this explanatory attempt, the third chapter begins by providing a general description of these kind of artifacts and outlines the main arguments in support of the characterization of such objects as musical instruments. It will be shown that, given that most of these instruments date back to the Upper Palaeolithic in Europe, this has led scholars to think that the possible origins of our musical abilities are associated with three evolutionary features: 1) the emergence of musical instruments, 2) the development of strategies for exploring the acoustic qualities of the specific material environment of that time (including caves and rocks), and 3) the use of the cognitive capacities commonly attributed to modern humans.

In the last part of this chapter, I show what are the problematic aspects of these assumptions. I will argue, first, that this narrative resonates with the alleged cultural explosion of the Upper Paleolithic in Europe. Second, I will show that this evolutionary account is based on a eurocentric and narrow idea of what a music instrument should look like. Then, I will propose that an ample range of objects, regardless of their particular morphology, may have been used for multiple purposes, including sound production and music making. Third, I will argue that several rhythmic patterns, potentially used as a device for music making, may have also been incidentally triggered by different modes, for example, by hammering a stone tool, slicing meat with a flint, cutting vegetation, cracking nuts, etc. In this regard, it is also very likely that many musical instruments, for which there are no archaeological records, may have been made of biodegradable materials (Morley, 2013; 2006; Trehub et al, 2015).[2]

In the fourth chapter, I will argue that Niche Construction Theory (NCT) offers a starting point for developing an alternative evolutionary explanation of this matter. I begin by outlining the main tenets of this new evolutionary approach as well as its relevance in evolutionary biology. It will be strongly stressed that NCT does not focus exclusively on the evolution of organisms, but also on the co-evolutionary processes of organisms and their environment. (Scott-Phillips et al, 2014; Day, Laland and Odling-Smee, 2003; Oyama, Griffiths & Gray, 2001, Laland and O'Brien, 2010, Odling-Smee, Laland and Feldman, 2003). In general terms, NCT maintains that organisms actively modify their own environment, and by doing so, alter the conditions of selection acting back on them as well as on other species that inhabit the same environment (Laland & O’Brien, 2010). Nowadays, it is widely accepted that niche construction is a pervasive

[2] They also might be simply common items clearly used for other purposes of daily life, and therefore not attributed to music in our present-day reckoning of functional implements from the past. For example, the recovery of a simple knife in an archaeological site will rarely be categorized as “musical instrument”, but only as “knife”.

phenomenon widespread from bacteria to humans (Laland and Sterelny, 2006). Beaver dams are one of the most cited cases. There is evidence that dams increases the survival advantages and reproductive possibilities of beavers by modifying entire landscapes. Beaver dams alter the water flow of rivers causing a significant impact on a wide range of biological and developmental aspects of different species. Moreover, beaver dams reduce erosion as well as decrease the turbulence that is a limiting factor for much aquatic life.

Importantly, this process encompasses a wide range of behavioral practices involving diverse sets of abilities (Laland and Sterelny, 2006) whose evolvability and transmission cannot be explained purely on the basis of a genecentric evolutionary account. This also reveals that the emergence of a wide variety of our cultural practices among different species should not be characterized as the mere result of an adaptation in the neo-Darwinian sense, but rather as resulting from interactive processes underlying the construction of biological and cultural niches (Laland and O'Brien, 2010). This suggests that most likely the interactive dynamics between organisms and their biocultural environment triggered the use of a wide range of sensorial means such as: touching, smelling, and hearing, among others (Scarre and Lawson, 2006). This in turn may have generated a variety of cultural niches. I suggest to understand cultural niches as learning scenarios supported by numerous sensorial and exploratory means that facilitate the acquisition of a wide range of skills distributed in an ample range of collective practices.

I will examine the way in which our interactions with the material environment may have played an important evolutionary role in developing cultural niches involving a set of cognitive capacities for sound production and social interaction. For a better understanding of these possibly co-evolutionary processes triggering multiple learning scenarios, and supporting the development of our musical capacities, I will argue for a cross-cultural characterization of music. Accordingly, I will show that the western picture of music does not allow us to perceive the ample range of cognitive capacities and social skills involved in making music. I will argue for the necessity of an alternative – namely a non-western characterization of music. In this regard, I will outline several conceptions of music from different human cultures around the world. This will shed valuable light on the complexity of this cultural practice and will show us that music is more than sound patterns, but rather a complex and multimodal cultural phenomenon entailing the intersection of sounds, sensory-motor system, cognitive capacities, skills, artifacts and social conventions. This cross-cultural perspective of music will also reveal that the abilities involved in music rely on our interaction with the socio-material environment. Based on Martinez and Villanueva (2018b), we will see that the abilities displayed by music only come into existence because the environment offers the possibilities for actions that would facilitate acoustic curiosity. For instance, violins can also afford different ways of being played, as shown in distinctive violin fingering techniques and body postures around the world. Thus, the materiality of the instruments also allows, among other things, for the acquisition of musical skills related to how it feels to hold and play an instrument (Martínez & Villanueva, 2018). In this light music appears as an exploratory activity, as music making instead of as a final and static acoustic product. I will propose to understand music making as an emerging and culturally embedded participatory social practice supported by cultural niches constituted of an assemblage of cognitive capacities, skills, artefacts, and social norms, involving complex patterns of interactive-exploration between individuals and their acoustic and material environment.

In order to explore the evolutionary origins of the capacity for making music, I will suggest to focus on a basic assemblage of cognitive capacities and skills constituted by the following components: sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization. This is not an exhaustive list of all the components underlying our capacity for making music. Different componential breakdowns of music might be appropriate for different research purposes. What I propose here is an operational breakdown for understanding how these basic components involved in our capacity for making music may have co-evolved throughout our hominin lineage. In the last part of this chapter, I will hypothesize that the components of this assemblage of cognitive capacities and skills may have emerged as a result of a prolonged and extensive diversification of patterns of interactive-exploration between individuals and their socio-material environment. Accordingly, I will suggest that, over time, these kinds of interactive processes may have given rise to an ample spectrum of cultural practices that emerged by consistently recruiting, sharing and combining distinct, but complementary, sets of cognitive capacities enhancing the evolution of multimodal learning scenarios. Importantly, I will hypothesize that the constitution of these multimodal learning scenarios may have also fostered the complexification of our sociality.

The aim of the fifth chapter is to develop an evolutionary account for explaining the possible processes through which this assemblage of cognitive capacities and skills involved in music making, may have co-evolved and become integrated. I will show that this theoretical enterprise require the integration of recent empirical findings coming from a wide spectrum of disciplines such as: ethnomusicology, music psychology, embodied music cognition, studies on gestural communication, social cognition, cognitive archaeology, material culture evolution, and social interaction, among others.

In the first sections of the chapter, I present a general survey of the lithic traditions archaeologically reported until now. I show that the morphology of the tools and the reconstruction of their possible manufacturing processes shed important light on the evolution of several cognitive capacities and skills. I will argue that the use of gestures may have played an important role in that respect (Wynn, 1991; Coward and Gamble, 2008). I also propose that several patterns of sound production and discrimination, as well as the refinement of motor dexterity implied in tool making, may have been supported by rhythmic sequences of bodily movements, which in turn may have fostered the development of joint coordinated actions. Importantly, coordinated actions involve a wide range of cognitive abilities, among them, a capacity for sensorimotor synchronization. In this regard, I will show that the notion of entrainment sheds light on the way distinct synchronized actions, mutually influencing one another, underlie our capacity for making music as well as many other cultural practices (Clayton, 2012). In a wide range of musical cultures, learning how to play a music instrument does not necessarily require explicit verbal instructions. Rather, it is mostly supported by several patterns of joint actions that become recognizable by musicians during the performance. Thus, a set of capacities such as observation, imitation, and nonverbal communication play a crucial role in triggering these kinds of collective behaviors. This shows that music making is grounded in a multimodal experience recruiting an ample range of cognitive resources. Recent studies have shown that audition reinforces vision in resolving uncertain visual motion patterns, for example, when grasping an object (Castiello et al., 2010). There is also evidence for improved learning in a visual task in subjects trained with congruent audiovisual stimuli compared to subjects trained with solely visual stimuli (Kim et al., 2008; Sigrist et al., 2012). This supports the idea that multimodal learning may strengthen multimodal cognitive representations and the connections between brain areas (Shams and Seitz 2008). Thus, coordination of tasks involving different capacities or subsystems (e.g. seeing, moving, touching, and sound production) has cascading effects in other tasks in which some of the same subsystems are involved (Sheya & Smith, 2010). This suggests that these overlapping multimodal perception-action loops may have also been responsible for the cascading socio-material interactions supporting the construction, over our hominin lineage, of cultural niches constituted by multimodal learning scenarios. I will hypothesize that these cultural niches gave rise to complementary assemblages of cognitive capacities, which in turn underlay the evolution of cultural practices transmitted across generations.

My hypothesis is that the assemblage of capacities and skills involved in our capacity for making music arose from a prolonged and extensive diversification of patterns of interactive-exploration between individuals and their socio-material environment over our hominin lineage. I will suggest that the constitution of this cognitive assemblage did not support solely the evolution of our capacity for making music, but rather a range of cultural practices. Thus, I finish the last chapter by arguing that the assemblage of these basic capacities and skills involved in making music (sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization) could have also enhanced the complexification of our sociality. This suggests that looking for the evolutionary path of certain cognitive capacities exclusively linked to music is a misleading enterprise. In contrast, it is most likely that our capacities for making music evolved alongside our basic capacities for social interaction. Thus, it is not surprising to realize that our capacities for making music are also part of the basic capacities for social interaction, and vice versa. Throughout this thesis, I will show that studies in ethnomusicology (Merriam, 1964; Blacking, 1974; Feld, 1982; Turino, 2008; Moran, 2013; Clayton, Dueck & Leante, 2013; Clayton, Sager & Will, 2005; Clayton, 2012), psychology of music (Trehub, 2000; 2001), embodied music cognition (Cross, 2007; 2001; 2008; Moran, 2013; Matyja & Schiavio, 2013), gestural communication studies (Kendon 1990; 2009; Armstrong & Wilcox 2007; Levinson & Holler 2014), basic social cognition (Hutto & Myn 2013; Gibson, 1986; Gallagher, 2012; Chemero 2009), cognitive archaeology (Wynn, 1991; Coward and Gamble, 2008), material culture evolution (Gamble, 2010; 2012; Ingold, 2007) and social interaction (Levinson 2006; Gallagher 2012) lead us to think that this is the case.

## CHAPTER 1. ADAPTATIONIST MODELS OF MUSIC EVOLUTION

### 1.1. Discussion about the origins of music in the nineteenth century

One of the first evolutionary theories about the origin of music was proposed by Herbert Spencer in the essay *The Origin and Function of Music* published in the mid-nineteenth century.[3] In this work, Spencer argues that the emotions of both humans and the rest of the animals are expressed through muscular movements and sounds. For instance, lions roar when they are enraged and some animals howl when they are wounded; human suffering produces contractions of the body and moans of pain, and human happiness triggers shouts of joy. For Spencer, music arose from emotional sounds. Thus, the variations of music depended on the type of emotion it expressed. He stated that music was originally vocal. For him, pitch, timbre and intervals of vocal music (specifically of songs) allowed to exaggerate the expression of emotions. According to him, a song is a sonic intensification of our emotional language. "Vocal music, and consequently, all music, is an idealization of the natural language of passion" (Spencer, 1904: 414). He asserted that "uncivilized people” have a very monotonous type of vocal music, very close to the everyday speech,[4] and that “we may infer that vocal music originally diverged from emotional speech in a gradual, unobtrusive manner; and this is the inference to which our argument points” (Spencer, 1904: 414). He added that:

> "... we may not only infer, from the evidence furnished by existing barbarous tribes, that the vocal music of prehistoric times was emotional speech very slightly exalted; but we see that the earliest vocal music of which we have any account, differed much less from the emotional speech than the vocal music of our days "(Spencer, 1904: 415-416).

---

[3] This essay was originally published in *Fraser’s Magazine*, London, 1857. Spencer re-published a revised and extended version of this essay in 1904 as part of his essays’ collection: *Essays. Scientific, Political and Speculative*, Vo. II, New York, D. Appleton and Company, Library Edition. The latter version is referred in this thesis.

[4] According to Spencer, vocal music of civilized cultures (western music), is not similar to spoken emotional language due to the high level of sophistication that this music has achieved.

Spencer thought that music arose from the melodic cadences of emotional language (Cross, 2007: 3): "the origin of music as the developed language of emotion seems to be no longer an inference but simply a description of the fact" (Spencer, 1904: 451). Furthermore he pointed out that both the evolution of musical instruments and the development of choral music are processes subjected to the law of progress. According to him, this general law encompasses the simplest to the most complex organism, including historical processes regarding the evolution of society, science, art, and earth (Martínez, 1998b: 160).

Several decades later, Darwin also proposed an evolutionary account for the origins of music. In his work *The Origin of Man*, published in 1871, he maintained that music played an important role in sexual selection. From his view, music in humans is analogous to the sounds produced by males of other species to attract and mate females. In short, the sounds produced by other species for reproductive purposes can shed important light on the evolutionary origins of human music. I will review this proposal with more detail in the section 1.6.1.

In the late nineteenth century, several scholars of Britain were developing theories about the origins of music under the influence of Darwin. In German-speaking regions, this issue led to a considerable reflection about the scope and structure of *Musikwissenschaft* (musicology) as a scientific discipline.[5] After the first full professor chairs were stablish in Vienna (Hanslick 1870, Adler 1898) and Strasbourg (Jadassohn 1897), several other universities in Germany quickly followed suit. One of the most urgent tasks of the young discipline, as Guido Adler pointed out in his inaugural speech at the University of Vienna on 26 October 1898, was to establish an archaeology with which to reconstruct music history from its very first beginning (Rehding, 2000).

[5] The Viennese musicologist Guido Adler (1855-1941) is particularly important because, in his paper of 1885 "The Scope, Method, and Aim of Musicology, he codified the research methods of this discipline by separating what he called historical musicology and systematic musicology (see Mugglestone & Adler (1981/1885)).See also Breuer (2011) for a discussion on how Adler transformed Haeckel's evolutionary biology into a working model for musicological research.

## 1.2. Discussion about the origins of music in twentieth century

During the first third of the twentieth century the discussion about the origins of music led to the development of musicological accounts supporting the view of an alleged superiority of western classical music.[6] This issue led scholars to think that the study of non-western music, considered as primitive, would help to understand how western music arrived to prominence.[7] This research orientation was labeled as comparative musicology. Thus, in the first edition of the *Harvard Dictionary of Music*, Willi Apel (1944) defined comparative musicology as the study of "exotic music", which implies the study of music outside Europe.[8]

Around the last quarter of the nineteenth century and the first half of the twentieth, the attempts to demonstrate an alleged superiority of western culture was strongly criticized.[9] In 1950, the label of comparative musicology was replaced by the term ethnomusicology (Kunst 1975/1950). In 1962, Sachs claimed that the concept of comparative musicology was no longer valid. Comparison, he said, is a methodological base of any branch of knowledge, all descriptions in science and humanities rely on similarities and differences (Sachs, 1962). In 1964, Alan Merriam wrote his influential book T*he Anthropology of Music* in which he defended the anthropological perspective of the ethnomusicological studies. He pointed out that the studies on music should not be reduced to the sound itself, but rather, music had to be understood as a fundamental aspect of human behavior shaped by values, attitudes, beliefs, and social conventions. For Merriam, music and human behavior are mutually dependent

---

[6] These ideas were reinforced by racist evolutionary approaches of social Darwinism that impacted social sciences before and after the Second World War.

[7] See Wallaschek, R. (1893). *Primitive Music*. London: Longmans, Green & Co. Also Graziano, A., & Johnson, J. K. (2006). The Influence of Scientific Research on Nineteenth-Century Musical Thought: The Work of Richard Wallaschek. *International Review of the Aesthetics and Sociology of Music, 37*(1), 17-32.

[8] From an academic point of view, the origin of comparative musicology goes back to the work of Guido Adler (Mugglestone & Adler 1981/1885) in the nineteenth century. For him, this discipline focused on the comparative study of non-western music (particularly folk songs) for ethnographic and classificatory purposes.

[9] See Rehding, A. (2000). The quest for the origins of music in Germany circa 1900. *Journal of the American Musicological Society, 53*(2), 345-385. Although, this account provides a rather conservative musicological perspective.

In the 1970s and 80s, this anthropological orientation led scholars to study musical expressions from regions practically unknown by western culture. Some of the most influential works of this period are the investigations conducted by John Blacking with the Venda of South Africa (1974), Anthony Seeger with the Suyá of Amazonas (1987) and Steven Feld with the Kaluli of Papua New Guinea (1982). These studies also revived the interest in comparison, but this time from new perspectives.[10] For example, certain studies aimed to compare music from different regions of the world that share similar biological environments and social structures. This methodological orientation also revived the interest in the origins of music. By the 1980s, the renewed interest in this matter drew the attention not only of musicologists and ethnomusicologists, but also of psychologists, biologists, and neurologists, among others. The principal concern was to determine whether the emergence of music was grounded on biological and universal features or was conditioned by culture.

The book *Biomusicology: Neurophysiological, Neuropsychological and Evolutionary Perspective on the Origins and Purposes of Music* written in 1991 by Nils Wallin, reflects a biological motivation of the evolutionary bases of music. Some years later, in 1997, Nils Wallin, Biörn Merker and Steven Brown organized a conference in Florence on the topic "The Origins of Music". The talks were published as a book in 2000 under the same title. This compendium aimed to develop a renewed interdisciplinary field of study under the label of evolutionary musicology (Wallin, Merker and Brown, 2000). In the introduction, the editors pointed out that all the contributions were theoretically based on the adaptationist perspective of evolutionary psychology. However, this compilation cannot be characterized as representing a well-integrated research field. Not only do all the contributions endorse an adaptationist perspective based on the general tenets of evolutionary psychology, but this compilation consists of a set of different, and in certain cases incompatible, studies around the evolutionary

[10] Blacking (1974), for instance, locates music as central to other domains of human behavior. Furthermore, taking into consideration the huge variety of music around the world, he considers it reasonable to suppose that music, like language and possibly religion, is a species- specific trait of mankind.

origins of music (see Mithen 2006, for a critical review). Nevertheless, the relevant aspect of this volume is that it endorsed the general idea that evolutionary studies on music can shed important light on significant aspects of human evolution.

In the next section, I will provide a general survey of the theoretical basis on which these models of music evolution have been built.

### 1.3. Adaptationism as an evolutionary explanation

Generally speaking, an evolutionary adaptation is understood as collection of traits that convey survival advantages to organisms. For that reason, it is thought that these traits were naturally selected and inherited across generations (Darwin, 1859). The proposed process through which natural selection gradually molded organisms to fit into their specific environment has been commonly known as adaptation (Darwin 1859, Ayala 1994, Godfrey-Smith 2001). Under this perspective, ‘‘adaptation is always asymmetrical; organisms adapt to their environment, never vice versa’’ (Williams 1992, p. 484). Adaptation has commonly been understood as depending on mechanisms of cumulative selection over long periods of time.[11] The timber wolf is one of the most famous examples provided by Darwin to explain this issue. Darwin argues that timber wolves descended from slower carnivores. He thought that, in the past, certain variations of the speed and strength in the ancestor were selected due to their reproductive advantage. According to this view, in the ancestral population speed was highly variable, but after selection acted, timber wolves increased, considerably, their range of speed. For Darwin, this was a clear example of how the mechanism of natural selection produces an adaptation.[12]

Darwin claimed that the evolution of human culture can be explained by a mechanism for group selection. This mechanism explains why a group of people

[11] Martínez (1998a) notes that for Darwin, adaptation as an evolutionary mechanism is sufficient to explain the survival of our species. Given that adaptation is a processes and not a state, it should not be considered as a perfect end. Thus, there are some species of woodpeckers that inhabit grasslands and treeless areas, such as rocky hillsides and deserts; likewise, some fish use their fins to walk. This means that evolutionary explanations based on adaptive processes give room for contingency.

[12] See Godfrey-Smith (2001) and Amundson (1994) for a debate concerning different ways of characterizing the notion of adaptationism.

retains certain social traits that, albeit pernicious at the individual level, are highly beneficial for groups. One of the examples of these social traits is altruism. For Darwin, altruism reveals how each member of a group, instead of being guided by their own survival instincts, tends to promote social behaviors that contribute to the group welfare (Darwin, 1871).[13] This perspective can be briefly described as follows: human groups whose social organization is grounded on strong moral principles of mutual cooperation tend to have more possibilities for survival. In contrast, those groups that do not promote this type of social principles tend to disappear (Martínez, 1998b).

### 1.4. Adaptationism and Modern Synthesis

During the 1920's, the development of experimental genetics triggered a revision of the main tenets of Darwinism. A decade later, Fisher's treatise, *The genetic theory of natural selection* (1930), argued for a synthesis between Mendelism and Darwinism. Several geneticists led by Fisher developed mathematical tools to provide a more accurate description of how evolution works. These mathematical models engendered a gene-centered view of evolution, according to which, adaptation occurs through a quantitative selection of genes that increase the frequency of the alleles whose phenotypic traits enhance the possibilities for survival. The main idea consisted of assuming that genetic processes that produce small evolutionary changes in the short term (microevolution), if extended throughout geological periods, could also explain the paleontological evolutionary patterns behind the fossil records (macroevolution). In sum, the general claim was that evolution could be explained by mechanisms of natural selection of small genetic changes.

During the 40's, these ideas became the core of the Modern Synthesis (MS) which, among other things, was based on a mathematical account of gene frequencies in populations of organisms. Thus, MS integrated Darwinian natural selection, population-level thinking, and Mendelian inheritance. Since then, it has provided the dominant conceptual framework for evolutionary biology (Mayr,

---

[13] At the same time, this mechanism suggests that individuals that belong to the same group may be indifferent to the problems affecting members of a different population (Darwin, 1871).

1982). From this perspective, also known as neo-Darwinism, solely genetic inheritance accounts for the transmission of selectable variation, and natural selection represents the sole directional factor in evolution that accounts for the adaptation of organisms by selection of genes that potentiate the expression of specific traits required for survival (Scott-Phillips et al, 2014; for a general characterization of the tenets of MS see Laland et al., 2015 and Müller 2017). [14]

In the next section, I will survey two of the most representative models of cultural evolution that based their arguments on a neo-Darwinian perspective. The general aim of these models is to find similarities between processes of cultural transmission and genetic inheritance. In section 1.6., I will show that several evolutionary models of music are based on these neo-Darwinist approaches.

### 1.5. Models of cultural evolution

Generally speaking, neo-Darwinist models of culture claim that the mechanism through which natural selection produces biological adaptations can be extended to explain mechanisms of cultural transmission and cultural evolution. In this vein, it has been argued that cultural transmission and genetic inheritance are driven from similar rules (Dawkins, 1976). Some scholars, inspired by population genetics, explain the adaptive function of social behavior by the mathematical frequency with which a specific behavioral trait is maintained or changed within the same population (Boyd and Richerson, 2005).

On the other hand, evolutionary psychology claims that natural selection is the sole mechanism that can explain the origins of our cognitive and cultural traits (Barkow, Cosmides & Tooby, 1992). Some of the main concerns that evolutionary psychologists address are the following: how have our brains acquired the capacity to process the available information of the environment? How does this information is computed by individuals? How do people base their reasoning on this information? What information do people find memorable?

---

[14] It has been argued that organisms should be considered as simple vehicles for transmitting genetic information (Dawkins, 1976).

What kind of affective reactions does this information produce? What kind of information is easy to learn? What kind of information can be socially transmitted? (Barkow, Cosmides & Tooby, 1992).

For evolutionary psychologists, once the brain receives sensory information from the environment, it produces, as a response, either mental representations or behaviors. Culture is understood as the result of an intricate and complex mechanism of information processing in the human mind (Barkow, Cosmides & Tooby, 1992). These scholars defend the idea that mind was gradually shaped during the Pleistocene by a set of cognitive specialized systems which were naturally selected and designed to solve specific survival problems of that time, such as: mating, escaping from predators, child-rearing, decision making, and etc. They argue that these cognitive designs were inherited across generations due to their adaptive advantages (Barkow, Cosmides & Tooby, 1992). Accordingly, this is what explains the current structure of our behavioral and physiological features. Thus, natural selection is regarded as the sole cause that explains the origins of the complexity of human cognitive structure, which also involves judgments, emotions and actions (Barkow, Cosmides & Tooby 1992).

In the next section I will show that several models of music evolution are based on this neo-Darwinist approach, particularly on the evolutionary psychology perspective. I will argue that these evolutionary models of music are coalesced by three general assumptions: 1) the idea that a set of cognitive capacities were amalgamated by natural selection to produce music, 2) the claim that music was naturally selected to face specific survival issues of our species; and 3) the assumption that music emerged as a package[15] in a certain period of our evolutionary history.

[15] This view takes for granted that all the cognitive capacities involved in making music arose together.

### 1.6. Adaptationist models of music evolution

In what follows, I present a survey of the main models of music evolution based on a neo-Darwinist perspective.[16] This selection does not comprise all the proposals developed on this matter, it rather represents a general overview of this issue.

### 1.6.1. Music and sexual selection

This explanatory model was initially proposed by Darwin (1871). According to him, sexual selection is a mechanism by which certain traits can be preserved and transmitted across generations. Darwin (1871) characterizes two kinds of sexual selection: rivalry and choice. The former refers to the usual competitive fights between males of the same species for mating with a female. The latter refers to courtship behaviors such as: displaying colorful ornaments, expelling smells, and producing sounds in order to attract partners for reproductive purposes. Darwin thought that the origins of music in humans are related with courtship behavior, particularly with birds' songs. He hypothesized that before we developed linguistic abilities, we used melodic intonations to attract partners for sexual reproduction.

Today, some scholars still maintain this explanatory orientation. (Andersson 1994, Miller 2000). For instance, it has been claimed that music (and dance) involves a set of outstanding capabilities for aerobic performance, motor coordination, strength and good health, which can also be considered as sexually attractive human attributes. In this regard, Miller (2000; 2009) argues that these kind of capabilities shaped our aesthetic preferences in the past. Importantly, Miller also thought that these capabilities are perceived as supporting a high level of sexual abilities. Thus, he argues that the attractiveness of the way we make music is connected to a set of desirable abilities for sexual

[16] This is not an exhaustive list of the evolutionary models of music. Moreover, these models also have been classified differently, depending on certain thematic emphasis, which has produced different versions of the same theoretical assumptions.

reproduction. He maintains that, most likely, these attributes were also sexually attractive for our ancestors. From this view, music should be considered as a key aspect of the evolutionary mechanism of sexual selection that facilitated our reproduction and survival.

### 1.6.2. Music and social cohesion

This model states that the emergence of music may have fostered the maintenance of a strong sense of togetherness among individuals of early human groups, by promoting mechanisms of group cohesion and cooperation (Dowling, & Harwood, 1986: 236). Morley (2009) thought that music intensified the experience of early human rituals, which in turn may have contributed to the formation of a sense of community among the participants and a sense of hostility against outsiders (Brown, 2000a).

Some scholars have based these kind of proposals on Freeman's studies (2000) on the oxytocin's effects of emotionally positive memory. These studies show that oxytocin's effects become stronger when there is a considerable activation of the limbic system, mainly produced by an experience of ecstasy. It is well known that this hormone is released during lactation and after orgasm. In both cases, it intensifies the experience of interpersonal coupling. Morley (2009) points out that oxytocin is also released when we listen to music, sing or play an instrument. Likewise, Huron (2001) claims that when we listen to the music we like, the levels of testosterone decrease significantly, which also is associated with low levels of aggressiveness and interpersonal conflicts. Similarly, Sloboda and O'Neill (2001) argue that when people make music, the mood of the participants is led to a positive social disposition. Thus, sharing a positive and common mood within a group potentiates a strong sense of community among individuals.

Given all the above, some scholar think that music played a relevant evolutionary role in reinforcing and maintaining mechanisms of social cohesion. Thus, this approach defends the idea that the emergence of music may have promoted high degrees of social bonding in early human groups, which in turn may have triggered the development of collective actions that increased the chances for survival.

### 1.6.3. Music and collective coordination

The main claim is that the emergence of music may have elicited coordinated movements that supported the development of collective activities that were relevant for survival.

Based on neuroscientific findings, the promotors of this evolutionary model highlight the fact that motor cortex is activated every time we listen to music. This shows that listening to music triggers a biological response that is deeply grounded in our motor system. The main claim is that music, bodily movements and collective actions are not dissociable. Importantly, movements and actions should not be reducible to the kind of movements that playing a musical instrument requires, but they also comprise the movements involved in dancing and other coordinated actions based on collective synchronization. This proposal is also based on ethnographic studies on different human cultures that have extensively shown that religious and social ceremonies are always accompanied by music. This invariably involves coordinated and collective rhythmic activities.

In short, the collective coordination model views music as a social practice that induces the synchronization of body movements. The main evolutionary claim is that the emergence of music may have promoted cooperative behaviors that potentiated the development of coordinated actions that may have led to collective practices highly relevant for our survival, such as hunting and fishing (Brown, 2000a; Dowling & Harwood, 1986).

### 1.6.4. Music for caregiving

This model claims that music played an important role in promoting emotional bonding between mothers and infants enhancing the beginning of our human sociability (Trehub 2000, Dissanayake 2000).[17] Music is ubiquitous in caregiving, caretakers across cultures have sung to infants since time immemorial (Trehub & Trainor 1998). Lullabies to soothe infants and induce sleep have also been considered as one of the most important emotional communicative means between children and their mothers. They are readily identifiable across cultures (Ayres, 1973). There is also strong evidence that infants pay more attention to their mothers when sung to than when mothers use spoken language, apparently due to greater enjoyment of the singing (Trehub 2000).[18] The evolutionary claim is that lullabies may have also been adaptively relevant for our ancestors enabling infants to develop a set of social abilities associated with safety (Cross 2001). In other words, this model contends that the socio-affective interactions between mothers and infants, triggered by lullabies, enabled our capabilities for sociability.

### 1.6.5. Music and language

The evolutionary relation between music and language has been tackled from different perspectives. In the eighteenth century, Rousseau (2000) pointed out that music and language share a common ancestor, then language evolved out of music to support the rational organization of human groups. One century later, Spencer (1904) pointed out that music emerged as an exaggerated intonation to express emotions. By the same time, Darwin (1871) argued for a primitive song-like communication system (protolanguage) and claimed that modern music is a “behavioral fossil” of an ancient communicative system used to attract sexual partners.

---

[17] This approach can be considered as a variant of the "social cohesion" model

[18] According to Trehub (2001), mothers sing to infants mainly at a high pitch, in a slow tempo and with emotional intensity.

Nowadays, there is a general consensus that music and language are structurally complex, acoustically varied and socially significant. This has led some scholars to think that both domains emerged as a response to the same adaptive communicative pressures (Ujhelyi, Molino & Brown, 2000; Mithen 2006). Brown (2000b) considers that both capacities most likely evolved from a common ancestor: the "musilanguage". Brown believed that the similarities between language and music should be explained by virtue of this common ancestor. For him, music and language, as now known, have to be understood as two evolutionary specializations that derived from a communicative precursor with a dual nature: emotive and referential. According to Brown, over time, music emphasized the former and language the latter (Brown 2000b).

On the other hand, Mithen (2006) also believes that both domains arose from a common ancestor, but instead of *musilaguage*, he proposed a sound-gestural emission: "Hmmmmm". For Mithen, "Hmmmm" was a complex and holistic system of communication that involved manipulative, multimodal, musical and memetic intentions. According to him, “Hmmmm” may have initially been used by *Homo ergaster* and *Homo heidelbergensis*. He even claimed that most likely “Hmmm” was highly developed by Neanderthals to facilitate the communicative needs of their large groups. For Mithen (2006), "Hmmmm" was the predecessor of language in *Modern humans*. He contended that language is a human biological feature that arose when our species started dividing the holistic expressions of "Hmmmmm" into discrete units, particularly when these units acquired referential meaning.[19] Thus, he hypothesized that music emerged once human language was fully developed. This means that language became a system of referential communication whereas music became a communicative system of emotions.

Another approach addressing the possible evolutionary connection between music and language has been based on a modularity perspective of the mind. This view assumes that distinct neurological modules are defined by their functional roles. Thus, biological functions are expected to be pre-wired

---

[19] Mithen based this argument upon the Wray (2000) proposal.

exhibiting consistency in localization, which means that if brain specialization for music is pre-wired, then the music specific networks are expected to have a relatively fixed arrangement (Peretz, 2011; Peretz et al 2015). From this view, in order to assess the evolutionarily relevance of music, it would be necessary to identify at least a neuronal network specialized for it. Accordingly, this would allow us to determine if music processing can be considered as a cognitive autonomous and naturally selected mechanism. Peretz and colleagues (Peretz & Pascale 2006; Peretz 2003, 2011) have shown that brain damage can shed important light on this issue. Peretz has conducted many studies to show that brain lesions can selectively interfere with musical abilities while the rest of the cognitive system remains essentially intact. Peretz has also shown that brain-damaged patients may loose the ability to sing familiar songs but retain the ability to recite lyrics and speak with normal prosody. The reverse condition has also been reported. Aphasic patients may remain able to sing familiar tunes and learn novel tunes whereas they fail to produce intelligible lyrics in both singing and speaking (Peretz, Cagnon, Macoir & Hébert, 2004). These kind of studies have led Peretz & Pascale (2006) to suggest that tonal encoding of pitch is the prime candidate for evolving music, as it appears unique to music.[20] For Peretz, a neural perceptual system that is unable to detect small pitch changes is bound to miss an essential part of musical structure.

Over the years, Peretz, alongside other scholars, has continued working on this issue in more detail (Armony et al., 2015; Peretz et al 2015). As expected, these authors have demonstrated that the processing of lyrics, tunes and songs, shares many features that are reflected in brain areas involved in their perception. However, they emphasize that neural overlap does not necessarily mean neural sharing. Their studies show that music and speech stimuli seem to activate distinct neural populations in overlapping regions (Armony et al., 2015; Peretz et al 2015; Peretz et al., 2018). Several fMRI experiments in which short music excerpts and human vocalizations are perceived in a random order by groups of people have revealed the presence of a region in the anterior superior

[20] Peretz also has proposed that the anchorage points of brain specialization for music ascribe a regular beat to incoming events (Péretz & Pascale, 2006)

temporal gyrus (STG) that responds more strongly to music than to human voice, including speech (Angulo-Perkins et al., 2014; Armony et al., 2015). This makes scholars argue for the existence of “music preferred neurons” located in that “music area” (Armony et al., 2015). Accordingly, several authors claim that the research on neural sharing between music and speech is also an important avenue for understanding the origins of our musicality (Peretz et al., 2015; Peretz et al., 2018). Importantly, the emphasis that these kinds of studies have placed on perception of pitch changes have led Péretz & Pascale (2006) to agree with Brown (2000b) that vocal blending may have offered an effective solution to the survival bonding problem of early humans by overriding individuality for the benefit of the group. In 2006, Peretz suggested that this adaptive solution might have been produced by natural selection (Péretz & Pascale, 2006). In recent works, she and her colleagues argue for the possibility that musicality recycled emotion circuits previously evolved for emotional vocalizations (Peretz et al., 2013). This would imply that music is a byproduct of our emotional life. However, the approach chosen by Peretz and colleagues for the accumulation of further empirical data to argue for the existence of “music preferred neurons” located in a specific brain area (Angulo-Perkins et al., 2014; Armony et al., 2015), again supports an adaptationist perspective on the origins of music.[21]

In contrast with this neuro-centered and adaptationist approach, several researches have tackled the relation between music and language by highlighting the types of structured social interactions grounded in mechanisms of nonverbal communication, gestures and bodily movements (Arbib, 2013, Cross, 2001, Kendon, 1990). This view defends the idea that speech, understood as an interactive process, is grounded in a huge variety of mechanisms for social interaction, for instance, such as when individuals rotate their turns in a conversation. This is crucial for making music, but also for doing sports and maintaining the flow of conversation. In this regard, it has been shown that rhythmic patterns involved in both conversation and musical performance arise

[21] As noted, this research program focused on the neurological recognition of tunes and lyrics of music and speech. Its aim is to find a specific area of the brain that may have evolved exclusively to processing music sounds.

spontaneously from mechanisms of social interaction between the participants (Orwin, Howes and Kempson, 2013).

Thus, a highly promising evolutionary explanation of the relation between music and language would require understanding the evolution of our nonverbal communicative skills understood as emerging capacities derived from dynamic processes of social interaction (Arbib 2013; Cross 2001, 2012; Moran 2013; Orwin, Howes & Kempson 2013; Clayton 2013, Martínez and Villanueva 2018a). In the last chapter, I will show that this perspective allows to expand our understanding of the relation between music and our communicative abilities beyond the explanatory limitations of the neuro-centered and adaptationist perspective.

### 1.7. Main tenets of the adaptationist models of music evolution

As above shown, each evolutionary model of music surveyed addresses a particular adaptive issue that was presumably faced once music emerged. Thus, some scholars argue that music, as a form of courtship, provided reproductive advantages in our species. Others claim that music was an important avenue for potentiating mechanisms of collective coordination that enhanced the development of social skills that in turn supported survival practices such as hunting, and territorial defense. Yet others argue that the emergence of music was highly relevant for developing mechanisms of social bonding. Others point out that the adaptive advantage of music, expressed in lullabies, allowed infants to develop their initial capacities for sociability in a context of safety. Still others state that music emerged to solve certain communicative pressures, and over time, it became almost exclusively responsible for communicating emotions. Similarly, other scholars claim that vocal blending may have offered an effective solution to the survival bonding problem of early humans by overriding individuality for the benefit of the group.

As can be noted, these models consider that the emergence of music played an important evolutionary role at the collective level, becoming a highly beneficial social trait for group survival. Furthermore, under this view, it is thought that music arose as a response to the external survival pressures that the environment imposed on early human groups. Thus, the general tenets on which these models are based can be summarized as follows:

a) Natural selection amalgamated the cognitive components required for making music in order to solve an adaptive issue that each proposal addresses.
b) Music arose with modern humans.
c) The set of cognitive capacities involved in music arose together as a whole.

In the next chapter, I will characterize the main explanatory limitations of this neo-Darwinist perspective based on the evolutionary psychology program. I will show that these criticism can be extended to all models of music evolution that are based on this adaptationist approach.[22] Furthermore, I will also survey the main arguments supported by the ex-adaptationist approach to evolution. Then, I will show that this view has also been applied to explain the origins of music. Finally, I will demonstrate the explanatory limitation of these evolutionary models.

[22] In the last chapter, I will show how different features presented by the adaptationist models of music can be complementary and empirically relevant for explaining the origins of music. Yet, the problem is in the adaptive framework that they endorse, which makes problematic to disclose its evolutionary relevance. My aim is to develop an evolutionary narrative on the origins of a basic set of musical capacities not committed with a neo-Darwinist evolutionary approach, but rather based on a dynamic evolutionary account of interdependent co-evolutionary processes that may have fostered the construction of socio-material environments scaffolded by mechanisms of social interaction.

# CHAPTER 2. CHALLENGING THE ADAPTATIONIST/EXAPTATIONIST FRAMEWORK OF MUSIC EVOLUTION

## 2.1. Standard Evolutionary Theory and Evolutionary Psychology

In the previous chapter it was shown that the Modern Synthesis (MS), also known as Standard Evolutionary Theory (SET), constitutes the core of the neo-Darwinist perspective on evolution. It was noted that this evolutionary paradigm privileges an externalist and gene-centric perspective. The shape and structure of organisms are explained as resulting from natural selection that arises from external conditions. Genes are conceived as the naturally selected units of information responsible for the transmission of biological and cultural traits across generations (Pigliucci and Müller, 2010). In the first chapter, I also highlighted the fact that evolutionary psychology resonates with one of the main claims of Modern Synthesis: namely that natural selection amalgamated the components of our cognitive architecture in a way that allowed our ancestors to solve specific survival problems (Barkow, Cosmides & Tooby, 1992).

The explanatory scope of the Modern Synthesis was challenged by the publication of the paper "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program", written by Stephen Jay Gould and Richard Lewontin, and published in 1979. In this paper, Gould and Lewontin pointed out that the adaptationist perspective of evolution conceives the organisms as a "patchwork" whose parts were selected specifically to solve certain problems imposed by the environment (Wilkins & Godfrey-Smith & 2009). According to Gould and Lewontin (1979), the adaptationist model fails in both aspects, in conceiving of natural selection as the sole mechanism to account for any evolutionary process, and in perceiving organisms as a simple collection of discrete objects that natural selection amalgamated over the course of evolutionary history.[23]

---

[23] Gould and Lewontin (1979) criticize the insistent claim of evolutionary biologists of proposing adaptive stories for explaining the emerge of any single organism's trait.

Likewise, the promoters of evolutionary psychology (Barkow, Cosmides, & Tooby, 1992) state that the most important evolutionary changes in human cognition took place during Pleistocene. This claim is in line with the idea of an alleged cultural explosion that occurred in Europe during the Upper Paleolithic, which has been characterized by the emergence, as a package, of all the cognitive, biological and cultural features currently attributed to modern human, including our musical, artistic capacities and symbolic thinking (Andersson, 1994; Brown, 2000b; Dissanayake, 2000; G. Miller, 2000; Mithen, 2006; 2010; Morley, 2012b; 2013; Renfrew & Morley, 2009; Trehub, 2000; Wallin, Merker, & Brown, 2000). This assumption takes for granted that the complexity of human cognition evolved to fit into a very specific and geographically homogeneous environment.

This alleged evolutionary explosion has been challenged by recent archaeological and geological findings showing that the Pleistocene environment was extensively diverse; comprising the African savannah, the Arctic, desert zones, and several areas in the vicinity of rivers, oceans and forests (Foley, 2012). Furthermore, the complex structure of human cognition leads us to think that it might have evolved in distinct periods of time through a wide range of evolutionary processes. Gould and Lewontin were inclined to conceive evolution as the result of a long list of complex interactions between organisms and their environment. In my opinion, this kind of approach would enable us to develop a broader evolutionary account for explaining the development of a wide range of cognitive capacities embodied in cultural practices. As has been shown in the previous chapter, our cultural practices are not driven by, or exclusively correlated with, genetic changes. Caporael, Griesemer and Wimsatt (2014) point out that a wide range of items (parts, properties, relations, cultural processes) that are reproduced and repeatedly assembled, can become entrenched to serve as scaffolding for later items, as a platform or as a constraint. For these authors, the study of these assemblies would constitute an alternative to generate broader conceptualizations to explain processes of cultural and cognitive evolution. Likewise, several scholars have claimed that our understanding of our bio-cultural inheritance would also benefit from recent findings in behavioral and developmental studies (Jablonka & Lamb, 2005; Oyama, Griffiths & Gray, 2001).

However, as we have seen in the previous chapter, most approaches to music evolution still endorse the neo-Darwinist perspective of the evolutionary psychology program. In the next section, I will disclose the main explanatory limitation of this kind of models.

### 2.2. Explanatory limitations in the adaptationist models of music evolution

The main tenets of the adaptationist models of music evolution outlined in the previous chapter are also their main problematic aspects: 1) the claim that music arose as a response to the external pressures imposed by the Pleistocene environment to early humans; 2) the assumption that natural selection is the sole evolutionary and unidirectional account for explaining the origin of music; 3) the assumption that the capacities involved in making music arose all together, as a package, to solve specific survival problems.

This kind of unilateral explanations has been criticized by several authors. Foley (2012) points out that it is very unlikely that the most significant transitions in human cognition occurred during a very specific period of our evolutionary history. For instance, it has also been argued that the emergence of eyes in vertebrates was the result of different evolutionary events, not the product of a singular adaptation for vision. Similarly, Foley claims that the wide range of our human capacities, including music, most likely evolved in different periods of time. Likewise, for Cross (2018) it is very unlikely that music did emerge as a full-blown capacity, as an adaptive response to the environmental pressures of a very specific evolutionary episode. Rather, he suggests that subcomponents of that general capacity emerged at different times. The exaptationist perspective, by contrast, intended to provide a different perspective from the adaptationist view. In what follows, I will outline the main evolutionary tenets of the exaptationist evolutionary approach and will show that some models of music have been erected on the basis of this alternative theoretical approach...

### 2.3. The exaptationist approach

Generally speaking, exaptation refers to a structure or attribute whose later use differs from its originally evolved function (Gould & Lewontin, 1979). The flight function of bird feathers has widely used as an example. It is argued that bird feathers may have originally arisen for thermal regulation, but their use was later *exapted* for flight (Buss, et al., 1998; Norell *et al.*, 2002; Regal, 1975). Thus, an exaptation provides new uses to ancient traits originally shaped by natural selection to face particular adaptive roles (Gould and Vrba, 1982).

### 2.4. Origins of music from the exaptationist approach

Unlike the adaptationist models of music, the exaptationist models go a step further by arguing that music is not an indivisible emerging package, but rather it involves a set of cognitive abilities that may have evolved in different periods of time. The aim of this approach is to show that music arose by using a set of capacities previously selected for other purposes. For instance, back in 1890, William James pointed out that music is zoologically useless and incidentally arose by using the pre-existing nervous system. Similarly, Panksepp (2009) claims that the origins of music have to be analyzed as resulting from the evolution of our emotional brain. In what follows, I will survey three of the most representative models that currently endorse this kind evolutionary perspective of music.

#### 2.4.1. Music as a cheesecake…

It has been argued that music is evolutionarily irrelevant because it did not convey any adaptive advantage for survival.[24] Pinker (1997) claims that music involves, among other things, capacities for motor control and auditory discrimination. According to Pinker, these auditory capacities most likely evolved to make us recognize the distinct sounds we perceive from our environment, and also to escape from danger. Similarly, he hypothesizes that our capacities for

---

[24] According to Pinker (1997), natural selection plays a key role in any evolutionary scientific explanation becoming the solely avenue that we have to understand why life is so special.

motor control (highly relevant for making music), arose to facilitate the development of complex physical dexterities required for walking, running, climbing, etc., which in turn may have been enormously useful to face survival problems such as looking for food, escaping from predators, etc. Thus, Pinker argues that music emerged by using, in a specific way, certain cognitive capacities previously selected for other survival purposes.

For Pinker (1997) the emergence of music, compared to other human capacities, such as language, did not convey any adaptive relevance. For him, music should be understood as a technology that induces pleasure by the activation of cognitive structures previously selected for specific survival purposes. Thus, the pleasurable experiences enhanced by music are comparable to the experiences induced by drugs. Neither of both have any adaptive function; their role is solely to induce pleasure by activating the reward system of the brain. In short, Pinker claims that music is a "cocktail of drugs”, a cocktail that we ingest through our ears to stimulate the neural circuits that produce pleasure (1997: 528). He concludes that music had never appeared in our lives, and everything would have remained the same.

#### 2.4.2. Theory of mind and music origins

According to this model, our capacity for Theory of Mind (ToM) provided the evolutionary foundations for music. General speaking, ToM refers to the ability to understand people as mental beings who have beliefs, desires, emotions, and intentions; and whose actions and intentions can presumably been explained by these mental states. This capacity has been conceived as an evolutionary cognitive advancement responsible for the efficient transmission of knowledge, fostering human cultural evolution (Tomasello, 1999). According to Livingstone and Thompson (2009), ToM has been characterized in two different ways: *Theory Theory* and *Simulation Theory.* The former suggests that humans construct a mental representation of the mental states of their conspecifics in order to predict and explain their actions. The latter suggests that the mirror neuron system simulates a hypothetical state of mind, providing a basis for empathy. Livingstone and Thompson propose that both modes operated in the creation and propagation of music, both are connected with what they consider as the two universal tendencies of music: emotion and ritual. In sum, Livingstone and Thompson (2009) state that music arose as an instance of affective engagement based on the pre-existing capacities of constructing mental models of the emotional states of our conspecifics.

### 2.4.3. Music as a transformative technology of the mind

Another way of understanding music is as a kind of technology, a proposal made by Patel (2010). His general aim is to show that the emergence of music was not the result of a process of adaptation by natural selection. Patel thinks that the adaptationist models of music presuppose the idea that natural selection designed the required skills to develop our capacity for making music (Patel, 2010: 377). He prefers to characterize music as a human invention, as a transformative technology of the mind (TTM) that is biologically powerful. He suggests that music, like the control of fire, become universal in humans because what it offered was universally valued: emotional regulation, a social framework for rituals, and a scaffolding for memorizing long sequences of information (Patel, 2018).[25] He argues that music can also be comparable to our reading skills, in the sense that these skills have been built on the basis of a previously existing brain system associated with visual-spatial cognition. Thus, reading, as music, is a technology which, once invented, has lasting effects on a wide variety of our mental capacities.[26] According to Patel, these technologies (including writing, airplane and internet) have modified the way people communicate, learn and create communities. Thus, they have become highly valuable due to the way they have contributed to shape our lives.

Patel argues that music is an invention shaped by cognitive capacities linked to several non-musical abilities. For example, he claims that auditory

---

[25] Similarly, Patel (2010) claims that the control of fire has been a highly valuable human invention because it transformed our lives in ways we value deeply, for instance, allowing us to cook, keep us warm, and enabling us to see objects in the dark.

[26] According to Patel, writing is also a technology that has allowed us to store and transmit thoughts, and to accumulate knowledge transcending the limits of any individual mind (Patel, 2010: 400).

discrimination of tonality is linked with cognitive mechanisms involved in linguistic processing linguistic. Likewise, he points out that the ability to entrain rhythmic movements to a beat is linked to brain mechanisms involved in vocal learning and vocal control (Patel, Iversen, Bregman and Schulz, 2009). Patel accepts that the origin of language can be explained by natural selection, what he rejects is the attempt to explain the origins of music from an equivalent evolutionary perspective.[27] According to him, music arose by using the neuronal circuits previously selected for linguistic purposes. In short, music should not be characterized as a biological adaptation but as an exaptation, particularly as a technological invention.

It is worth noting that, in contrast to Pinker, Patel argues for the long-lasting effects of music in our lives. Just like other transformative technologies, once invented and experienced, he says, it becomes virtually impossible to get rid of them (Patel, 2010). Furthermore, Patel highlights that music is biologically powerful because it fosters the development of neural networks linked to distinct capacities such as: concentration, emotional regulation, motor control, and language recovery, among others.[28] This leads Patel to remark that music is a human invention that has transformed human life; it is not only a product of alternative uses of our mental capacities, "it has also the power to change the brain, it has the ability to change the nature of ourselves" (Patel, 2010:412).[29]

---

[27] It is important to note that the initial formulation of Patel's TTM leads to a very simplistic characterization of what a technology is in terms of by-product. In my opinion, a novel technology should be characterized in a broader sense, as a kind of material culture innovation arising as a result of co-evolving processes involving a wide range of cognitive capacities. Some years later, Patel (2018) recognized the limitations of his initial characterization of TTM theory and tried to make it consistent with contemporary discussion about bio-cultural evolutionary processes.

[28] In chapter five, we will see that several empirical findings, on which Patel based his proposal, are relevant for my evolutionary account.

[29] Patel (2018) recognizes that his TTM theory implies a divide between cultural invention and biological evolution. Thus, being inspired by current discussions on gene-culture coevolution within evolutionary biology, he tries to update his TTM theory by arguing that music, as an invention, may have also triggered processes of gene-culture evolution.

### 2.5. Explanatory limitations of the exaptationists models of music

Pinker's proposal has been criticized for endorsing a narrow way of understanding music; it is exclusively as a pleasure device. Cross considers that Pinker´s view endorses an ethnocentric perspective of music, a set of complex sequences of sound produced by the few and consumed by the many, simply for pleasure. Cross states that this narrow perspective of music misperceives the importance of music as a complex and socially significant interactive medium (Cross, 2018: 9). In contrast to Pinker, Cross (2011) argues for a cross-cultural perspective of music, socially multifunctional and enormously diverse. Likewise, ethnomusicological studies have shown that music has always fulfilled multiple social purposes: offering possibilities for communicating with dead people (Feld, 1982), intensifying ritual experiences (Feld 1982; Renfrew & Morley, 2009), enhancing and promoting social bonding (Merriam, 1964; Blacking, 1974), preserving and transmitting social and historical memory (Merriam, 1964), etc. Similarly, a wide range of studies in music cognition (Arbib, 2013, Patel, 2010, Cross, 2008, 2011, Peretz and Zatorre, 2003, Zatorre et al, 2007) has shown that music displays a large spectrum of cognitive abilities such as memory, motor control, imitation, gestural communication, bodily synchronization, etc. These findings show that music is not merely a "cocktail of drugs" that we ingest through our ears to experience pleasure.

On the other hand, we saw that Livingstone and Thompson (2009) recognize that music has multiple functions. It was also shown that, for these authors, our pre-existing capacities for mindreading allowed for the origins of music, which in turn improved our emotional and ritual life. Likewise, for Patel (2010), music is a technology that was "invented" by re-using pre-existing cognitive systems previously selected to fulfill other survival purposes. Furthermore, Patel considers that once music was "invented", it had favorable and lasting effects for our species, fostering the development of neural networks linked to distinct capacities such as: concentration, emotional regulation, motor control, and language recovery, among others.

As we can see, these two models, unlike the adaptationist approaches, conceived music as a multifaceted and multifunctional phenomenon constituted by different cognitive components. However, these models situate the origins of music alongside the emergence of modern human; is that when our species was adaptively well equipped to survive. This perspective resonates with the problematic idea of an alleged cultural explosion that occurred during the Upper Paleolithic in Europe. Furthermore, the discussion about adaptationist and exaptationist models of music preserves the idea that natural selection is the sole evolutionary factor for explaining the emergence of any biological and cultural trait, including music. The explanatory limitations of this perspective have already been outlined in the previous chapter. Killin (2016a; 2016b; 2017) claims that the usefulness of this standard set of distinctions is challenged by recent studies on co-evolutionary processes, which undermine the artificial separation of biological and cultural evolution prevalent in several evolutionary models of music. According to Killin, integrating biological and cultural interactions into a complex evolutionary account would provide a more effective theoretical framework for thinking about the emergence of music than the offered by the adaptation/non-adaptation approach.[30]

[30] Recent works have recognized the potential role that Niche Construction Theory can play for developing alternative explanatory accounts for the evolution of music (see e.g. Currie & Killin, 2016; Killin, 2016a; van der Schyff y Schiavio, 2017).

# CHAPTER 3. MUSICAL INSTRUMENTS AND COGNITIVE EVOLUTION

## 3.1. Musical instruments and the origins of music

While the models of music evolution surveyed in the last two chapters have their explanatory limitations, there are other alternatives. For instance, one might think that the origins of music should rather take into consideration the development and use of artifacts, archaeologically characterized as musical instruments. In what follows, I will provide a general description of these artifacts. Then, I will outline the main arguments supporting the characterization of these objects as musical instruments. Finally, I will discuss whether or not these archaeological findings support the idea that music may have arisen as a result of the production and use of these kinds of artifacts.

### 3.1.1. Flutes and whistles

Several artifacts cataloged as flutes are associated with Mousterian Technology (200,000 to 40,000 years ago) and the Middle Paleolithic Age in Europe, as well as with the Middle Stone Age in Africa. This technology has also been linked to Neanderthals in Europe and, in some cases, to the ancestors of modern humans in Africa and Asia. This shows that many of these objects were made before the alleged cultural explosion of the Upper Paleolithic in Europe.

There are four flute-like objects whose legitimacy as musical instruments have been largely debated. These objects were discovered in the late nineteenth and early twentieth centuries. They came from the cave of Haua Fteah in Libya, the cave of Ilsenhöhle in Germany, the cave of Kent in England, and the cave Divje babe I in Slovenia (Morley, 2013). Due to the less-rigorous excavation techniques used at that time, several objects are not well preserved. Furthermore, their physical characteristics were either not reported or detailed. The characterization of these objects as flutes was made almost exclusively on the basis of their external appearance. This is problematic given that our current idea of what musical instruments look like may not necessarily match with the way certain objects could have been manufactured in the past to produce sound.

One of the objects that has perhaps provoked the most intense debate is a flute-like made by a bear's femoral bone found in the cave Divje babe I, in Slovenia (Kunej and Turk, 2000).[31] See the image below.

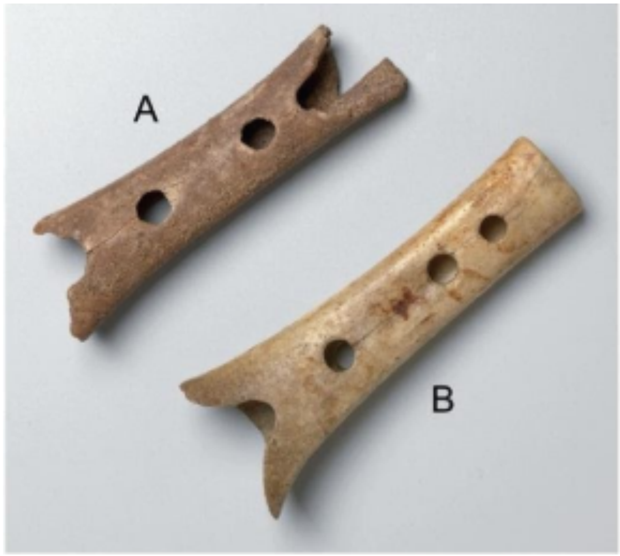



**Figure 1**. A. The posterior side of the 'flute' from the Mousterian layers at Divje babe 1.B. The posterior side of a modern replica of the 'flute' made of the diaphysis of a one - to two-year-old cave bear femur with its metaphysis preserved. Photograph by Tomaz Lauko, National Museum of Slovenia. (Taken from Tuniz, et., al 2012).

The claim for characterizing this object as an ancient flute is based on its external appearance and acoustic properties. In fact, several reconstructions of it have been used to show that it can produce a wide range of Western music scales (Kunej and Turk, 2000). This has led some scholars to conclude that most likely it was intentionally made as a musical instrument (Turk et al, 2006).[32] An extensive ethnomusicological literature, however, has conclusively shown that Western musical scales are not universal; the musical sounds employed in a wide range of musical practices of the world go beyond the conventional characterization of Western music scales. Moreover, detailed studies on the morphology of this object have shown that its manufacture was not a result of deliberate design, but rather of carnivorous activity, most likely of some kind of

[31] More precisely, this object is a left femoral diaphysis of a two years old bear approximately.

[32] These researchers have also argued that this object can be considered as the oldest musical instrument known until now.

hyena (e.g. d'Errico and Villa, 1998, Chase and Nowell, 1998, d'Errico et al, 2003; Morley, 2013).

Another set of artifacts, dating back between 45,000 and 12,000 years ago, have also been characterized as musical instruments. They were made of different sections of animal bones, naturally hollowed and without the inner medulla. (Morley, 2006, 2013). These objects are linked to the Middle and Upper Paleolithic in Europe. Most of them have been classified as flutes and whistles. There is a debate about how they were made and musically executed. A recently updated database reports 144 flutes and whistles associated to the Aurignacian (43,000 to 28,000 years ago), Gravettian (28,000 to 22,000 years ago) and Magdalenian (17,000 to 11,000 years ago) periods (Morley, 2013). Various flutes consist of a hollowed and pierced bone. The unperforated flutes usually produce a single tone. The perforated ones can produce different tones which not only depend on the number of holes they have, but also on the way the air is expulsed through the tube: by blowing from one end of the tube, by blowing into a V-shaped slit (commonly located in one end of the tube),[33] or by blowing and pressing the instrument against the lips (the usual way of playing a snail or trumpet). [34]

In the caves in Ach and Lone Valleys in Swabian Jura, in Southwest Germany, a set of symbolic artifacts from the Upper Paleolithic were found next to several *flutes*. These symbolic objects include stone figures of humans and animals, rock engravings and bone tools. In three caves of this region (Geissenklösterle, Hohle Fels and Vogelherd) the oldest known bone flutes were found. The legitimacy of these flutes is uncontroversial. These flutes were mostly made with bird bones and mammoth ivory. A flute found in the excavations of Hohle Fels is perhaps the most well preserved (Conard et al., 2009). This flute consists of twelve pieces of a vulture radius. It was rebuilt by the archaeologist

---

[33] This structure resembles the current andean quenas made of bamboo, which are used in the music of South America andean region.

[34] Also flutes of distinct sizes can be placed together as a "pan pipe flute", in this way each tube produces different tunes. Similarly, several tunes can be produced by a group of musicians, each one playing their own flute. Pygmies' flute ensembles are an example. In these ensembles, each musician plays a bamboo flute of different size; when the musicians play together, a wide variety of tunes is produced alongside complex rhythmic patterns.

Maria Malina and  is associated with the Upper Paleolithic dating back from around 35,000 years ago and has been linked to Aurignacian technology (Higham et al., 2012; Conard et al., 2009). Its surface was scraped and polished. It has lines marks around the holes. Several researchers think that these lines were made to guide where to make the holes (Morley, 2013). Furthermore, the holes are surrounded by small depressions, which allows them to be covered with fingers in a very precise and comfortable way. This instrument has five holes; four of them are well preserved and one is incomplete (the one at the end of the tube). See the image below.

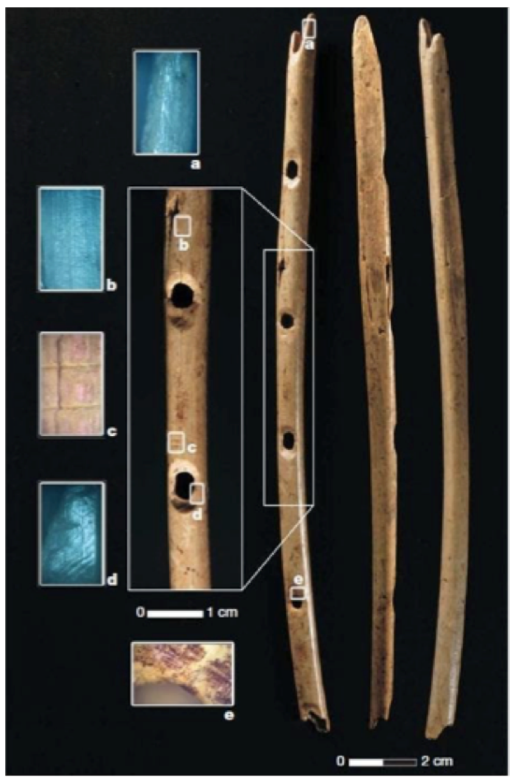



**Figure 2.** Bone flute from Hohle Fels archaeological horizon Vb. Photomicrographs documenting striations and notches from manufacture and polish from use: a, b, d, incident-light fluorescence mode (ultraviolet-and violet-light excitation); c, e, incident-light, obliquely crossed polars, l plate. The photomicrographs were made with a Leica DMRX-MPVSP microscope photometer. The long axis of the micrographs is 2.8 mm long. (Taken from Conard, Malina & Münzel, 2009).

Several fragments of two more flutes were found at the same site. These flutes are made of mammoth ivory (Conard et al., 2009). Three more flutes found in the cave of Geissenklösterle show almost the same manufacturing process. Both are made of swan wing bones.

They have been conventionally identified as Flute 1[35] and Flute 2[36] (Morley, 2013). No other swan bones were found in the excavations, suggesting that these flutes were most likely manufactured elsewhere and brought to the cave (Münzel et al., 2002).[37] The archaeological context in which these objects were found includes Aurignacian tools. Recent studies using C14 radiocarbon date these findings back to around 42,000 years (Morley, 2013). These flutes have been reconstructed with swan bones. It has been shown that they can produce seven tones. They can also be sounded by pressing the lips against the flute and expelling air through the tube, comparable to the way current trumpets are played.[38]

The reconstructed flutes were made to test the acoustic qualities of the original ones. Some melodies have been played with these objects, the recordings are available on internet under the permission of the University of Tübingen.[39] Nevertheless, as previously mentioned, the melodies produced by these instruments do not necessarily reveal the kind of music they played in the past. Playing Western music with these instruments has to be understood only as an exploratory way of using these objects.

A third flute found in the cave of Geissenklösterle is made of mammoth ivory. This flute has been identified as Flute 3 and is associated with the Aurignacian context. It was reconstructed with 31 fragments found between 1974 and 1979 (Conard et al., 2009). The appearance of this flute is comparable to the one previously mentioned, but its production required a higher degree of skill and precision. Mammoth ivory has several layers (as the rings of the trunks of trees). To hollow it and turn it into a flute, it is necessary to cut it in half and remove its central layer from both parts. Then, the separated parts have to be reattached,

---

[35] This flute was polished and has also line marks on the surface. It was rebuilt with 23 fragments extracted in 1990 (Morley, 2013).

[36] This flute was re-built with 7 fragments found in 1973 (Morley, 2013).

[37] This is in contrasts to the case of vulture bones, which have been found in distinct caves of the Suabia region including the caves of Geissenklösterle and Hohle Fels.

[38] Likewise, it it possible to reproduce the sound by inserting, at the end of the flute, a reed (as the one used in clarinet) or a couple of them (as the ones used in oboe) (Morley, 2013).

[39] The music recordings can be reached through this link: http://www.boston.com/news/health/articles/2009/06/24/archaeologists_unearth_oldest_musical_intstruments_ever_found/?page=full

most likely by using some kind of resin. Some marks on the flute have led archaeologists to think that these marks were made to guide the manufacturing processes, resulting in the forms achieved in the bone flutes described previously (Morley, 2013).

Before these discoveries, several flutes found in the cave of Isturitz, and in the Basque region of the French Pyrenees, were the only ones known. All these flutes (or fragments of them) are associated with modern human technology, dating back from the Aurignacian to the Magdalenian period. The vast majority of these flutes were found in an area surrounded by stalagmites. Many of them are decorated with images of animals. In the same site, several archaeological artefacts made of bone, ivory, horns, teeth, shells and stones, were also found. It is believed that this site was inhabited by Neanderthals during the Middle Palaeolithic, before the arrival of modern humans to Europe. This site contained tools and sound artifacts dating back from different periods of time, which has led some scholars to think that the location may have been inhabited by distinct populations in different time periods. Due to the availability of resistant materials (mainly wings of birds) found at this site, the location, and the cave's acoustic properties, it has also been suggested that this site may have enhanced an important part of these populations' cultural lives (Morley, 2013).

On the other hand, *whistles* have also drawn the attention of several archaeologists. Many of these objects are linked to the Paleolithic context. The vast majority of them are made of the phalanges of reindeer (Dauvois, 1989). This material is easy to manipulate and most likely was an abundant resource given the large population of reindeer existing during the Middle and Upper Paleolithic. There has been an intense debate about whether or not these objects were intentionally made to produce sound (Morley, 2013). It has also been discussed whether they have any possible connection with symbolic thinking. Caldwell (2009), for instance, thinks that these objects were not whistles but rather anthropomorphic figures associated with some communal rituals. He bases this interpretation on ethnographic evidence conducted on human populations in Siberia and Greenland. These populations produce objects almost identically designed to the ones above mentioned. They use these objects as amulets; many of them look like a well dressed woman (Caldwell, 2009). Whether or not these objects may have been associated with any kind of symbolic thinking, the sound produced by them can travel over long distances. This has led some scholars to think that these objects may have also been used for communicative purposes between populations (Dauvois, 1989). Interestingly, the sound produced by these whistles induces certain local animals to rest. This suggests that such whistles may have also played a role in developing hunting strategies (Morley, 2013).

In general terms, note that the objects surveyed above are made with strong materials and exhibit similar designs, suggesting the existence of similar manufacturing procedures. Furthermore, several scholars think that the line marks on several flutes were made with different tools at different times (Lawson and d'Errico, 2002). These incisions have been found in many bone artifacts and wooden tools dating back to the Gravettian period. This has led archaeologists to think that these marks may have not been added for decorative purposes, but rather to guide the manufacturing processes.

Most of the flutes were made from bones of adult birds of prey, mainly vultures and, less frequently, eagles. It has been thought that their circular flight patterns may have helped our ancestors to locate dead animals that may have otherwise been difficult to find. Remarkably, the abundant and frequent use of these materials reveals a special type of relationship that may have been developed between humans, birds of prey, and food resources (Morley, 2013).

### 3.1.2. Percussion instruments

One common assumption is that the emergence of music is closely connected with the manufacturing of melodic musical instruments. In many musical practices around the world, however, the use of melodic instruments is not the rule. The human voice is often use as a melodic instrument regularly accompanied by percussive sounds such as drumming, clapping, snapping, patting different parts of one’s own body, etc.. Ethnographic evidence shows that a wide variety of ethnic groups around the world use predominantly percussion instruments, whereas melody is produced by the voice (Morley, 2013; 2006). The archaeological record of percussion instruments includes a wide variety of objects such as drums, scrapers, rattles, rocks, among others. Furthermore, almost any object can be used as a percussion instrument. This makes it difficult to claim that, in the past, certain objects were made, or used, solely for that matter. Most of the percussion instruments around the world are made of biodegradable materials, which leads us to think that the very ancient explorations of the sound properties of the environment may have triggered the use of this kind of biodegradable materials too (Morley, 2013). Hence, the origins of musical practices should not exclusively be linked to the emergence and use of melodic instruments.

As can be expected, archaeologists have also found several percussion instruments made from fossilized material (Dauvois, 1999, Harding, 1973, Morley, 2013). One of the best known instrument was found in Mezine, Ukraine, dating back to about 20,000 years ago. It was made with six mammoth bones. It has several marks showing that it was constantly struck at the same places. Furthermore, it was found next to a couple of drumstick-like artifacts and several rattles. Similar findings have been made in several places (percussion instruments placed next to each other in the same site: mammoth bones with painted red and yellow incisions, jaws, skull's parts, among others). For Bibikov (1978), this suggests the existence of an ancient percussion orchestra of bone instruments.

Some grooved objects may have also been used as scrapers. The grooves are usually made on the surface of a piece of wood, bone or stone. The sound is produced by scrapping (or carving) the grooves, usually with a small object. Several scrapers already registered in the archaeological record have been found in areas of Italy, Belgium, Czech Republic and Spain. They are made with mammoth bones dating back to the Middle Paleolithic (Morley, 2013). Importantly, the habit of making these kind of linear incisions on bones, horns and stones was a popular practice already existing in Gravettian technology (about 27,000 to 22,000 years ago). This does not necessarily reveal an obsession with making scrapers, but rather that the technique to make these kind of artifacts was already developed.

Apart from the relevant role that these objects may have played in exploring the acoustic properties of the environment, caves and other material resources may have been used to explore methods of sound production (Reznikoff, 2008; 2005; Reznikoff and Dauvois, 1988). For instance, Dauvois (1989; 1999) has studied the acoustic properties of some rock protuberances located inside several caves. The remaining marks on these protuberances shows that they were repeatedly struck on specific areas that produce different tunes. Several scholars think that these protuberances may have been used as lithophones. (Dauvois 1989, 1999; Reznikoff, 2008).[40]

An interesting case is the cave of Nerja in Málaga, Spain. This cave has an interior area with excellent acoustic qualities, and also features 19 paintings. The cave contains "the organ", a lithophone-like consisting of a spectacular group of tilted and tightly packed folds. In total there are around 200 folds, most of them are decorated with different designs, including an ibex and a deer (Morley, 2013). Some edges have been broken at various heights and their worn appearance testifies that they were constantly used. Dams (1984), has showed that crystalline sounds of different pitches can be produced by striking the folds with a wooden stick or other object. Dams (1984), claims that the physical wear of several folds was probably caused by breaking the folds intentionally to produce sounds of different pitches.[41] Interestingly, this location continues to be used as a concert hall in the summer.

The cave of Roucador in France also contains a lithophone-like area decorated with lines and black points (Dams, 1985). Apart from the caves and their internal rock formations, certain portable rocks may also have been used for sound production. Many flints and different kinds of stones also have relevant acoustic properties. Zubrow, Cross, and Cowan (2001) argue that our ancestors may also have explored and exploited flints for sound production. For them, many of the Upper Paleolithic flint blades could have been used as lithophones. This leads us to think that early humans from the Upper Paleolithic may have

[40] Apart from the research conducted by Reznikoff and Dauvois, several studies on caves of Spain, France and Portugal de España, Francia, Portugal, reveal this tendency (Dams, 1984, 1985).

[41] As it can be expected, high and low pitches can be produced, respectively, by playing the long and shorts folds.

paid special attention to the acoustic properties of their environment and to the characteristics of the materials with which they interacted during their lifespan.

### 3.1.3. Bullroarers

The archaeological record also includes objects that could have been used as bullroarers dating back to the Upper Palaeolithic. Generally speaking, this instrument consists of a flattened piece of wood, stone, bone or ivory. It has a small hole in one of its ends from which a cord is tied. This instrument emits an intense hum when it is pulled and rotated with a string. The sound varies depending on the size of the object, the length of the string and the resulting radius of rotation. The archaeological record has registered fewer bullroarers than flutes. In fact, determining whether or not an artifact was used as a bullroarer is always problematic. D'Errico and Vila (1997) point out that the perforations found in bones or flattened stones do not necessarily indicate that these artifacts were used as bullroarers. Perforations may also have been used to design distinct artifacts such as fishing rods and collars, among others. These authors do not dismiss that, in certain cases, some perforations may have also been produced by carnivores. Furthermore, it is also possible that bullroarers may have been made with biodegradable materials as wood or bamboo.

In the archaeological record, there are several examples of what might have been used as bullroarers. One of these objects was found in the cave of La Roche de Birol, in the Dordogne, France. It is linked to the Magdaleniense period (Morley, 2013). It is a piece of a carved, oval reindeer horn. This object is approximately 18 mm long and 40 mm wide (at its thickest part). It has several lines marks on its surface. It seems that each line was symmetrically carved. One of its ends has a small perforation, which was presumably made in order to attach a string. It has also been argued that this object may have been used as a large collar; nevertheless, its extremely effective acoustic properties has led scholars to think that it was used as a bullroarer. Dauvios elaborated a replica of this artifact to show that the original one may have produced a very intense sound (Morley, 2013). Dauvios (1989) registers seven other similar objects found in the surrounding areas. Three of these objects dating from the Upper Magdalenian are particularly interesting. Apart from their outstanding acoustic qualities, they have bovine images carved on them. Morley points out that these designs have led some scholars to think that the sound produced by bullroarers may have been used to imitate animals ruminating. In any case, the physical qualities of these objects and the way they were built, lead us to think that they might have been used as bullroarers. Importantly, the objects dating from the Paleolithic are mostly made of reindeer bones, ivory or horn. Nevertheless, this should not be interpreted as an ancient tendency to prefer, exclusively, this kind of material. Rather, the fauna available at that time may have induced the use of this kind of material as one of the ways to explore the acoustic qualities of the environment.

### 3.2. Archaeological findings and the origins of music

The aim of having briefly described the morphology of some artefacts that were presumably used as musical instruments, was to highlight both their acoustic properties and the complexity of their manufacturing processes. As it was noted, the manufacturing process of flutes, percussion instruments and bullroarers, may have required advanced motor skills and a rich set of cognitive abilities. Given that most of these instruments are linked to the Upper Palaeolithic in Europe, one could be tempted to associate the origins of musical abilities to three evolutionary aspects: 1) the emergence of these musical instruments, 2) the mechanisms of sound exploration-exploitation of the material environment available at that time (including caves and rocks), and 3) the set of cognitive capacities commonly attributed to modern humans. As can be noted, this three-part evolutionary narrative would resonate with both the alleged human cultural explosion of the Upper Paleolithic in Europe and adaptationist theories of music. Furthermore, this kind of evolutionary account would leave an important gap between the emergence of modern humans in Africa (more than 100,000 years ago) and the alleged cultural explosion dating to the Upper Paleolithic in Europe.[42] This significant gap leads us to think that the set of cognitive capacities—and the wide variety of artifacts—associated with human populations of Upper Paleolithic may have not arisen as a whole package during that specific period of our evolution.[43] Rather, it is very likely that a biocultural engine of cognitive capacities and skills arose as a result of a conglomerate of cultural processes that evolved throughout our hominin lineage. In fact, there is evidence that several capacities and skills that have been associated with the alleged cultural explosion were already present in Africa during the Stone Age (that is, tens of thousands of years earlier) such as: lithic technology, bone tools, colorful pigments, decorative art, aquatic tools, and hunting practices (McBrearty and Brooks,2000; d'Errico et al., 2003). Several artifacts found in the cave of Blombos in South Africa have led researchers to think that the production of bone tools dates back to about 70,000 years ago (Henshilwood et al., 2001; d; Errico et al 2001). These findings reveal that the artifacts linked to the Upper Palaeolithic most likely resulted from a prolonged development of different kinds of technologies arising in different periods of time (McBrearty and Brooks, 2000). Given all the above, the following questions arise: why have several musical instruments only been found in the above-mentioned sites and not in other places or periods of time? Should we think that the European fauna was a determining factor that made possible the construction of those musical instruments? (Morley, 2013).

It seems very problematic to answer these questions by appealing to a crucial cause occurring during an specific moment in the history of our human evolution. Foley (2012) states that it is an error to think that the most significant cognitive improvements occurred during a specific period of our evolutionary history. For him, the evolution of a complex phenomenon such as music cannot be explained by a single event. Instead, he argues that the set of abilities involved in music most likely evolved over different periods of time (Foley 2012). Certainly, the abundant presence of birds may have played an important role in

[42] Note that this approach also stands for a very questionable primitive characterization of African modern human.

[43] For discussion of an earlier African origins of symbolic art, see Collins, J. E. (2018). Symbolic Arts and Rituals in the African Middle Stone-Age. *Utafiti, 13*(1), 1-22.

this matter (Morley, 2013). Tyrberg (1998) claims that bird populations may have not been equally abundant in Africa prior to the arrival of modern humans to Europe. But it does not imply that the production of musical instruments made of bone resulted from the sudden emergence of new cognitive abilities in a specific place and period of time. It is more feasible to think that it resulted from the exploration of the acoustic qualities of novel natural resources (Morley, 2013). D' Errico et al. (2003) argue that the time, effort, and expertise devoted to the manufacture of these instruments suggests that music may have been of considerable importance for human groups that arrived in these new environments. This also suggests that our musical capacities linked to voice and body movements most likely had an ancient provenance. Cross argues that "the ubiquity of music in native American and Australian societies in forms that are not directly relatable to historic Eurasian or African music strongly suggests that modern humans brought musicality with them out of Africa" (Cross, 2018: 11). It is also very likely that many musical instruments, of which we do not have any archaeological record, may have been made of biodegradable materials (Morley, 2013; 2006; Trehub et al, 2015).[44] The existence of musical instruments made of bone, ivory and horn is not a predominant characteristic in many contemporary traditional societies around the world. Contrastingly, the instruments made of wood, bamboo, hollowed-out firewood, or dry spider nests, are spread worldwide (Morley, 2013, 2006).

D'Errico et al (2003) argue that the archaeological record of musical instruments can lead us to think of a prolonged acoustic exploration of the environment that goes beyond the use of these objects.[45] It is also important to note that tool use and production cannot be dissociated from sound. Hammering is likely to have been a commonly heard sound over the development of stone industries. Likewise, slicing flesh from bones, cutting vegetation, cracking nuts,

---

[44] Morley (2013) urges us to not forget the unexplored vast regions of African whose study will shed important light on the musical instruments that may have been used by modern humans before their arrival to Europa.

[45] In fact, a broad evolutionary account of our ancient exploratory mechanisms of the acoustic environment should not be limited to the idea of a musical instruments deliberately constructed for that specific purpose.

and the use of grinding utensils in food processing should be conceived as ancient sources of sound production (Larsson, 2015). This means that sound production should not been exclusively linked to the paradigmatic notion of a musical instrument. In fact, most artifacts' morphologies can function for different purposes: they do not have stable and inherent properties, which means that there is not a rigid association between an artifact's shape and its function. For Hiscock (2014), hominis tool users could have used diverse artifact forms for different functions. Shaping and reshaping tools may have also altered its morphology, and I would say, may have potentiated the exploration of its acoustic properties. In this line of argument, Hiscock (2014) points out that the associations between particular lithic "types" and particular functions underestimate the complexity and dynamism of tools form/function relationship. Furthermore, tool-making might also have fostered the development of specific forms of listening, discrimination, and production of certain acoustic properties of stones (Blake & Cross, 2008; Cross, Zubrow, & Cowan, 2002; Killin, 2016, 2017; Morley, 2013; Zubrow, Cross, & Cowan, 2001), which in turn may have enhanced knapping expertise. Thus, instead of linking sound production and the origins of music with the production and use of specific kind of artifacts, the origins of music should rather be explored in terms of patterns of acoustic exploration of a wide variety of material resources (e.g., stone, wood, bone, bamboo, human body, etc.) supported by the evolution of an ample spectrum of cultural practices and cognitive capacities.

### 3.3. Critical remarks

In this section, I summarize the problematic aspects of considering the origins of music as a result of the use and construction of the music instruments found in the archaeological record. First, this narrative resonates with the alleged cultural explosion of the Upper Paleolithic in Europe whose explanatory limitations have been outlined in the last chapter. Second, this evolutionary account is based on a eurocentric, narrow idea of how music instruments should look. However, as we discussed in the previous section, distinct objects, regardless of their particular morphology, may have been used for multiples purposes, including sound production. Third, sound patterns may have also been incidentally produced by different means, for example, by hammering a stone tool, slicing meat with a flint, cutting vegetation, cracking nuts, etc. This in turn could have enhanced the development of new exploratory avenues for sound production underlying the development of a great diversity of cultural practices including ancient ways of music making.

# CHAPTER 4. NICHE CONSTRUCTION THEORY AND MUSIC MAKING

## 4.1. Introduction

In the previous chapters, I have shown that the traditional evolutionary accounts of music are based—directly or indirectly—on a neo-Darwinist perspective, particularly on the assumption that natural selection is the sole force explaining the emergence of any cultural or biological trait. As we saw in the second chapter, Gould and Lewontin (1979) had already claimed that this view entails a simplistic view of evolution. In contrast, they thought of evolution as the result of a long list of co-depending influences triggered by reciprocal interactions between organisms and their environment. I will argue that for a better evolutionary understanding of these interactive and co-evolving mechanisms that might explain the origins of our musical capacities, a broader framework is required. Niche Construction Theory (NCT) offers a starting point for developing this kind of evolutionary narrative. In the next sections, I examine the structure of the set of ideas underlying NCT as well as its relevance in evolutionary biology.

## 4.2. Niche Construction theory: a general overview

A considerable range of approaches have challenged several assumptions of Modern Evolutionary Synthesis.[46] One of these approaches is Niche Construction Theory (NCT), which states that organisms actively modify their own environments, and by doing so, alter the patterns of selection acting back on themselves as well as on other species that inhabit this environment (Laland & O'Brien, 2010). Accordingly, the niche construction process provides a second evolutionary route to establishing an adaptive match between organism and their environment. Such matches should not be understood as rigid products of a one way process, solely involving the responses of organisms to environmentally imposed problems (as often articulated in the adaptationist literature). Rather,

[46] It is argued that the study, recognition and, integration of an ample spectrum of evolutionary processes, often neglected by MS, has revealed the necessity for an extended evolutionary synthesis (see, e.g. Pigliucci, 2007, Pigliucci & Müller, 2010; Laland et. al., 2015).

adaptive match should be conceived of as a dynamical product of bi-directional processes involving organisms both responding to pressures posed by their environments, and solving some of those problems while creating new ones as they change their environments. (Laland & O'Brien, 2010: 193). It is important to note that from the perspective of NCT, organisms do not solely inherit genes, but also inherit their modified environments and the practices that produced these environmental modifications (Odling-Smee, Laland and Feldman, 2003; Day, Laland and Odling-Smee, 2003, Oyama, Griffiths & Gray, 2001, Laland and O'Brien, 2010; Odling-Smee, Laland and Feldman, 2003).

From the adaptationist perspective, niche construction has often been neglected as an *evolutionary* factor. It has been claimed, for instance, that the niche construction process results from an adaptive mechanism; it is as a response of organisms to environmentally posed problems. Thus, from the neo-Darwinist view, organisms' niches are seen as extended phenotypes relying on the action of previously selected genes (Dawkins, 1976). Mayr's (1962) distinction between *proximate and ultimate causes* is commonly used to endorse this view. Natural selection is seen as the *ultimate cause* of a phenotype, while developmental processes, such as learning and behavior, are conceived as *proximate causes*. Accordingly, the niche construction process is conceived of as the result of *proximate causes*, as an extended phenotype. Thus, the process through which organisms modify their environment is explained by the effect of natural selection.[47]

However, NCT does not focus exclusively on the evolution of organisms, but also on the co-evolutionary processes of organisms and their environments[48] (see: Scott-Phillips et al, 2014; Day, Laland and Odling-Smee, 2003; Oyama,

---

[47] Godfrey-Smith (1996) has pointed out that this view entails an externalist theory of evolution as it seeks to explain the internal properties of organisms, their adaptations, exclusively in terms of properties of their external environments, it is in terms of natural selection pressures.

[48] For Laland, Odling-Smee and Feldman (2001) two means through which organisms modify its respective environment and the selective pressures it exerts on them are perturbation and re-localization. The former refers to the way organisms alter the components of their environment: for instance, when organisms secrete chemical substances, consume the available resources or made artefacts. The latter is when organisms move to different locations in a way that they are exposed to new environments that will also modify. These authors point out that in most of the cases, niche construction involves both processes.

Griffiths & Gray, 2001, Laland and O'Brien, 2010, Odling-Smee, Laland and Feldman, 2003). It has been shown, for instance, that human cultural variation, depending largely on social learning, may result in cultural niche-constructing practices that modify the natural selection of certain human genes. Thus, 'causation' processes becomes complex, which implies to replace the traditional dichotomous *proximate* and *ultimate* distinction by a notion of *reciprocal causation* (Kendal et al., 2011). It is widely accepted that niche construction is a pervasive phenomenon widespread from bacteria to humans (Laland and Sterelny, 2006). Beavers' dams are one of the most cited examples of niche construction. There is evidence that dams increase the survival advantages and reproductive possibilities of beavers. Furthermore, beaver dams also alter the water flow of rivers causing a significant impact on a wide range of biological and developmental aspects of different species. Moreover, beaver dams reduce erosion as well as decrease the turbidity that is often a limiting factor for much aquatic life.

Termite mounds are another illustrative example of this process. Mounds contain an extensive system of tunnels and conduits that serve as a ventilation system for the underground nest. This internal ventilation system regulates mounds' temperature and humidity allowing termites to deal with the external characteristics of the environment. Moreover, mounds protect termites against predators (Laland and Sterelny, 2006). Other dwellings have beneficial effects on the organisms that construct them. Ants have deficient behavior when they are exposed to temperatures below 20 degrees Celsius. Their nests functions as temperature regulators, keeping temperatures above 20 degrees Celsius (Day, Laland & Odling-Smee, 2003). Similarly, earthworms' physiology restricts earthworms' abilities to cope with their physical environment. There is evidence that these organisms evolved from an aquatic ancestor, and the tunnels they built allowed them to solve specific problems of water requirements and salt balance. Importantly, during the excavation process, earthworms ingest fine soil particles and digest organic waste. Over this process, they also aerate and enrich the soil by releasing phosphorus and potassium from the subsoil. This compensates for the earthworm's physiological limitations, leaving positive consequences for the environment, and generating a beneficial impact on the life of other organisms in the same area (Laland & Sterelny, 2006). On the other hand, the construction of webs by arachnids is highly relevant for the development of social life among these species. Webs allow spiders to develop collective strategies for communication and hunting, resulting from repetitive web-building practices carried out across arachnid' generations (Day, Laland & Odling-Smee, 2003).

These examples lead us to see that the niche construction process is a way through which organisms build and inhabit "micro-universes" that allow them to solve important needs. As it can be noted, this process encompasses a wide range of social practices involving diverse sets of abilities (Laland and Sterelny, 2006) whose evolvability and transmission cannot be explained solely on the basis of a gene-centric evolutionary account.

Constructing niches compensates for the physiological limitations of organisms and potentiates the emergence of new anatomical and behavioral features (Day, Laland & Odling-Smee, 2003). These practices can alter the morphology of organisms including the emergence of new genes. Lactose tolerance in humans is a frequently cited example of this. It is widely known that lactose tolerance depends on genes associated with the digestion of milk. There is an important correlation between the emergence of these genes and the history of livestock farming. This cultural practice, spread among shepherd communities, created alternative selective pressures that enhanced the emergence of the genes responsible for the absorption of lactose (Aoki 1986; Feldman and Cavalli-Sforza 1989; Tishkoff et al. 2007). Thus, the rate of alleles associated with lactose digestion increases proportionally with the period of time this practice has been preserved and transmitted across generations. Conversely, if a considerable number of individuals stop consuming milk, the genes associated with lactose tolerance will not be transmitted to the next generations (Laland & Sterelny, 2006). This shows that there were not genes naturally selected and designed for lactose absorption before the emergence of livestock farming. The differences between lactose tolerance and intolerance among human populations cannot be explained as an adaptation in the traditional neo-Darwinian sense; by appealing to a previous genetic differentiation imposed by natural selection. Rather, these studies show that the emergence of livestock farming gave rise to genes that favor lactose absorption process (Laland & Sterelny, 2006). The link between livestock farming and the absorption of lactose allows us to see that niche construction is not a process genetically caused, but rather it reveals how cultural practices can foster the emergence of morphological and genetic changes. Likewise, several diseases caused by socio-cultural factors—such as lack of hygiene, poor diet, and poverty conditions of life in densely populated cities—strengthened the immune system of individuals living under these conditions. This shows that phenotypic variations can be biased by developmental and cultural constraints, which implies that genes do not always have a generative role but often a stabilizing one (Müller, 2019).

It has also been extensively reported that music making produces structural differences in the brain. For example, certain structural differences due to musical training extend to motor and sensorimotor cortices, to premotor and supplementary motor areas, and involve subcortical structures such as the basal ganglia and the cerebellum (e.g. see Amunts, et al 1997; Bangert & Schlaug, 2006; Bermudez et al, 2009; Elbert et al 1995; Hutchinson et al., 2003). There is evidence that this neuronal circuitry is engaged in motor control and fine motor planning and performance (e.g. finger motions) during music execution as well as in motor learning (Schmidt & Lee, 2011). There have also been observed differences in brain connectivity. For example, musicians exhibit greater midsagittal size of the corpus callosum (see e.g. Lee et al, 2003). These cases reveal that music can drive physiological changes in individuals. Importantly, musicians can transmit these physiological alterations to their offspring only by inheriting the practice of music, which means that these morphological modifications cannot be genetically transmitted across generations.

As we can note, niche construction, understood as a co-causal evolutionary process, generates modified habitats and cultural practices that are transmitted by organisms to their descendants through non-genetic, “ecological” inheritance. In humans, ecological inheritance incorporates heritable material culture, inheritable cultural knowledge (Odling-Smee et al., 2003) and cultural practices. This includes a broad spectrum of institutions, beliefs, social organizations, artifacts, and social behaviours that have been constructed and transmitted across generations. Hence, instead of explaining the emergence of any cultural trait as the result of a genetic adaptation to an autonomous and fixed environment, NCT invites us to consider and develop evolutionary explanations of co-evolutionary processes between organisms, environments and capacities. NCT replaces an externalist-internalist view of evolution for an interactionist approach (Kendal, Tehrani and Olding-Smee, 2011). This means that organisms and their environment are mutually engaged through co-evolving evolutionary mechanisms based on processes of reciprocal causation. In summary, NCT goes beyond evolutionary explanations based on traditional dichotomies between causes and effects, as well as biology and culture (Day, Laland and Odling-Smee, 2003). In the next section, I will argue for the relevance of cultural niches constituted by multimodal scenarios of social learning.

### 4.3. Cultural niches as multimodal learning scenarios

As previously discussed, niche construction has important implications for the relationships among genetic evolution, development, and cultural processes. One implication is that niche-constructing organisms can no longer be seen as mere ‘‘vehicles’’ of their genes, as Dawkins (1976) suggests. Organisms also modify the selection pressures of their own (and of other species’) environments. A second implication is that there is no requirement for niche construction to result directly from genetic variation in order for it to modify natural selection. Humans modify their environments mainly through cultural processes, and it is this reliance on culture that lends human niche construction a special potency (Laland, K & O’Brien M.J., 2011). Processes such as learning can be of considerable importance to subsequent generations because learned knowledge can guide niche construction, the consequences of which can be inherited through ecological inheritance. In this respect, learning provides a rich source for the development of social practices that can be expressed in niche construction.

Agriculture, for instance, was neither invented by an individual farmer nor did it appear as the result of a genetic mutation. Likewise, it is not genes that are responsible for domesticating dogs, producing cheese, or growing rice. In short, there are not genes exclusively selected for cultural practices. The development and spread of agriculture was based on processes of social learning. The major cultural transitions, such as agriculture and the colonization of new environments, promoted the consumption of new resources that significantly impacted our human genetic configuration (Laland & O'Brien, 2010).

This reveals that the emergence of a wide variety of cultural practices should not be characterized as the result of an adaptation in the neo-Darwinian sense, but rather as the result of the construction of socio-cultural niches (Laland and O'Brien, 2010). I suggest to define *cultural* niches as those *constituted by learning scenarios that facilitate the acquisition of a wide range of skills distributed in numerous collective practices.* The construction of these cultural niches may also have played an important role in the evolution of our cognition, which I characterize as distributed among individuals, practices and artifacts (Hutchins, 1995; 2006). Importantly, in this work I embrace the idea of cognition as action oriented, it is primarily structured by complex patterns of interaction between organisms and their environment (Hutto & Myin, 2013). Accordingly, cognition is not separate from sensory-motor processes but as arising from them (Sheya & Smith 2010).[49]

This cognitive approach has been reinforced by recent studies on learning and development. It is widely known, for instance, that children are not passive recipients of the instructions given by adults, but rather they also co-direct their own processes of learning by interacting with other children, the adults they live with, the artefacts they use and the activities with which they daily engage (Flynn et al, 2013). Through these participatory processes, children become co-constructors of their own learning environment. Furthermore, brain plasticity of infants is not fully developed at birth, but it is highly fostered by the

[49] The roots of this kind of approach can be found in phenomenologists like Husserl (1998) and Merlau-Ponty (1962). Later, these ideas were elaborated and re-introduced by Maturana & Varela (1980), Varela et al. (1991), and Thomson (2007), among other authors. The central claim is the continuity between mind and life, where cognition is rooted in direct embodied engagement with the environment.

infants' interaction with their surroundings (Flynn et al, 2013). Recent studies in computational neuroscience agree that multimodal-perception action loops can drive neural change and connectivity (Lugarella et al., 2005; Lugarella & Sporns, 2006). From this perspective, learning and development are accrued products of the real-time interactional events that, in turn, arose from perceiving and acting in a physical world (Sheya & Smith, 2010).[50] The "sticky mittens" are a famous example of experimentally controllable contexts in which developmental change can be created and studied. In this study, researchers placed Velcro-covered ''sticky mittens'' on the hands of infants not yet capable of actually grasping objects. The mittens allowed the infants to ''pick up'' toys by swiping at them, thus precariously coordinating vision and reaching. Researchers also found that use of the mittens later increased the infants' ability to explore objects, thereby not only facilitating the development of reaching for objects but also of visual–oral exploration (Needham, Barrett, & Peterman, 2002). This shows that the coordination of tasks involving different capacities or subsystems (like seeing, reaching and orally exploring objects) has cascading effects in other tasks in which some of the same subsystems are involved (Sheya & Smith, 2010). These overlapping coordinations may also be responsible for the cascading interactions characteristic of human development, wherein even seemingly distant achievements may be developmentally related (see Smith & Pereira, 2009). This may explain why babies' multisensory exploration of a rattle and the encounter with its sound properties display activities directed at the sonic possibilities of the object at hand, just as the babies' experience with the mittens in Needham's experiment leads to the coordination of multimodal experiences and the development of new types of actions (Schiavio et al, 2017). For Schiavio and

---

[50] Many of these ideas were inspired in Gibson's "ecological psychology". The basic claim of Gibson (1986) is that organisms do not perceive the environment in a neutral way, but rather they perceive the properties of objects as opportunities for action: in other words, *organisms perceive what objects afford. We do not perceive a thing with a certain shape and then attribute to it the function of hammering, say, but rather we perceive the artifact, a hammer*. While sympathetic with this general approach, some scholars suggest that the classical Gibsonian notion of affordances implies that they are intrinsic features of the environment, which does not give enough attention to the active role living creatures play in shaping they worlds they inhabit (for a discussion on this subject see Chemero, 2009). In this thesis, I argue for a view of cognition that is not wholly driven by the environment, but rather by the embodied activity of living agents.

colleagues, this case study reveals that sound resources can be seen as possibilities for action resulting from self-organizing and creative avenues through which infants explore and interact with their sonic environment. For these authors, music in early infancy is an emerging property of the ongoing relation between the infants and their environment, it is a cultural trait that emerges when infants actively engage in exploratory behaviours involving sound-related outcomes.

I suggest that these overlapping coordination of multimodal perception-action loops give rise to flexible assemblies of interdependent clusters of capacities, artifacts and cultural practices underlying the construction of cultural niches through which this conglomerate of cognitive capacities, artifacts and cultural practices are cultivated and transmitted. In the next section, we will see that these interdependent and dynamical processes most likely were enhanced by our interaction with the material world over the course of our evolutionary history.[51]

### 4.4. Niche construction and material culture

The emergence of scenarios for social learning was most likely one of the most significant ways through which cultural niches may have been constituted.[52] Sterelny (2003) points out that the modification of the environment may have allowed our ancestors to develop strategies for social learning and social organization.[53] Spoken language should not be considered as the principal means through which these scenarios for social learning were initially constructed. Learning abilities may have been potentiated by a wide diversity of social channels (Sterelny, 2003).[54]

---

[51] Iriki & Taoka (2012) propose the concept of "triadic niche construction" to study the evolutionary impact of tool use in accelerating interactive links between ecological, neural and cognitive features over the hominin evolution.

[52] Sterelny (2012) argues that once individuals developed a sense of identity to a group, then social learning become a key evolutionary feature of our species.

[53] For Sterelny (2003), this view is in tune with the idea of an extended mind (Clark & Chalmers, 1998; Johnson, 1987).

[54] According to Roger Bartra, from caves to what he calls "multifamiliar hives" of contemporary cities, humans have been constructing artificial environments not only as an external protection but also to deposit an ample variety of signals and marks. These marks and signals constituted our "microcosms",

Donald (2009) argues that art is one of the mechanisms that allowed us to store relevant information on our environment. He considers that Franco-Cantabrian paintings of the Upper Paleolithic may have served as a material device for transmitting ideas across generations. For him, the paintings of animals and hybrid images (human-animals) manifest a religious view of the world, particularly an animistic perspective, as shown in the artistic manifestations from the Early Stone Age. Donald claims that Upper Paleolithic paintings constituted an important way through which myths, stories, and allegories were transmitted across generations. Additionally, he hypothesizes that the paintings also captured the way through which these early humans may have conceived the meaning of life. From his view, the paintings served as a device for social structure and cognitive regulation supporting the development of a particular conception of the world (Donald, 2009).[55]

It is important to note that, conceiving paintings as depositories of relevant information, as Donald proposes, implies a sharp distinction between mind and materiality. Thus, paintings are understood as the materialization of pre-existing ideas located inside human minds. As such, the material environment is the place where several ideas (conceived as information contained in the brain) has been deposited. This traditional distinction between materiality and mind is also commonly used to explain the emergence of human symbolic behavior.

Contrastingly, several scholars argue that the emergence of the conceptual system, commonly associated with our mental representations, is grounded in patterns of physical and sensorial interaction with the material world (Renfrew, 2012, Lakoff and Johnson, 1980, Lakoff and Johnson, 1999). This means that, evolutionarily speaking, the emergence of our abstract concepts, for example, the ones we use to describe spatio-temporal aspects of the

---

which has been externalized as life styles, religious systems moral habits, etc. For Bartra, the entanglement of signals constituted our "cognitive prosthesis" that complement our brain circuits. He states that these "texturas simbólicas externas" are regularly used by our neuronal circuits in such a way that these neuronal circuits could not properly function without these prosthesis (Bartra, 2013). Bartra called "exocerebro" to this ample diversity of "symbolic extra-soma circuits". According to him, neuronal activity can properly function once it is complemented by a "cultural prosthesis" (Bartra, 2012; 2013).

[55] In a similar way, Flynn and colleagues (2013) claim that the Upper Paleolithic paintings may have enabled the generation and propagation of ideas that supported the structuring for social behavior.

environment, most likely were based on the development of the sensory-motor system (Gallese and Lakoff, 2005). This urges to rethink our understanding of material culture and to examine the way it may have played an important evolutionary role in cognitive evolution. Ingold (2007) rightly points out that a better understanding of material culture requires us to focus on materials and skilled transformation of materials which are not necessarily shaped by mental and semantic representations. This reveals the need for developing dynamic and multi-directional accounts to explain the evolution of the cognitive capacities involved in several material transformation processes.

Given the above, Upper Paleolithic paintings could be understood as resulting from a wide range of intertwined social practices that may have enhanced the stabilization of different cultural traditions over time. Recent archaeological studies have taken into account both the material characteristics of the macro-landscapes where the Upper Paleolithic paintings are located, and the specific material qualities of the caves and rocks. These studies have shed important light on the relevant role that our interactions with the material qualities of the environment (e.g. texture, luminosity, and sound) may have played in structuring and stabilizing a wide diversity of cultural practices. Accordingly, it is problematic to assume that the multi-sensorial experience provided by the material environment was predominantly perceived through vision (Boivin 2011). Instead, it is most likely that the multiple ways of interacting with the environment may have been fostered by a wide range of sensorial means such as touching, smelling, and hearing, among others (Houston and Tabue, 2000; Scarre and Lawson, 2006). Fahlander and Kjellström (2010) state that sound is often forgotten in our images of the past, although it is often a vital component of any place. In this regard, Jacopo Moggi-Cecchi analysed a fossilized ear stirrup of a specimen known as Stw 151 (from Sterkfontein, South Africa). He found that the size and appearance of this fossilized ear bone resembles more closely to modern apes' ear stirrups than the modern humans one. For Mithen (2006), this indicates that early hominis were much more sensitive to sounds of higher frequencies than modern humans are. He urges for an archaeological exploration of our sound environment.[56] Several researchers think that a set of numerous rituals constituted by collective practices (such as singing and dancing) took place inside the painting areas of the Upper Paleolithic caves (Donald, 2009; Lewins-Williams, 2009). Therefore, these locations have been characterized as "areas of activity" (Lewins-Williams, 2009).

All the above leads us to think that the many ways through which our ancestors interacted with the material qualities of their environment may have enhanced the development of a considerable number of practices that supported the structure of our cultural niches over our evolutionary history. Several scholars consider that the perception and manipulation of the sonic environment may have played an important role in the evolution of a wide range of our perceptual and cognitive capacities (Scarre and Lawson, 2006).

In the next section, I will focus on the way in which the acoustic exploration and alteration of the material environment may have fostered the structure of distinct cultural practices that diversify the scenarios for social learning through which a wide range of cognitive capacities and motor skills were cultivated and transmitted.

### 4.5. Acoustic qualities and material culture

Acoustics is a subdiscipline of physics interested in the study of sound production, transmission, storage, perception and reproduction. Acoustic studies have shown that the way we perceive sound depends, among other things, on the physical environment through which sound is produced and propagated. Some physical environments may increase the intensity of sound and prolong its duration while other environments do not. For instance, a room surrounded by cement walls does not produce the same acoustic results as one surrounded by cardboard; the sound produced in the latter space is more resonant and loud than in the former. Likewise, the sound quality of an acoustic guitar varies depending on the guitar´s shape, the material with which it has been made, as

[56] Several years after the publication of *The prehistory of mind* (1996), Mithen recognized that not having included music as an important component of the cognitive evolution of our species was the main mistake and weakness in his work. In his book of 2006, *The singing Neanderthal. The origins of Music, Language, Mind and Body*, he corrected his previous error.

well as the place where the guitar is played (inside a tent, in a concert hall, outdoors, etc.). Similarly, singing or speaking inside a large cave,a narrow tunnel, or a house changes the levels of voice's resonance.[57]

In recent years, there has been an increasing interest in studying the possible use of sound during the past. The archeology of sound, also known as archaeoacoustics, is an interdisciplinary field in which some of these findings have begun to be integrated.[58] One of its aims is to explore the acoustic universe of the Upper Paleolithic period (Reznikoff & Dauvois, 1988). For instance, some scholars have tried to study the ways in which our ancestors may have experienced the sound qualities of both the caves and their own voices by analyzing the acoustic qualities of vocal sounds produced inside these areas. Given the remarkable acoustic qualities of these locations, several scholars argue that the paintings were made after early humans discovered those acoustic properties (Waller 1993, 2002; Reznikoff & Dauvois 1998; Reznikoff, 2005, 2008; Dauvois 1989). Reznikoff (2005, 2008) and Waller (1993, 1987), claim that the acoustic resonance of these locations resembles the kind of sounds that the animals represented in the paintings could have produced in the wild. Thus, they hypothesize that the acoustic qualities of the cave led early humans to make the paintings. Waller (1993, 1987) suggests that early humans conferred a mythical meaning to the remarkable acoustic resonance of these painting areas. For him, the paintings and the acoustic qualities of these locations contributed importantly to the conformation of an early animistic

[57] This is a minimal characterization of some research topics in archaeoacoustics. From a couple of decades ago, this field of study has been significantly increased in such a way that, currently, it is conceived as a interdisciplinary domain that encompasses a wide variety of scientific interests coming from a vast diversity of disciplines such as physics, mathematics, psychology, neurology, and musicology (Roederer 1995).

[58] Archaeoacoustics is an interdisciplinary domain that predominantly integrates research interests in archaeology and acoustics. Its principal focus has been to develop a scientific idea of how the past may have sounded, and which were the possible uses and meanings that ancient human cultures attributed to these sounds. Acoustics studies include Paleolithic caves, ancient Roman and Greek temples, rocky outcropping, mountains, churches and medieval theaters, etc.

conception of the universe, which in turn gave rise to the development of several ritual practices.[59]

Waller studied not only the acoustic qualities of several caves but also of approximately 100 outdoor sites of painted rocks. These studies were conducted in France, Australia and the United States. He obtained the same results in all of them; most of the painted rocks were located in areas of the most outstanding level of acoustic resonance.[60] For Waller (2002), it is impossible to know what our ancestors had in mind when they made these paintings. He also accepts that we will never know whether or not the acoustic resonance of certain locations may have induced early humans to produce certain kinds of paintings. Along this line, Scarre and Lawson (2006) state that it is impossible to know whether the acoustic qualities of the caves were discovered and manipulated before the production of the paintings. However, they claim that once remarkable acoustic qualities of any material environment were detected, it would most likely have been extensively explored and exploited. Likewise, Blake & Cross (2008) state that, every time individuals produce sounds, whether intentionally or unintentionally, this activity impacts the way the members of a group relate with each other and with other groups and species. These authors remark that "the study of sound and its relevance in human behavior should not necessarily be a search for acoustic features, but rather for the activities carried out that were conditioned by sensitivities to and uses of sound from both biological and cultural perspectives" (Blake & Cross; 2015: 91).

In the next section, we will see that the exploration of the acoustic qualities of the environment may have triggered numerous ways through which our ancestors created and modified their environment, and by doing so, significantly

---

[59] Waller (1993) points out that several cultures produced mythical explanations concerning the origins of echo. For instance, Greeks believed that on the top of a mountain, this kind of sound was emitted by the nymph Eco.

[60] It has been shown that in Scandinavia, a significant amount of stone art dating back from the Neolithic period was located next to waterfalls and rivers (Goldhahn, 2002). These objects may have been deliberately placed in those regions given the type of multi-sensorial experience provided by this kind location (Goldhahn, 2002). Furthermore, in Southern India, rocks dating back from the Neolithic produce different types of timbres and pitch when they are struck, which suggests that they may have been used as music instruments (Boivin, 2011).

impacted the diversification of their cultural practices and their cognitive evolution.

### 4.6. Sound production and spatiotemporal orientation

One way through which some researchers have examined the acoustic qualities of the Upper Paleolithic caves is by producing vocal sounds in the interior of the painting areas. According to Reznikoff (2008), once the vocal sound is produced, it travels throughout the cave and the experimenter's body in such a way that the experimenter also feels in his own body the acoustic properties of the surroundings.[61] Despite issues concerning the empirical relevance and methodological basis supporting this experiment, I want to highlight the way in which this exercise can exemplify an embodied experience of sound perception.

Reznikoff suggests reproducing this experience at home. It consists in placing our left hand on our chest and our right hand on our head. Then, to pronounce, or even better, to sing the vowel A and the consonant M alternatively. By doing so, we can feel how the sound vibration travels from our chest to our head and vice versa. We can also sing the sounds A, O, U and M, alternatively. In this case, we will feel how the sonic vibration moves from the chest (A), to the throat (O); then from the chin (U), to the top of the head (M). These simple exercises show how certain sonic vibrations travel differently throughout the body.[62] Reznikoff (2008) argues that the immersion of the sound in our bodies can be felt before we are born, it means that prenatal body feels the vibration of mother's voice and her heartbeat. It has been claimed that at that stage of life, we already perceive external sounds such as voices of other people, dogs' barking, sounds of musical instruments, sounds of different objects, etc. Thus, hearing allows the prenatal fetus to establish sonic relationships with an external world before it can be visibly perceived. Hearing also allows the fetus to become familiar with an initial sense of space by perceiving the physical proximity of the objects: near or far, up or down, behind or in front of, etc. There is evidence that newborndeaf children have spatial orientation problems (Reznikoff, 2008).

---

[61] See Reznikoff (2008) for the methodological details of this experiment.

[62] Reznikoff recommends to emit lower pitch sounds to intensify this embodied experience of sound.

This leads us to think that the initial stages of space perception are not founded on a conceptual framework previously developed in our minds. Several scholars have pointed out that our conceptualization of space and time is grounded on, and closely connected with, an embodied orientation and understanding of the world (Lakoff & Johnson, 1980; Gallese & Lakoff, 2005).[63] Importantly, the notion of spatiality has also been used to characterized musical scales as a set of sounds that "ascends" or "descends" during certain period of time. This way of organizing sounds resonates with the experience of the sound traveling throughout our body in Reznikoff's experiment.[64] The sequential order in which these sounds are heard also reflects the order in which they are being produced and bodily experienced. Thus, if we would only hear the first couple of sounds, then we would not be capable to hear the whole sequence.

Consider the following example of Gallagher, Martínez & Gastelum (2017). In listening to a familiar melody, there is some sense of what is to come, a primal expectation of the notes to follow. Then, if someone hits a wrong note, we are surprised or disappointed. In the same way in which a person fails to complete a sentence, we experience a sense of incompleteness as our expectation is unfilled, or what we experience fails to match our anticipation. Trehub (2001) has shown that music perception skills of prelinguistic infants are surprisingly similar to those of listeners who have had years of informal exposure to music. For example, infants also recognize the invariance of melodies across shifts in pitch level (transpositions) and tempo (Trehub, 2000). This reveals that, from a very early stage of lives, hearing discrimination plays an important role in structuring the intrinsic temporality of retention and anticipation. Gallagher, Martínez & Gastelum, (2017) point out that if this experience were of only one moment at a time, without experiential connection to previous moments, it would be impossible to make sense of the world. Furthermore, these authors claim that

---

[63] Lakoff and Johnson (1980) claim that, in may contexts, the notion "up" refers to bodily experiences such as: standing up, being well, feeling happy, being alive, etc. Thus, terms as "below" and "down" refer to opposite experiences. They also argue that this metaphorical meanings are grounded on the way we orient ourselves in the space, which also plays an important role in the way we refer to different aspects of the world.

[64] Including the experience of hearing sounds traveling in our surroundings.

our intrinsic temporality of retention and anticipation also helps to structure movement and action. Gaze anticipates the rotation of our body when we turn a corner (Berthoz, 2000). Similarly, in any performance situation—particularly improvised music—performers must hold open multiple possibilities for "next steps" in a scenario, dependent upon what has just gone on before and in anticipation of what may happen next. Thus, performing musicians react and adjust in real-time to very subtle modulations created by others (and also by their own actions). Indeed, the expressive nature of ensemble musical performance is contingent upon this capacity (Moran, 2013). Importantly, bodies have a natural tendency to move with music, which also provides a very rich bodily experience of time (Johnson, 2007).

As we can see, it is very likely that diverse mechanism of sound exploration and production played a key role in scaffolding the evolution of our cognitive capacities of spatiotemporal orientation. Furthermore, these embodied explorations of sound production may have also fostered the evolution of sensory-motor abilities which, in turn, may have substantially contributed to the development of cognitive capacities for collective coordination involved in a wide diversity of cultural practices. All the above allows us to think that the capacities involved in music making most likely co-evolved alongside a wide range of cognitive capacities that scaffolded the development of diverse cultural practices based on complex levels of collective synchronization and coordinated actions. If we consider that a wide range of our cognitive capacities co-evolved interdependently, then the stabilization and complexification of an ample diversity of cultural practices depended on the construction of cultural niches constituted by multiple and intertwined learning scenarios, through which these practices and capacities were cultivated and transmitted across generations. For a better understanding of this co-evolutionary processes it is necessary to develop a characterization of music as an active multimodal phenomenon with a cross-cultural orientation. That is the aim of the following section.

## 4.7. Towards a cross-culturally oriented and multimodal characterization of music

In the next sections, we will see that a cross-cultural perspective of music showcases the explanatory limitations of the conservative and exclusively Western perspective, and sheds important light on the ample range of cognitive capacities and social skills that this cultural practice involves.

### 4.7.1. The worldless picture of music

Understanding music as "the art of combining sounds (voices and instruments) and silences to produce harmony and beauty" is one of the simplest and widely spread definition of music; it is often found in conventional dictionaries of Western music. This definition presupposes a Western and paradigmatic understanding of beauty which excludes a wide variety of musical forms and practices around the world.[65]

In Martínez & Villanueva (2018b), we show that Western conceptualization of music endorses an understanding of music as high culture, as opposed to social life and material culture; it is a cultural achievement that can be studied independently from the social-material environment in which it is produced. This view of music was built upon 19th-century aesthetics, according to which music stands by itself, beyond any human utilitarian purpose (Dalhaus, 1978/1989; Fubini, 2008; Hegel, 1975; Kant, 2000/1790; Schelling, 1999; Schopenhauer, 1969). Since then, this disembodied model of music has dominated the landscape of Western art music, conceived of as a pleasurable auditory experience passively perceived by an audience. This approach has been widely reinforced by contemporary philosophers of music. For instance, Peter Kivy argues that the liberating power of music comes from 'its complete freedom from connection with our workaday world and its problems' (Kivy, 1997, p. 209). Similarly, Alvin Goldman (1992), defends that the detachment of music

[65] See Tomlinson (2001) for a critical review of the way Western culture has historically conceptualized music.

from the real world is the source of the great value we place on music and musical experience.

This worldless picture of music does not capture the situated character of our most common musical experiences. In early 21st century, contemporary musicology argued that all the meanings attributed to music have been historically and socially constructed (Cook, 2001). Nevertheless, in current musicological and philosophical literature prevails the idea that the great value we place on music is based on the pleasurable auditory experience. This implies that listening is the solely way through which music can be aesthetically valued.

However, music is typically embedded in other domains of social action; it is rarely an end in itself, but a component of other activities (Cross, 2012). What counts as music can vary among musical cultures (Cross 2001; Trehub, Becker & Morley 2015; Rice 2014). Likewise, what counts as an aesthetic experience of music is not an end in itself that transcends the conglomerate of social practices in which music takes place. Neither it is something passively perceived by a static audience. Instead, it is the result of an actively engagement with a socio-cultural, normative, participatory, and highly emotional environment.

#### 4.7.2. Music making from a cross-cultural and socio-material perspective

Ethnomusicological studies have convincingly shown that music has always existed in all known human societies (Merriam: 1964; Blacking: 1974; Cross: 2001; Rice, 2014). The many ways in which music has been characterized around the world shed valuable light on the complexity of this cultural practice. The Igbo people of Nigeria use the word *nkwa,* meaning singing, playing a musical instrument, and dancing (Cross: 2001). Bantu speakers in East Africa use the word *drum* (*ngoma)* for singing and clapping. They also use the same term referring to the laments of people in specific social contexts. In some regions of Bulgaria, the term music (muzika) refers to what we understand as instrumental music, they use other terms for singing and drumming. In some social contexts of Islam, the notion of music refers to singing and musical performances that take place in secular events. The kind of music they use in a sacred environment is referred to as "chants" or "recitations" (Rice, 2014). In the

Huasteca region of Mexico, the term *huapango* (which could be roughly translated into English as "dance on wood") refers not only to music and singing, but also includes the items involved in a social gathering such as food, drinks and dance. Thus, in the Huasteca region this term is equivalent to what in other regions of Mexico is called *fandango* (Hernández 2000; Sanchez 2002).

On the other hand, the lexeme *K'ehoh* in Tzotzil language refers to both a "song with lyrics and music" and to an instrumental melody. In the latter, it is assumed that melody contains "words" implicitly (Nava, 2010). The phrase purépecha *kústatarakweechaksï piresïti* can be roughly translated as "the instruments sing". Similarly, the phrase *warhiitiichaksïo piresïnti*, also in purépecha, can be translated as "women songs". What can be noticed in both phrases is that the verb *piresïnti* is used to refer to any melody no matter whether it is produced by a human voice or by a musical instrument, the fact is that "both of them sing". The same happens with the word *rol* in Zapoteco, which means to sing or to play an instrument (Nava, 2010).

A cross-cultural approach helps to demonstrate that components and uses of music go beyond the Western traditional view. Western music, for instance, exploits a dynamic and binary perceptual distinction between sounds that are consonant and those that are dissonant so as to articulate musical structures in time. Other cultures exploit the consonance-dissonance distinction in different ways, as in the music of the *campesino* culture in Northern Potosi (Stobart, 1996). Similarly, Indonesian gamelan music does not employ the binary distinction at all (Perlman, 2004). Ethnomusicological studies have also shown that in cultural contexts where people work outdoors (e.g. *pastoral* populations), human songs become part of the varied sonic environment produced by other animals in the wild (Rice 2014). Likewise, the Kaluli of Papua New Guinea do not establish a sharp distinction between birdsongs and the sounds produced by their musical instruments (Feld: 1982). For them, birdsongs are also an important component of the musical environment. Several indigenous communities in Mexico have incorporated the sounds of animals into their musical melodies. Mixtecos imitate the roar of bulls with their instruments. A wide range of reed flutes played by Yaquis, Zapotecos and Totonacos, among many others, imitate birdsongs (Nava, 2010). There are also a large variety of instruments that resemble the sound of rivers, waterfalls, wind, thunderstorm, raining, etc. In human settlements surrounded by mountains, people have used the acoustic qualities of echo to create channels of communication based on whistles, screams and songs. Thus, it has been documented that Mazatecos have developed a complex system of communication based on whistling. This Mexican indigenous group takes advantage of echoes as well as the tonal structure of the Mazateco language (Nava, 2010). Similarly, many musical instruments such as snails shells and drums have been widely used for communicative purposes.

In Alaska and the Canadian Arctic, the Inuits' vocal games, locally designated as *nipaquhiit* (which means games done with sounds or with noises) consist in alternating rhythmic patterns of glottal sounds, sounds of musical instruments, and gasping respirations (Soh Fujim Record oto, 1993). Yodel traditions of distinct regions of Switzerland, Austria, Africa and Oceania, consist in changing vocal registers in a few seconds, alternating exhaled and aspirated sounds by singing the syllable *yo* (Record of Gerard Kubik, 1967). Screams are worldwide considered as an important component of music performance. In carnival rituals, screams are interspersed between the sounds of musical instruments in order to guide the development of a musical performance, and also to reinforce the emotional level of the participants. Another interesting case is the way George Benson intersperses vocal sounds with the music he plays in the guitar. Benson's performances evoke certain traditional music from Laos where vocal sounds get almost completely mixed with the sounds of the musical instruments (Record of Jacques Brunet, 1992). This also resembles the way the Mongolian limbe transverse flute is played: interspersing vocal sounds imitating the sound of the flute (Robert Hamayon Registro, 1968).

As we can see, these characterizations and uses of music entail a broad ambit of what "music" encompasses. A cross-cultural perspective of music also sheds important light on the ample range of abilities involved in music, which in turn are based on our interaction with the socio-material environment. In Martinez and Villanueva (2018b), we argue that the abilities displayed by music only come into existence because the environment offers the possibilities for actions that it does.[66] For example, many musicians play the violin by following the fingering technique and body posture usual in Western Classical music. But violins can also afford different ways of being played, as shown in the southern Indian Classical Carnatic style, where each type of *gamaka*[67] reveals a distinctive violin fingering technique and body posture (Swift, 1990; Weidman, 2012). Similarly, during the 1970s, in certain indigenous areas of the Huasteca region of Mexico, violinists sat in a chair and played the violin by holding it between their legs or knees (Villanueva, 2012). Moreover, it is quite common for blues and rock guitarists to use unorthodox playing techniques, chording and fingering.[68] Tëmkin (2004) also points out that an ancient fingering technique for playing the gusli (the Russian version of the psaltery) prevailed among the new musicians who replaced the gusli for the accordion.

Thus, the materiality of instruments also allows, among other things, for the acquisition of musical skills related to how it feels to hold and play an instrument (Martínez & Villanueva, 2018b). More generally, this reveals that the interaction-exploration of the surroundings also triggers a set of social affordances ("taskscapes," in Ingold's terminology) that emerge from the activities of a social group as a whole, the movements required to undertake these activities, the material properties of the environment, and the sounds produced (Ingold, 1993). Nowadays, it is widely accepted that music affords coordinated bodily practices such as foot-tapping, marching, dancing, and clapping (Clarke, 2005; Krueger, 2011; Reybrouck, 2005). ''Music is not simply something that is heard and consumed, it is something that is done in interaction with others'' (Cross, 2008, p. 151).

---

[66] See also Rietveld & Kiverstein (2014).

[67] Melodic ornamentation that characterized Carnatic raga such as slides, deflections, and fingered stresses (Swift, 1990).

[68] It also happens as when left-handed guitar players have to restring their instruments to be able to play, like Albert King and Otis Rush used to do (Lilliestam, 1996).

All the above leads us to conceptualize music as an activity, as music-making instead of a final and static sonic product.[69] Turino's (2008) distinction between presentational and participatory social fields of music is relevant for this purpose. In the *presentational field*, performances have the responsibility to prepare and provide music for the audience who does not participate in producing sounds or motions that are deemed fundamental to the performance. In the *participatory field*, there is not a formal artist-audience distinction, only participants and potential participants. In the presentational field, music is seen as a medium for display, as an aesthetic entity standing on its own. In the participatory field, music is understood as a medium for social interaction (Turino, 2008).

In this work, I conceive music-making[70] as a participatory field and understand it as *an emerging and culturally embedded participatory social practice scaffolded by clusters of capacities, skills, artefacts, and social norms, involving complex patterns of interaction-exploration between individuals and their sonic and material environment.* Accordingly, music-making does not presuppose any abstract rule transcending the conglomerate of cognitive capacities, skills and social practices through which—and upon which—music takes place. This account of music resonates with Johnson's (2007) claim that music exists in the intersection of organized sounds, our sensory-motor system, our bodies, our brains, our cultural values and practices, our historical-cultural conventions, our previous social experiences, and a long list of further socio-cultural aspects. Likewise, Cross (2001) points out that any attempt to characterize any general attribute of music must recognize its corporeal nature, and the indivisibility of movement and sound.

---

[69] Christopher Small (1998) coined the word ''musicking'' to define music not as a thing but as an activity. This notion attempts to capture a great variety of activities involved in making music, from composing, listening, and playing an instrument, to body movement and giving meaning to sounds, among others.
[70] Hereinafter, I will use music and music making interchangeably.

### 4.7.3. Music-making as a multicomponent engine for social interaction

The way I conceptualize music-making allows us to see that this cultural practice is a complex, multifunctional and multifaceted cultural phenomenon. Ethnomusicologists and musicologists agree that music-making encompasses multiple interactive components. This idea is familiar from music theory; Western music has commonly been dissected into the separate components of rhythm, melody and harmony. However, this traditional breakdown has not to be considered as necessarily appropriate from an evolutionary perspective, or that these components are themselves monolithic capacities (Fitch, 2018).[71]

It has been thought that the capacity to synchronize our musical behaviours with others is one of the basic components of music making. This requires the individual capacity for synchronization to some external time given. Fitch (2018) claims that the most sophisticated form of synchronization involves beat-based predictive timing. This capacity to extract an isochronic beat and synchronize to it is commonly called beat perception and synchronization or BPS (Patel 2006). This capacity is clearly connected to our intrinsic temporality, of retention and anticipation. As we saw in the previous section of this chapter, hearing discrimination plays an important role in structuring intrinsic temporality which also allows to structure our movements and actions (Gallagher, Martínez & Gastelum, 2017). Although the majority of research in both human and animal have studied BPS using a metronome or recorded musical stimuli, it is most likely that human rhythmic abilities did not arise to allow people to synchronize to metronomes but rather to the actions of other humans in groups (Fitch, 2018). Our capacity for detecting beat in rhythmic sound sequences is already functional at birth (Winkler, Háden, Ladinig, Sziller, & Honing, 2009), and it has been shown via musical experimental studies that the auditory and the motor systems are completely integrated (Zatorre, Chen, & Penhune, 2007). This capacity is also considered as a human universality in the sense that in every human culture people regularly gather in groups for pulse-based rhythmic drumming, singing

---

[71] Fitch (2018) invites us to accept that there is not one "true" or "correct" breakdown, but different componential breakdowns might be appropriate for different research purposes.

and dancing, often in ritual settings. Note that dancers also make use of the synchronization abilities to synchronize with music and with other dancers. The distinction between music and dance is not made in many languages around the world, since music and dance are conceived together to comprise a common mode of human interaction (Nettl 2000; Cross 2001). A close link between music and dance is also evident in most European music outside the concert hall, and although in those cases dance may be distinguished from music, it is almost always accompanied by it. Importantly, dance, as a cultural manifestation constituted by communicative body movements, includes sound production (Fitch 2018). Furthermore, the one form of instrumental music that is very nearly universal is the use of percussive instruments: ideophones and drums (Nettl, 2000; Savage et al, 2015). For Fitch (2018), the core of most human dancing is not strictly distinguishable from drumming. In this case, tap dancing, flamenco dancing, and traditional dances of Mexican *son*, for instance, constituted both dancing and drumming simultaneously.

Imitation is also an important pillar of music making. It enhances our synchronization abilities by providing high-fidelity copy of the body actions of others (Trehub, et al 2015). Likewise, imitation allows a competent couple of group dancing to coordinate the actions of individuals, and in the process matching, reversing or complementing each other (Laland et al., 2016). Imitation enhances the integration of visual, auditory and motor systems. This allows musicians and dancers the acquisition of long sequences of movements within complex scenarios of social learning.

Given all the above, in this work I suggest that music making is grounded on an engine for social interaction constituted by clusters of cognitive capacities and skills such as: sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization.

In evolutionary terms, I hypothesize that this cluster of capacities and skills involved in music-making may have emerged as a result of a prolonged and extensive diversification of patterns of interaction-exploration between individuals and their socio-material environment. Paradigmatically, the importance of toolmaking for an explanation of human cognitive evolution has been widely accepted. However, it is usually assumed that its importance is reducible to the role tools have in allowing us to perform certain tasks that have adaptive value, like hunting. But the relevance of toolmaking and tool use in human evolution goes beyond its role in specific adaptive tasks (Martínez & Villanueva, 2018b). The ancient processes of tool-making could have required a considerable refinement of motor control that led to complex sequences of muscle movements and perhaps synchronized patterns of knapping (D'Errico et al., 2003; Tomlinson, 2015) grounded in the capacity for synchronizing to an external beat. Humans do not need special training to detect and motorically synchronize to an external beat (Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2018). This is a human capacity that may have ancient roots in our hominin lineage (Martínez & Villanueva 2018b, Tomlinson 2015; Morley 2013). Toolmaking might also have fostered the development of specific forms of listening, discrimination, and the production of certain acoustic properties of stones (Blake & Cross, 2008; Cross, Zubrow, & Cowan, 2002; Killin, 2016a, 2017; Morley, 2013; Zubrow, Cross, & Cowan, 2001), which in turn may have enhanced knapping expertise.

This possible multimodal scenario reveals that our capacity for making music may have been cultivated as part of complex processes of co-construction and co-evolution of niches of abilities such as the ones already discussed. It is most likely that, over time, these kinds of interactive processes may have given rise to an ample spectrum of cultural practices that emerged by consistently recruiting, sharing and combining different clusters of cognitive capacities in multimodal learning scenarios. This resonates with the claim that, since music consists of a multiform, multifaceted, and complex phenomenon, its origins should be explained as a part of co-evolutionary processes of behaviors and capacities (Killin, 2016a, 2017; Tomlinson, 2015; Van der Schyff & Schiavio, 2017). The elaboration of this co-evolutionary narrative is the aim of the next chapter

# CHAPTER 5. MUSIC MAKING AND SOCIAL INTERACTION: A CO-EVOLUTIONARY APPROACH

## 5.1. Introduction

The aim of this chapter is to develop an evolutionary account for explaining the possible processes through which a set of cognitive capacities and skills supporting our capacity for making music may have become dynamically integrated. I will suggest that music-making is grounded on a social interaction engine[72] constituted by clusters of cognitive capacities and skills such as sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization. I will hypothesize that the evolution of this cluster of capacities and skills should not be linked exclusively to our musical capacity, but rather, this cognitive engine should be seen as the result of a prolonged and extensive diversification of patterns of interaction-exploration between individuals and their socio-material environment. I will delineate the possible co-evolutionary processes through which this cluster of cognitive capacities and skills may have become integrated.

## 5.2. Tool-making

It has been maintained that the emergence of speech and symbolic thinking played a decisive role in our cognitive and cultural evolution. However, archaeologists and primatologists are increasingly interested in studying the manufacturing of tools as a way to have a better understanding of how our cognition may have evolved (see e.g. Wynn, 1991; 2002, Wynn & Coolidge, 2010; Yamamoto, S. et al, 2013). It had been accepted that the construction of stone tools arose alongside the emergence of genus *Hom*o. The oldest archaeological register of stone tools associated with the genus Homo -the lithic

---

[72] See Levinson (2006).

industry known as Oldowan or Mode 1- dates back to around 2.6 million years ago (Lewis and Harmand, 2016). Nevertheless, in the West of Lake Turkana in Northern Kenya, there have recently been discovered stone tools (cataloged as LOM3) dating back 3.3 million years ago — 700,000 years earlier than the Oldowan industry. This means that Mode LOM3 was developed before the emergence of the genus *Homo*, which in turn reveals that the evolutionary origins of these capacities goes beyond the origin of our species. In the next sections, I will present a brief survey of the lithic traditions that have been archaeologically reported. Then, I will discuss to what extent the morphology of the tools and their possible manufacturing processes can shed important light on the evolution of our cognitive capacities and skills such as: motor control, spatial-temporal orientation, nonverbal communication systems, sound production and discrimination, and patterns of imitative behaviour (Wynn, 1991; Coward and Gamble, 2008).

### 5.2.1. Mode LOM3

Several replication experiments suggest that the LOM3 knappers were predominantly using both the passive hammer technique, in which the core is held in both hands and struck downwards onto an anvil and the bipolar technique, in which one hand stabilizes the core on the anvil and the other strikes the hammer down vertically onto the core (Lewis and Harmand, 2016). However, the multiple percussion traces on the stones lead archaeologist to hypothesize that the manufacturing processes may have involved more steps than the ones mentioned. See the image below.


**Figure 3.** *In situ* unifacial core (LOM3-2012-H18-1, 3.45 kg.) Bipolar technique. (Taken from Lewis and Harmand, 2016).

### 5.2.2. Mode 1

Oldowan stone tools, also called Mode 1, dates to 2.6 million years ago, approximately. The first area in which this archaeological culture was documented is the site of Olduvai Gorge in Tanzania, Africa. Archaeologists think that a hammerstone may have been used to strike on the edge, or striking platform, of a suitable core rock to produce conchoidal fractures with sharp edges. The chip removed is the flake, which consists of a stone with one sharpened edge. Archaeologists consider that these stone tools were used for cutting and removing skin (or meat) of the animals captured for consumption (see the images below taken from Fuentes, 2007). The manufacturing process of these tools may have required the acquisition of more complex skills. Striking the stones precisely on the exact place may have required high level of sensorimotor control and visual-spatial accuracy (Lewis and Harmand, 2016). See the images below.

**Figure 4**. Oldowan tools. The stone shown in the upper left was used unmodified; it is a little larger than a tennis ball. At the bottom right are two flake tools. The remaining tools shown are core tools from which flakes have been removed, leaving sharp edges. (Taken from Fuentes, 2007).

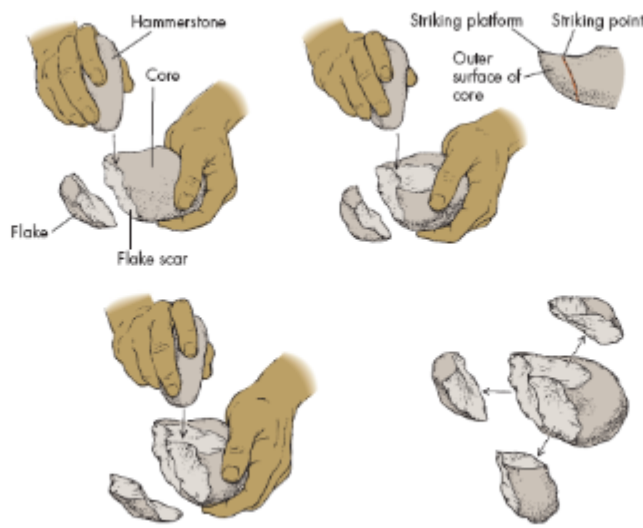



**Figure 5**. Making Oldowan tools. The toolmaker strikes a hard stone in just the right place to remove thin, sharp, flakes. (Taken from Fuentes, 2007).

### 5.2.3. Mode 2

Acheulean industry or Mode 2 dates back to around 1, 7 million years ago. This lithic industry did not replace the former; Mode 1 never completely disappeared. Archaeologists consider that the Mode 2 industry first appeared in East Africa, and then spread throughout different areas of Africa, Asia and Europe. It is widely accepted that this lithic tradition lasted for around one million years. The prolonged period of this stone industry gave rise to a wide variety of tools that included ovate, cordate, and symmetrical bifacial handaxes. Several archaeologists associate Acheulean industry with different species of hominins, but there is a consensus that the oldest Acheulean tools were elaborated by Homo erectus. [73] The images below provide a general overview of the morphology of these tools.

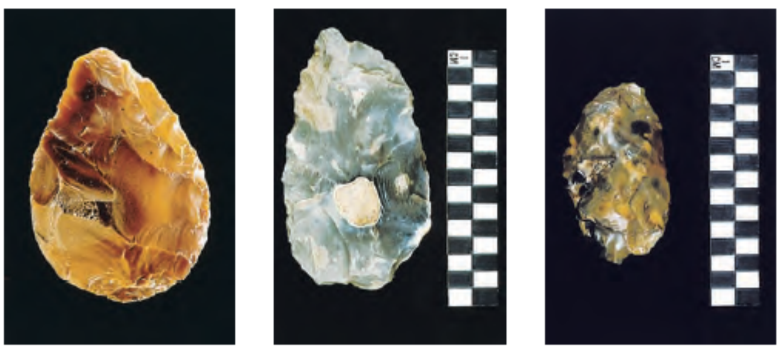


**Figure 6.** Acheulean industry or Mode 2. (Taken from Fuentes, 2007).

[73] There is evidence that before the emergence of speech and the considerable enlargement of the brain, Homo erectus was the first hominin in Africa that spread throughout Eurasia about 1.5 million years ago (Armstrong and Wilcox 2007; Lewin & Foley 2004, Foley 2012). Apart from the ability to manufacture complex tools, Homo erectus also controlled fire (Lewin & Foley 2004, Foley 2012).

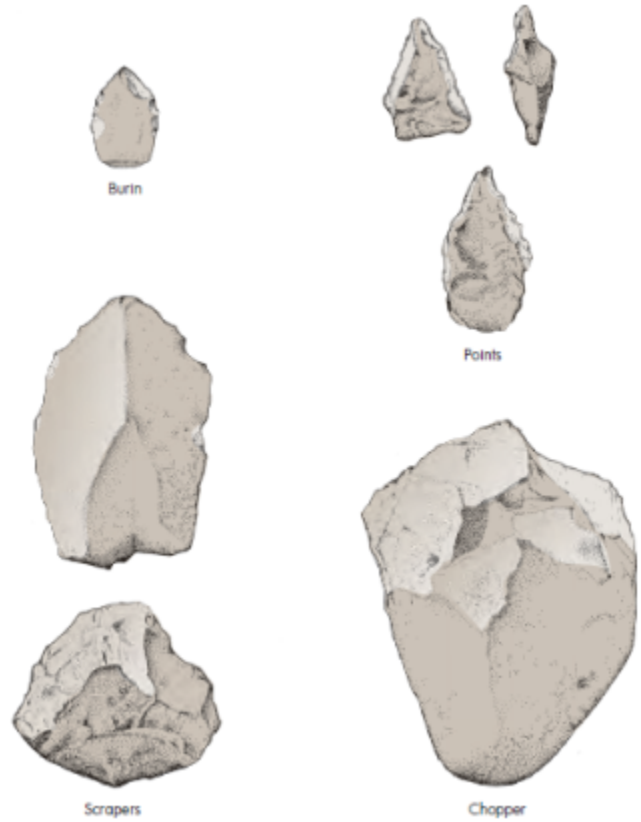



**Figure 7.** Acheulean stone tools. They have been found at many sites, from Africa to western Europe and into central Asia. They are characterized by bifacial flaking, which produces a sharp edge. The color images represent actual "hand axes" used primarily as chopping and cutting tools. The line drawings reflect the different types of tools in the Acheulean tradition. (Taken from Fuentes, 2007).

The images above exhibit that Acheulean industry, in contrast to the previous lithic tradition, displayed an increasing tendency for bifacial shaped tools that are bilaterally symmetric. Archaeologists suggest that the major manufacturing refinement took placed between 600,000 to 700,000 years ago. The wide majority of the hand-axes produced during this period of time are made of stone, bilaterally symmetric and with two sharp edges. Due to its design, archaeologists claim that the manufacturing processes most likely consisted in a complex sequence of steps. The images below illustrate the reconstruction of this manufacturing process.


**Figure 8.** How Stone Age Humans made Hand Axes.
(Taken from Encyclopedia Britannica, Inc., 2006)

As can be seen, different percussion objects such as semi-spherical stones, pieces of wood or other soft material may have been used as hammers for shaping the finest parts of the tools. Most likely, these tools were used for cutting either meat or animal skin as well as for manufacturing different objects of wood or bone (Fuentes, 2007). The emergence of different tool designs over the long period of Acheulean tradition has been explained differently. It has been proposed that the last stage of major sophistication resulted from a sudden punctuated variation. Other scholars suggest a gradual process of sophistication. In any case, three common aspects of Acheulean tradition are commonly highlighted: prolonged periods of stability, a tendency towards symmetrical designs and a continuous increasing of the manufacturing complexity (Tomlinson, 2015).

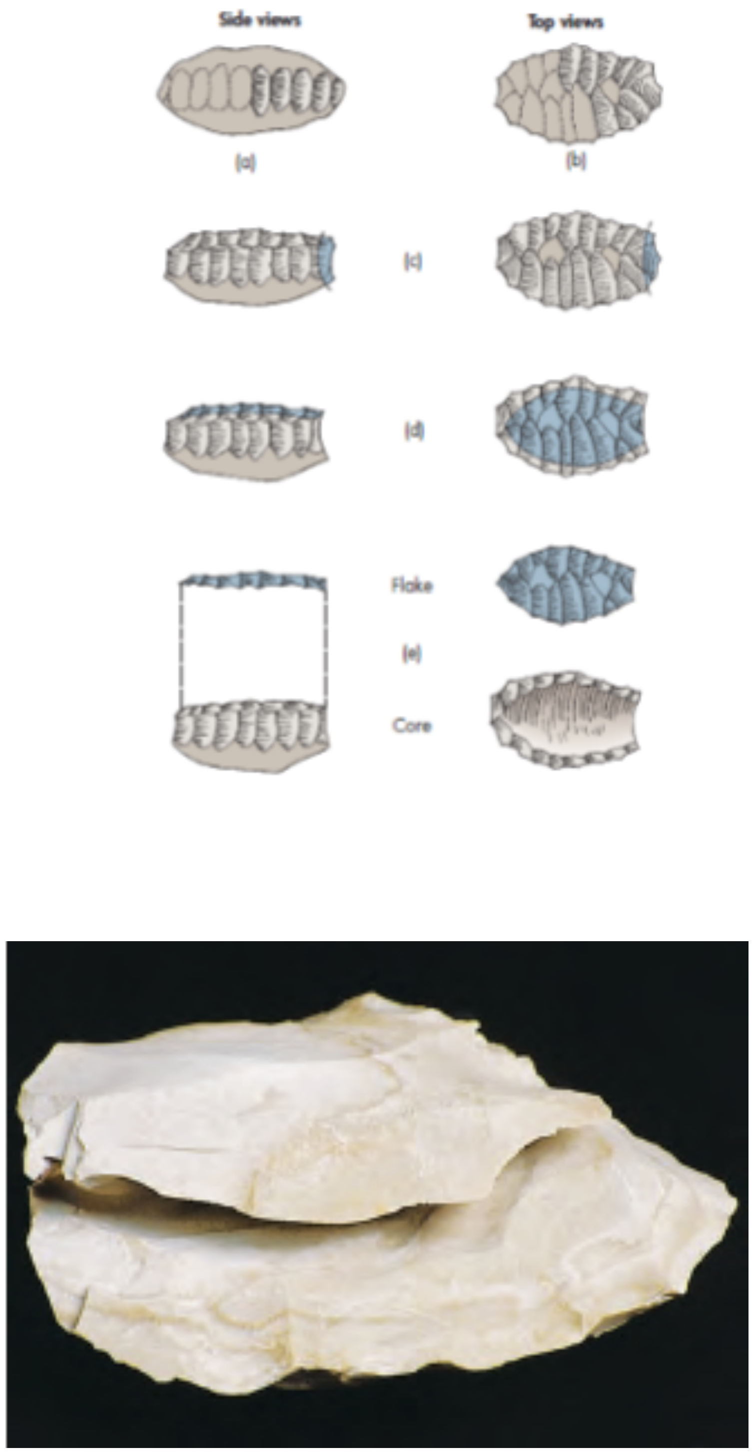
The complexity of these tools have led scholars to inquire for the cognitive capacities that may have been involved in its manufacturing process. It has been proposed that the manufacturing skills may have been favored by the acquisition of the upright posture, which allowed these hominins to freely use their arms and hands to perform a large number of activities (Levinson and Holler 2014). Upright posture may have also allowed them to sustain face-to-face relationships that very likely fostered the development of communicative behaviours that do not require a fully articulated language. Leroi-Gourhan (1964) and Gamble (1999), suggested that the development of a basic kind of gestural communication system may have played an important role in the evolution of this manufacturing technique.[74]

### 5.2.4. Mode 3

Mousterian or Mode 3 lithic industry dates back to around 600,000 to 40,000 years ago. This lithic tradition is based on a manufacturing technique known as Levallois.

Archaeologists argue that this technique was initiated in Africa and later spread in Euroasia approximately 200,000 years ago. It has been claimed that this technique resulted from the acquisition of a higher level of a manufacturing sophistication manifested in a wide diversity of artifacts. Technically speaking, experimental archeology suggests that this manufacturing procedure may have started by striking the periphery of a rounded and elongated stone while the stone front is peeling. Once it gets peeled, it has to be struck right in the core to extract a large flake. The extracted flake is the expected tool. For archaeologists, the manufacturing process reveals that these hominins had already acquired a high level of motor control and very likely complex levels of social organization. The images below show the appearance of these tools as well as its possible manufacturing process.

[74] According to Levinson and Holler (2014), it is very problematic to think that the manufacturing process of this kind of tools were not based of any kind of communicative system supporting the transmission of these cultural tradition.




**Figures 9.10.** The Levallois technique. Produce a margin along the edge of the core (a); shape the surface of the core (b); prepare the striking platform (the surface to be struck) (c,d); remove the flake (e). The lower photo shows a replica of a Levallois core and tool. (Taken from Fuentes, 2007).

### 5.3. Motor control and procedural memory

Generally speaking, the evolution of lithic traditions shed important light on the evolution of a wide range of hominin cognitive capacities. It has been argued, for instance, that Acheulean technology may have been based on a succession of gestural movements triggered by exploratory mechanisms of the material (stone) qualities of the environment. Thus, over time, these exploratory avenues may have turn into the creation and stabilization of patterns of joint action and social interaction. Back in 1964, Leroi-Gourhan proposed the notion of *operational chains* as a way to understand how the structured sequences of gestures may have guided the manufacturing process of Acheulean tool industry. According to him, this sequential repertoire of gestures may have served as a bridge between hominin social behavior and the exploration of the material world.

Thus, it can be thought that the constant exploration of the material environment allowed early hominins to develop motor skills embodied in coordinated sequences of movements, which in turn, underpinned the manufacturing processes. Most likely, the maintenance and transmission of these cultural traditions may have depended on the development of manufacturing strategies tacitly learned by observing and replicating certain kinds of sequential motor movements. Donald (1993) argues that the Acheulean industry was based on the acquisition of a sophisticated level of body motor control. Over time, these skills may have been mastered with high level of accuracy. We can hypothesize that this kind of manufacturing expertise became socially distributed and transmitted across generations. Wynn and Coolidge (2010) suggest, for instance, that the continuous repetition and imitation of the sequential motor movements involved in knapping traditions may have fostered the development of a kind of procedural memory, which also scaffolded the maintenance and transmission of Levallois technique across generations. For these authors, procedural memory may have also scaffolded a wide range of other cultural practices that required the acquisition of high level of motor coordination and manual dexterity structured through complex sequences of motor movements. These skills are thought to support a wide range of cultural practices such as playing music, playing volleyball or ridinga bicycle. Generally speaking, the development of tool-making can be understood as part of co-evolutionary processes involving a set of cognitive capacities, motor skills, artifacts and cultural practices. It can be argued that these co-dependent evolutionary processes constituted cultural niches that substantially modify the cognitive and social-material environment over our hominin lineage. In the next section, I will highlight the relevant role that nonverbal communicative strategies may have played in these co-evolutionary processes. This will allow me to show that our basic cognitive capacities involved in music-making may have emerged as part of co-evolving processes allowing the development of the hominin tool-kit for social interaction.

### 5.4. Nonverbal communication and social interaction

The success of our species has been explained as depending on the brain enlargement and on the emergence of speech. This has also been associated with the development of our symbolic thinking. The evolutionarily relevance of these features cannot be denied; however, I argue for the important role that basic patterns of interaction with the world may have played in the development of a wide range of cognitive capacities that scaffolded the emergence of an ample diversity of cultural practices, including music. I hypothesize that the development of gestural communication may have enhanced complex mechanisms for social interaction underlying the diversification of distinct cultural practices (Kendon, 2009, Levinson and Holler 2014).

Levinson and Holler (2014) point out that our current daily social life is based on a large diversity of mechanisms of gestural communication. It is widely accepted that our bodily movements and postures can describe sizes, trace forms and delineate spatial relationships between individuals and things. Kendon (2009) claims that a basic way to link an agent with some aspects of his environment is by observing what an agent is doing (Kendon 2009). Importantly, body gestures are mostly interconnected. For instance, there is a strong connection between the use of hands (e.g. to carry out a physical task), vocal movements (Levinson and Holler 2014; McNeill 1992; 2000) and facial gesticulation. An intricate set of movements of tongue, lips and jaw is triggered while we use our hands to cut, carry up or push, an object (Levison and Holler 2014).

In primates, visual perception of their conspecifics' actions is fundamental for establishing collaborative social behaviors (Tomasello 2008). In monkeys, chimpanzees and humans, the gaze orientation of an individual can lead conspecifics to observe the same object. In humans, distinct body movements of our conspecifics can redirect our gaze towards an external object. Perhaps the most common guide movement is extending the arm to point out with the index finger. This practice exists in all human cultures. Infants start pointing out before the first year of life, before having learned their mother tongue (Kendon 2009). It is thought that index finger pointing is a distinctive human capacity (Levinson and Holler 2014).[75]

Newborns' ability to direct their attention towards adult's facial movements and emotional expressions is an initial component of our cognitive life (Gallagher, 2012, p 208). There is evidence that our human capacity for imitating facial expressions of people with whom we interact starts from the first 36 hours of birth (Field, et al 1982). Furthermore, the repertoire of gestures that children use for communicative purposes increases over their lifetime (Levinson 2006). The communicative function of gestures did not disappear after the emergence of speech. Speech and gestures are two indissociable components of a highly complex and integrated human communication system (Kendon 1990; 2009; Armstrong and Wilcox 2007; Levinson and Holler 2014). Experimental studies have revealed that linguistic skills and the ability for decodification decrease significantly when adults are required to talk without gestures (Levinson and Holler 2014). Similarly, facial gesture and hand movements decrease as a consequence of stuttering, both making linguistic fluency almost impossible to maintain (Mayberry and Jaques 2000). The use of gestures in linguistic communication remains in contexts where there is not visual contact

[75] Although some monkeys ask for things by extending their arms and hands, they do not point out with the index finger, except the ones living in captivity (Leavens et al, 1996)

between participants, such as during a phone call. There is evidence that people blind from birth use gestures while speaking (Levinson and Holler 2014).

Thus, most likely, gestures may have played an important role in the development of primary communicative systems that scaffolded the evolution of a wide range of cultural practices. For instance, Gamble (1999) points out that the origins of Acheulean industry may have arisen from mechanisms of gestural communication, material exploration and rhythmic patterns of stone percussion and carving. Ingold (1993) makes a distinction between neutral — uninhabitable — landscape, and *taskscape*. The latter, he says, emerges from the activities of a social group, the sequence of movements involved in these activities, the material characteristics of the environment and the sound produced by these practices. Thus, the development of lithic industries may have been supported by the construction of *taskscapes* of activities, which can be defined as multimodal cultural niches. There are reasons to think that the refinement of motor dexterity implied in tool making could have been supported by rhythmic sequences of bodily movements. This in turn may have fostered the development of joint and coordinated actions that enhanced the evolution of a wide diversity of cultural practices, including music-making. It is worth to note that coordinate actions require a wide range of cognitive abilities, among them, a capacity for sensorimotor synchronization. In the next section, I will show that the notion of *entrainment* can shed light on the way in which synchronized actions scaffold music as well as many others cultural practices.

### 5.5. *Entrainment*

*Entrainment* is a concept coined by Christiaan Huygens in 1665. He used it to characterize a specific physical phenomenon. Huygens placed several pendulums on a common base and made them work at different speeds. After a short time, the motion and speed of all the pendulums got synchronized. This physical synchronization is what Huygens defined as *entrainment* (Clayton, Sager and Will, 2005; Will & Turow, 2011; Clayton, 2012). He thought that *entrainment* was a ubiquitous phenomenon of nature. Since then, this concept has served to characterize the way in which two or more independent oscillatory processes become synchronized by interacting and mutually influencing each other (Will & Turow, 2011, Clayton, 2012). This tendency has been found in a wide variety of physical and natural systems such as: firefly's light flashes, or toad croaks, among others (Clayton, 2012; 2013).

In the case of humans, *entrainment* also refers to rhythmic processes that take place within the body of an individual such as oscillatory synchronization of brain waves, hormone secretion, heartbeats, blood circulation, breathing, eye blinking, walking movements, locomotion, and female menstrual cycles (Clayton, 2012; Will and Turow, 2011). *Entrainment* also emerges from patterns of interaction between individuals such as synchronized grain grinding, handicraft activities, sawing wood or rowing boats. In many cases, these activities also involve the synchronization of breathing and limb movements (Clayton, 2012). Thus, when two or more individuals interact, the rhythmic movements of their actions become entrained (Clayton, Sager and Will, 2005; Will and Turow , 2011; Clayton, 2012; Moran, 2013). It is important to say that even though entrainment can be intentionally provoked, it mostly arises spontaneously.

Clayton identifies different kinds of *entrainment* based on the way two or more oscillatory systems interact with each other: one to one, one individual and a group, and two or more groups. He also points out that this phenomenon can produce symmetrical and asymmetric interactions based on distinct rhythmic patterns and time scales (Clayton, 2012). Thus, two clocks whose rhythmic processes influence each other display a *symmetric entrainment* (one-to-one). An *asymmetric entrainment* [76] (one-to-group) is produced when several light flashes of a firefly group become synchronized because each member of the group is influenced by the light flashes of the others. On the other hand, in a carnival parade, many musical bands become synchronized once they all start playing music together. This synchronization is based on a relation between groups that also produce a *symmetric entrainment*. (Clayton, 2012). In general terms, *symmetric entrainment* refers to a collective performance where participants influence each other and, by doing so, maintain the collective synchronization of the whole group. On the other hand, *asymmetric entrainment* occurs when some members of a musical ensemble exert influence on the rest in order to maintain the collective synchronization. The latter phenomenon can occur for several reasons: when there are musicians who play the role of teachers or experts guiding the performance of the rest; when a musician joins for the first time an ensemble and has to follow the group, etc. These mechanisms of interaction can also be identified in dance (Laland et al., 2016). In most cases, *entrainment* emerges without the need of any explicitly verbalized instruction. Observation, imitation and nonverbal communication play crucial roles in triggering this kind of collective behavior. In the next section, I will show that *entrainment* allows us to see that music-making mostly emerges spontaneously from patterns of social interaction. We will also see that *entrainment* sheds important light on a vast range of behavioral commonalities between music-making and other complex and multimodal processes of social interaction.

[76] According to Clayton (2012), the circadian cycles of individuals influenced by day and night cycles can also be characterized as *asymmetric entrainmen*t.

### 5.5.1. *Entrainment* and musical practices

*Entrainment* allows individuals to synchronize the movements of their bodies to an external beat. This is a basic requirement for musical ensembles in which groups of individuals synchronize their movements during their musical performance (Clayton, 2012). Importantly, *musical entrainment* does not refer to an “absolutely precise” synchronization of movements among individuals. Musical performances are not necessarily based on following a beat in unison.[77] *Entrainment* implies the idea of "dragging”, and "mutual influence. For example, when a group of musicians play a specific printed musical score, the execution time of each note is variable. It does not reflect the exact mathematical duration represented on the sheet. This phenomenon is imperceptible to both musicians and audience due to the minimal range of variation in time in which each note is played. Thus, the collective performance is perceived (by musicians and audience) as a complete synchronized social practice.

Furthermore, *entrainment* can spontaneously arise from either a percussive polyrhythmic sequence of beats or from relatively uniform percussion patterns. Rhythmic collective coordination is not necessarily based on completely synchronized movements and sounds distributed into an exact metric unit (Clayton, 2012). Many musical traditions around the world (e.g. Pygmy music, variants of jazz, and variants of Mexican *son*) are not necessarily based on a regular beat. In fact, for many musical traditions, playing "out of phase" can be an implicit way to regulate the collective coordination in a musical ensemble. This means that musicians can execute different rhythmic patterns in an ensemble (e.g. by singing or playing any instrument) either in "phase" or "out of phase", but the ensemble, as a whole, maintains a collective coordination. I propose an evolutionary characterization of *entrainment* as a *cognitive device that may have emerged from dynamic, flexible and intertwined assemblies of skills and cognitive capacities based on sequential, coordinated, and spontaneous coupling action, which in turn, scaffolded the development of a wide variety of cultural practices in*

[77] Becker (2011) prefers to use the notion of rhythmic entrainment instead of rhythmic synchronization.

*our hominin lineage*. This understanding of *entrainment* can lead us to think that most likely a set of basic capacities involved in music-making may have arisen as part of complex cognitive assemblies playing an important role in the evolution of our social interaction.

### 5.6. Musical capacities and social interaction: a co-evolutionary approach

My evolutionary hypothesis is that our capacity for making music may have been cultivated as part of complex processes of co-construction and co-evolution of niches of cognitive capacities and skills underlying the structuring of a complex and multimodal cognitive engine for social interaction. In what follows, I will show that a set of basic abilities involved in making music (sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization) may have scaffolded the development of multimodal learning scenarios allowing the evolution of interdependent cultural practices and the complexification of our sociality.

#### 5.6.1. Music making in multimodal learning scenarios

It has already been mentioned that *entrainment* allows individuals to synchronize the movements of their bodies to an external beat (Clayton, 2012). Some scholars suggest that the ability to align to an external beat was acquired once hominins developed a high level of vocalization and articulated language (Patel, 2009). The vocal and motor behavior of parrots and cockatoos has been presented as a valuable sign on the relation between vocal skills and the capacity for synchronizing to an external beat. Patel et al (2009) conducted an experiment with a cockatoo named Snowball. In this experiment, they show that Snowball was occasionally able to synchronize his movements to a regular beat of a musical stimulus (particularly to the song "Everybody" played by the band "Bad street boys"). It must be mentioned that Snowball was 6 years old when it was obtained by its current owner (PO). By the time the experiment was conducted, Snowball was 12 years old. Importantly, when Snowball was obtained by (PO), it was already able to synchronize its movements with music by dancing. It is uncertain when and how Snowball developed this ability. Although it is widely accepted that cockatoos have developed a complex social life, this particular behavior has not been observed in other cockatoos in the wild. Patel et al (2009) recognize the need for conducting more experiments with more species of birds to see whether or not the ability for vocalization, which they assume closely related with linguistic capacity in humans, can be understood as a prerequisite for developing the capacity of synchronizing to an external musical beat. Patel et al. endorse a "vocal learning and rhythmic synchronization hypothesis", which suggests that the origins of the human capacity for music synchronization depended on the evolution of our capacity for vocal learning. From this perspective, the evolutionary origins of music is explained as a consequence of the capacity for vocalization and speech.[78] Schachner et al (2009) have tried to show that Snowball's behaviour can be found in other vocal species. However, their reports are uncertain as they are based on videos on Youtube.

I agree that vocal sounds have certainly become an important component of music over our evolutionary path, but it does not imply that vocalization (fully articulated) was a necessary prerequisite for developing the capacities for music-making, in particular the capacity for *entrainment*. For instance, researchers suggest that *entrainment* behaviour has recently been found in the less vocally flexible California sea lion, which in turn represents a limitation for the vocal learning and rhythmic synchronization hypothesis (Cook et al. 2013). Furthermore, several studies suggest that our ability to produce rhythmic patterns is based on complex sequences of muscle movements whose development preceded the emergence of spoken language (Alcock, Passingham and Vargha-Khadem 2000, Mayberry and Jaques 2000, Morley 2012a). This leads us to think that the considerable refinement of motor control required by ancient processes of tool-making may have enhanced the development of complex sequences of muscle movements, and perhaps the synchronized

[78] The explanatory limitations of this kind of approaches has been already surveyed in the second chapter .

patterns of knapping (D'Errico et al., 2003; Tomlinson, 2015), fostered the capacity for synchronizing to an external beat. Humans do not need special training to detect and motorically synchronize to an external beat (Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2018). This is a human capacity that may have ancient roots in our hominin lineage. Our capacity for detecting beat in rhythmic sound sequences is already functional at birth (Winkler, Háden, Ladinig, Sziller, & Honing, 2009), and it has been shown via experimental studies that our auditory and motor systems are completely integrated (Zatorre, Chen, & Penhune, 2007). Contrastingly, this capacity is not fully developed in non-human primates. Macaques, for example, can produce rhythmic tap tempo matching. However, the taps of macaques mostly occur about 250 ms after stimulus onset, whereas humans show asynchronies close to zero (Merchant et al., 2018, p. 172). Moreover, macaques reach this level of synchronization by observing the movements of the experimenters. They are less capable of synchronization by hearing, which can be explained by a marked absence of coupling between the auditory and motor system (Merchant & Honing, 2014). Ravignani et al. (2016) argue that despite our cultural diversification, humans demonstrate a general proclivity for rhythm. They suggest that percussion instruments may have provided the first form of musical expression in human evolution. They even consider that percussive behaviour may have already been present in our ancestors some million years ago, before the split between the human and Pan Lineage. In order to test, in the lab, this proclivity for rhythm, they asked people to participate in a 30 minutes drumming experiment. The participants (no musicians) were told to imitate sets of randomly generated drumming sequences that they heard in headphones. The imitation attempts were recorded and heard by the next participants in independent transmission chains. By perceiving and imitating drumming sequences from each other, participants turned initially random sequences into rhythmically structured patterns. Two characteristics seem to emerge in most of these experiments: random patterns evolve into sequences that exhibit increasing learnability and structure over generations of learners. In other words, initially random sequences transformed into increasingly structured and learnable music-like patterns. Ravignani et al. (2016) interpret their experiment as a possible way to obtain divergent musical cultures, where each 'musical culture' (corresponding to an experimental chain) constituted a system by itself. Their claim is that random sequences of sounds can turn into the stabilization of rhythmical patterns scaffolded by imitative behaviours and memory.

This experiment allows us to think that tool-making, and any other repetitive percussion-like exploration of the acoustic properties of the environment, might also have fostered the stabilization of rhythmical patterns over our evolutionary history. Importantly, given that this study is centered on the way sound sequences can be replicated after been heard, the results also reveal the interdependence between the auditory and the motor systems, which has already been shown by other studies (Zatorre et al., 2007). It has also proved that sound is particularly useful for perceiving periodicity, regularity, and velocity of movement (Nesbitt 2003; Kapur et al. 2005). For instance, motor learning itself can be aided by sonification, especially the time-dependent dynamic coordination of movement. Furthermore, there is a strong connection between our capacity for reproducing cyclical mechanisms of motor movements and the capacity for producing gestures (Morley, 2012). On the other hand, it is also widely accepted that human gestures potentiate our capacity for imitating motor patterns. This enables us to anticipate and replicate corporal behaviors of people with whom we interact. I have argued that music-making also involves patterns of imitation supported by gestures and body movements. In other words, gestural communication enables musicians to assemble their individual actions and bodily movement into collective and synchronized behaviours (Moran 2013; Rahaim 2012; Clayton 2013).[79]

During a musical performance, musicians have multiple possibilities for coordinating their actions collectively depending on what currently happens and

---

[79] Even in classical Western music, the execution of some musical passages does not only require the ability to "read" the score accurately, it is also necessary to pay special attention not only to the verbal instruction given by a music teacher, but above all, music students have to observe the way in which the teacher executes these musical passages. Knowing to listen and observe what the teacher performs becomes crucial for music learning, this will allow the student to imitate the teacher's movements to achieve the required music sound. In most of the cases, for teachers it is more effective to show how a certain musical passage must be executed than verbally explain how to do it.

what is expected to occur. Musicians constantly adjust their actions through spontaneous and mostly, implicit nonverbal communicative behaviors. In a wide range of musical cultures, learning how to play music does not necessarily require explicit verbal instructions, but rather it is mostly supported by several patterns of joint actions that become recognizable by musicians during the performance. A high level of sensitivity to bodily respond to the movements of other bodies is fundamental. Musicians constantly have to adapt their actions to subtle body modulations of other participants (Moran, 2013). This allows us to see that music arises from complex patterns of movements embodied in social and coordinate actions. It is not something simply heard and consumed, but rather it is something produced from dynamic processes of mutual social interaction (Cross 2008).

Archaeologist suggest that the use of gestures as a communicative and imitative device may also have played an important role in the development of lithic industries (Gamble, 2010; 2012; Leroi-Gourhan, 1964). The maintenance and transmission of these cultural traditions may have been supported by mechanisms of tacit learning resulting from replicating certain kind of sequential motor movements constantly heard and observed. There is evidence that audition reinforces vision in resolving uncertain visual motion patterns, for example, when grasping an object (Castiello et al., 2010). It has also been demonstrated that learning of a visual task is superior in subjects trained with congruent audiovisual stimuli compared to subjects trained with solely visual stimuli (Kim et al., 2008). Furthermore, it has been shown that multimodal learning may strengthen multimodal cognitive representations and the connectivity between brain areas (Shams and Seitz 2008). This reinforces the idea that the coordination of tasks involving different capacities or subsystems (e.g. seeing, moving, touching, and sound production) has cascading effects in other tasks in which some of the same subsystems are involved (Sheya & Smith, 2010). Thus, we can hypothesize that tool-making might also have fostered the development of specific forms of listening, sound discrimination and production (Blake & Cross, 2008; Cross, Zubrow, & Cowan, 2002; Zubrow, Cross, & Cowan, 2001) visual acuity, and some kind of *entrainment* (Killin, 2016a, 2017; Morley, 2013), which in turn could have enhanced knapping expertise.[80] It is very likely that the overlapping of multimodal perception-action loops was also responsible for the cascading socio-material interactions that scaffolded the construction of cultural niches, understood as multimodal learning scenarios. I suggest that these cultural niches, constituted by a conglomerate of cognitive capacities, artifacts and cultural practices, which also include the mosaic of capacities involved in music making, fostered the complexification of human sociality.

### 5.6.2. Spontaneous coupling movements as evolutionary scaffoldings for music-making and social interaction

I have pointed out that the phenomenon of *entrainment* involves the capability to synchronize our body movements to an external pulse (e.g. clapping or finger snapping as a response to a musical stimuli) and that it does not necessarily imply a perfect coupling (Clayton, 2012). This leads us to think that the evolution of upright walking may also have improved the development of complex mechanisms of joint action and *entrainment*. Kelso (1995), for instance, remarks that walking together may have been highly relevant for developing several tasks of collective coordination. As an example, he refers to the coupling behavior that emerges when an adult and a child walk together. In this case, the walking speed of each one varies depending on the length of their respective legs, their body shape, their particular habit of walking, etc. This means that their individual walking speed varies when they walk separately. But, when both individuals walk together, their body movements get synchronized in a way that their stride lengths are almost equal. The child may be walking faster than usual while the adult may do it slower. The point is that this kind of coupling is progressively reached and maintained by flexible and variable joint walking movements. In other words, when an adult and a child walk together their body movements are mutually influencing one another and become relatively stable. In general, by

[80] Larsson (2015) considers that hammering is likely to have been a commonly heard sound throughout millions of years. Scraping and striking stones, bones and wood, may have given rise to rhythmic patterns of sound sequences. Likewise tool uses such as: slicing flesh from bones, cutting vegetation, cracking nuts, and the use of grinding utensils in food processing.

walking side by side, people, often subconsciously, synchronize their steps. This suggests that when people walk together the perception of the partner's walking influences the gait and walking movements of the people involved, without requiring any conscious effort or intent (Larsson, 2013). Furthermore, in paired walking, participants can be phase locked with a phase difference close to 0 (in phase), or they can be phase locked with a phase difference close to 180 (antiphase) with walkers contacting the ground simultaneously with opposite-side feet.[81] Kelso (1995) proposes the notion of relative coordination to refer to this type of coordinated movements that emerge from patterns of social interaction.[82] The notion of relative coordination is compatible with the notion of *entrainment*, as it also captures the phenomenon of "dragging" and "mutual influence" that emerges when two or more oscillatory mechanisms interact. Kelso emphasizes that these mechanisms of relative coordination emerge without the need for a linguistic and explicit guide, as it is the case of a wide range of social activities (Levinson 2006; Kendon 1990).

Note that walking also triggers hearing capacities, as it conveys information about the properties of the sound source, and without explicit training, listeners learn to draw conclusions based on the features of the sound, including such aspects as gender (Li et al. 1991), posture (Pastore et al. 2008) emotions of a walker, and properties of the ground surface (Giordano et al. 2012). Importantly, music is often played at a tempo similar to walking (Changizi 2011, p. 191). There is evidence that people can synchronize walking movements with music over a broad spectrum of tempos (Larsson, 2013). Furthermore, listening to a metronome has a positive effect on walking rehabilitation, especially in patients who have suffered a stroke. Patients improve their walking abilities (e.g. synchronization of their stride lengths, coordinated movements of left and right leg, etc.) by hearing an external rhythmic pulse either from a metronome or a musical stimulus. Furthermore, movement disorders caused by Parkinson's disease have been corrected by auditory rhythmic

[81] The latter means that the right foot of one walker and left foot of the partner will reach the ground almost simultaneously (Larsson, 2013).

[82] Kelso (1995) makes a distinction between relative coordination and absolute coordination. The former has already been explained. The latter refers to a mechanical, rigid and permanent coupling of a machine.

stimulation (Thaut, M. et al, 1996). When the acoustic stimulus ceases, then walking abilities degenerate considerably (Morley, 2013, Minetti and Alexander, 1997).

All the above leads us to think that, most likely, bipedal walking played an important role in the development of *entrainment* and thereby the evolution of rhythmic abilities and music making. We have seen that beat induction is a cognitive skill that let us perceive a regular pulse in music to which we can then synchronize. Perceiving this regularity in music allows humans to create music and dance together. Dancers make use of the synchronization abilities to synchronize with music and with the other dancers. In the previous chapter, I showed that the distinction between music and dance is not made in many languages around the world. In those cases music and dance are conceived as a common mode of human interaction (Morley 2014; Nettl 2000; Cross 2001). This allows us to see that collective dancing, walking and, music-making, are supported by common clusters of cognitive capacities constituted by the *entrainment* of rhythmic and communicative body movements.

I suggest that the coupling movements emerging spontaneously from walking in a group and the coordinated movements involved in collective dancing (eg. in communal rituals) differ only in degree, not in kind. This reinforces the idea that joint walking, as a specific way of dancing, may also have played an important role in the evolution of rhythmic abilities. In fact, in many musical cultures around the world there is not a clear distinction between walking, marching, dancing and making- music. Fitch (2018) claims that the core of most human dancing is not strictly distinguishable from drumming. In the previous chapter, I showed that in many cultures around the globe, the rhythmic sounds produced by dancing can be understood as a constitutive component of music, which means that dancing is conceived as a percussion instrument. For instance, in a wide range of variants of Mexican *son*, dancers execute their movements on wooden platforms; they strike their feet on these wooden platforms so the rhythmic sound produced becomes audible. Note that these percussion sounds fit perfectly with the sounds of other musical instruments. Likewise, flamenco dancing, Pizzica dancing from Salento, in Puglia, Italy, Schuhplattler dancing from Austria, dancing of music *llanera* (grasslands) from western-central Venezuela and eastern Colombia; tap dancing, among many others, can be conceived as both dancing and drumming simultaneously. Furthermore, in certain regions of Africa, percussion sounds of dancing are conceived as music by itself. No other musical instrument needs to be added.

Imitation is an important pillar of music-making. Imitation enhances synchronization abilities by providing a high-fidelity copy of the body actions of others (Trehub, et al (2015). Likewise, imitation allows a competent couple of group dancing to coordinate individual actions and, in the process matching, reversing or complementing each other (Laland et al., 2016). We have seen that imitation enhances the integration of visual, auditory and motor systems. Importantly, this allows musicians and dancers the acquisition of long sequences of movements within complex scenarios of social learning. Thus, the learning of a dance routine invariably begins with a demonstration of the steps from an instructor or choreographer, which the dancers then set out to imitate. It is not coincidence that dance rehearsal studios almost always have large mirrors along one wall. These allow the learner to switch rapidly between observing the movements of the instructor or choreographer and observing their own performance (Laland et al., 2016). Furthermore, dance and music require the capacity to merge, temporally separated sounds and bodily movements, into a coherent sequence of rhythmic patterns, which means to merge sounds and bodily movements into a complex (and often long) sequentially integration of coordinate actions. Importantly, there is evidence that the ability of learning and reproducing sequential integrated actions is grounded in processes of social learning. This suggests that long chains of actions are very difficult to learn asocially, but rather social learning substantially increases the chances that individuals will acquire the appropriate sequence (Whalen, et al., 2015).

Traditional musical performance of Mexico known as *zapateado*,[83] illustrates the way social learning scaffolds the individual acquisitions of complex sequences of actions involved in music and dance. In this social event, the level

---

[83] Also known as fandango jarocho, this is a traditional music and dance celebration from Southern Veracruz and neighboring regions (see García de León, 2006).

and roles of participation for musicians and dancers can vary widely, allowing people with very basic skills to perform together with experts. Some roles can imply clapping in time, or playing basic chords and rhythmical patterns on the *jaranas*,[84] while others may require a very high level of musical and dance specialization.[85] In Martínez & Villanueva (2018) we argue that *fandango* allows us to see how a great variety of roles and abilities at different levels become distributed among the participants.[86] We point out that this social practice fosters the transmission of musical and dance knowledge during the performance, mostly without any explicit verbal instruction, but rather by displaying loops of tasks of social coordination driving and defining new tasks that shape the complex social structuring of this cultural phenomenon.

Laland et al. (2016) suggest that our ancestors were predisposed to be highly competent manipulators of a long string of behavioral elements because many of their tool-manufacturing and tool-using skills, extractive foraging methods, and food-processing techniques required them to carry out precise sequences of actions in the right order. This suggests that repetitive and predictable sequences of actions may have strongly influenced the development of *entrainment.* In fact, Laland et al (2016), hypothesize that these sequence-learning capabilities are clearly exploited in learning dance and, I would add, in music-making and in a wide range of cultural practices involving basic patterns of social interaction.

This evolutionary view leads me to argue that music-making may have emerged from the dynamic integration of a cognitive engine constituted by clusters of cognitive capacities and skills such as sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of

[84] The jaranas are traditional guitars employed in the fandango jarocho. They look like the thin Baroque guitars and have five courses, three of which may be doubled, with eight strings in total. They come in a variety of sizes and are traditionally made from a single piece of wood.

[85] As with the improvisatory melody played by the guitarra de son, also known as requinto jarocho, consisting in a small guitar whose four (or five) strings are plucked with a horn plectrum, and the complex zapateado dancing.

[86] See the image in the Appendix section

motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, collective synchronization, and sequential structuring of coupled actions. I hypothesize that the dynamic integration of this cluster of capacities and skills involved in music-making arose from a prolonged and extensive diversification of patterns of interaction-exploration between individuals and their socio-material environment along our hominin lineage. Furthermore, it is most likely that this basic cluster of cognitive capacities may also have scaffolded the development of a wide range of cultural practices, social tasks and artefacts, and still remains supporting the current complexity of basic patterns of social interaction and social cognition in our daily life.

# CONCLUSIONS

Throughout this dissertation, I have argued for an evolutionary account of the possible ways through which a basic assemblage of capacities and skills involved in music-making become integrated, cultivated and transmitted across generations. It is most likely that the wide range of the human features, capacities and skills, including the ones involved in music-making, co-evolved in different periods of time. For many years it was thought that distinct features associated with *Homo sapiens*, including large linear bodies, elongated hind limbs, large energy-expensive brains, sexual dimorphism, increased carnivory, and unique life history traits, evolved near the origin of the genus in response to heightened aridity and open habitats in Africa.

Recent analyses of fossil, archaeological, and environmental data indicate that such traits did not arise as a single package. Instead, some arose substantially earlier and some later than previously thought. Antón and colleagues (2014), suggest that, from ~2.5 to 1.5 million years ago, three lineages of early homo evolved in a context of habitat instability and fragmentation on seasonal, intergenerational, and evolutionary time scales. Likewise, there is evidence that the emergence of eyes in vertebrates resulted from different evolutionary events, not as the product of a singular adaptation for vision. This leads us to think that most likely the origins of distinct items commonly attributed to Homo, including the ample variety of our cognitive capacities, have arisen throughout different populations (and species) over the course of the hominin lineage.

It had been argued that the capacities required to make stone tools arose alongside the emergence of the genus Homo. However, as shown in the last chapter, stone tools dating back 3.3 million years ago discovered in the West of Lake Turkana in Northern Kenya, reveal that the evolutionary origins of these capacities may have appeared long before the emergence of our species. This kind of findings challenges the way in which the evolution of the hominin technological behavior has been characterized and opens new valuable avenues for the study of cognitive and cultural evolution.

Regarding music, there is an increasing agreement among several researchers that music consists of a complex phenomenon whose origins should be explained as part of co-evolutionary processes of behaviors and capacities (Killin, 2016, 2017; Tomlinson, 2015; Van der Schyff & Schiavio, 2017). In general terms, my proposal resonates with this kind of evolutionary approach. However, these models draw the conclusion that music did not arise until our vocalization skills were fully articulated, which means until the arrival of modern human (Killin, 2016, 2017; Tomlinson, 2015; Van der Schyff & Schiavio, 2017; Morley, 2013). I have outlined, in distinct parts of the thesis, the problematic aspects of this assumption. In this dissertation, I have also argued, based on empirical findings coming from distinct disciplines, that most likely the development of the capacity for vocalization might not have been a key component of a basic cluster of capacities underlying the initial stages of musical behavior.

On the other hand, I do not think that it is possible to localize the precise moment and place when music first appeared, as for instance, it is not possible to indicate the precise moment and place when a specify musical genre, a dance style,[87] or a particular dish of a local cuisine, first appeared. Neither it is possible to situate, for instance, the precise moment and place when a person becomes bald or when a person acquired his current weight and body shape. These cases are a consequence of historical processes involving the interaction of multiple components in different spaces and periods of time. Likewise, what I have proposed here is an evolutionary model for a better understanding of the possible ways through which a basic set of capacities, involved in music-making, co-evolved and become integrated over the course of our evolutionary path. My

[87] The historical and ethnomusicological study of Hernández-Jaramillo (2017 ) on distinct musical expressions (*El jarabe Loco and El zapateado jarocho, El punto and El zapateado Cubano, and La Guajira Española*) spread in three countries (Mexico, Cuba and Spain), from the seventh century until now, reveals the misleading enterprise of situating the precise moment and place where each of those musical genres first appeared. Instead, Jaramillo presents a flowing picture of how the constant interaction between the musical components of these musical expressions gave rise to a complex musical system widely spread in both space and time.

evolutionary account of music making is centered on the idea of a cognitive assemblage constituted by the following components: sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization. This list is not an exhaustive account of all the components underlying the capacity for making-music. Different componential breakdowns of music-making can be appropriate for different research purposes. What I have proposed here is an operational breakdown of an evolutionary picture whose distinct components may have co-evolved and become integrated over the hominin lineage.[88]

As we saw throughout this thesis, the aim of several evolutionary accounts of music are committed to find a unilateral, and mono-causal explanation, on the origins of this cultural practice. In part this is due to the fact that these models preserve the idea that natural selection is the sole evolutionary factor for explaining the emergence of any biological and cultural trait, including music. The usefulness of this standard approach to evolution has been challenged by recent theoretical models and conceptual tools. Niche Construction Theory (NCT) offers a broader framework to understand co-evolutionary processes between organisms, environments, artifacts and capacities. From this perspective, organisms and their environment are mutually engaged through co-evolving evolutionary mechanisms based on processes of reciprocal causation. I have argued that this approach allows the integration of biological and cultural interactions into a complex evolutionary account and provide a more effective theoretical framework for thinking about the emergence of music than the offered by the adaptation/non-adaptation perspective.

Niche Construction Theory also provides a better framework for studying interdependent mechanisms underlying a wide range of social practices whose evolvability and transmission cannot be explained solely on the basis of an adaptationist evolutionary account. This includes a broad spectrum of institutions, beliefs, artifacts, and social behaviours that have been constructed and

[88] My proposal resonates with the idea of a relational thinking and processual understanding of life systems (see Dupré and Nicholson, 2018).

transmitted across generations through bio-cultural entrenchments of capacities, artifacts, relations, and processes.

The ethnographical example overviewed in the introduction to this thesis (the Saint Patron's celebration of Huehuetla, Puebla), and plenty of multidisciplinary studies discussed throughout this dissertation, reveal that distinct cultural practices (e.g. music and dance) are transmitted on the basis of similar assemblages of cognitive and cultural components. The traditional festivity in Huehuetla shows that a wide range of cultural items are commonly reproduced in assembly. In this case, the preservation and transmission of ritual music goes hand and hand with the preservation and transmission of ritual dance. Thus, items that are reproduced and repeatedly assembled can become entrenched in a cultural system and are thereby available to serve as scaffolding for later items, as a platform or as a constraint (see e.g. Caporael, Griesemer and Wimsatt, 2014; Wimsatt, W. C. & James R. Griesemer; 2007).

It has been argued that modeling the evolution of cultural traits as repeated assemblies and entrenchments of heterogeneous relations, parts, and processes, provide an alternative, albeit complementary, to the neo-Darwinian population genetic approach for conceptualizing evolutionary change (Griesemer and Wimsatt, 2014, p. 2). In this regard, I have suggested that the heterogeneous cognitive assembly supporting our capacity for making music was evolutionary grounded in multimodal cultural niches resulting from distinct cognitive entrenchments that occurred over the course of the hominin lineage. These cultural niches may have functioned as learning scenarios that potentiated the development and transmission of a basic set of cognitive capacities and skills required for making music, which in turn, would have enhanced the complexification of our patterns for social interaction and social cognition.

In summary, in this thesis I have suggested that diverse exploratory mechanisms of the socio-material surroundings have generated a wide variety of cognitive capacities and skills. It is most likely that over time, these patterns of interaction with the socio-material environment have produced repeated assemblies of cognitive capacities and skills underlying the development and stabilization of a wide range of cultural practices. It is very unlikely that an ample range of cultural practices became constituted from rigid cognitive assemblies. I rather suggest that distinct cultural practices have emerged by consistently recruiting, sharing and combining different clusters of cognitive assemblies. Thus, an assemblage of basic capacities and skills involved in making music (sound discrimination and sound production, intrinsic temporality, spatiotemporal orientation, gradual refinement of motor control, rhythmic body motions, imitative motor behavior, nonverbal communicative systems, and collective synchronization) could have scaffolded the construction of cultural niches constituted by multimodal learning scenarios allowing the evolution of interdependent cultural practices and the complexification of sociality. Therefore, the search for the evolutionary path of certain cognitive capacities exclusively linked to music is a misleading enterprise. By contrast, my main argument is that most likely the capacities for making music evolved alongside our evolution of social interaction. This implies that these musical capacities are part of the basic capacities for social interaction, and vice versa. I have shown, throughout this thesis, that studies in ethnomusicology, psychology of music, gestural communication, social cognition, cognitive archaeology, and social interaction, among others, lead us to think that this is the case.

## APPENDIX

**Huehues dance (Huehuetla, Puebla, Mexico)**


**Negritos dance. Religious procession (Huehuetla, Puebla, Mexico)**

**Voladores danza (Huehuetla, Puebla, Mexico)**

**Religious procession (Huehuetla, Puebla, Mexico)**

**Quetzales dance (Huehuetla, Puebla, Mexico)**

**Religious procession (Huehuetla, Puebla, Mexico)**

**Fandango Jarocho (Veracruz, Mexico)**

## REFERENCES


Alcock KJ, Passingham RE, Watkins KE, & Vargha-Khadem F. (2000). Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain Lang*. 75(1):17-33

Amundson R. (1994). Two Concepts of Constraint: Adaptation and the Challenge From Developmental Biology. *Philosophy of Science* 61: 556-578.

Amunts K., Schlaug G., Jäncke L., Steinmetz H, Schleicher A, Dabringhaus A. & Zilles K. (1997). Motor cortex and hand motor skills: structural compliance in the human brain. *Human Brain Mapping* 5 (3): 206-215.

Andersson, M (1994) “Acoustic Signals” en: *Sexual Selection*, Princeton University Press, New Jersey, 349- 368

Angulo-Perkins A., Aube W., Péretz I., Barrios F.A., Armony J.L., & Concha L. (2014). Music listening engages specific cortical regions within the temporal lobes: differences between musicians and non-musicians. *Cortex* (59): 126-137

Antón, S. C., Potts, R. y Aiello, L. (2014). Evolution of early Hommo: An integrated biological perspective. *Science* 345, pp. 1-13

Aoki, K. (1986). A stochastic model of gene-culture coevolution suggested by the “culture historical hypothesis” for the evolution of adult lactose absorption in humans. *Proc. Natl. Acad. Sci*. 104: 10944-10949

Apel, W. (1944). Comparative Musicology and Exotic Music*. Harvard Dictionary of Music.* Cambridge, Mass. pp. 167, 250.

Arbib, M.A. ed. (2013). *Language, Music, and the Brain. A Mysterious Relationship*. Cambridge, Massachusetts: The MIT Press.

Armony J.L., Aubé W., Angulo-Perkins A., Péretz I & Concha, L. (2015). The specificity of neural responses to music and their relation to voice processing: An fMRI-adaptation study. *Neuroscience Letters* (593): 35-39

Armstrong, DF & Wilcox, S.E. (2007).*The Gestural Origin of Language*. Oxford University Press.

Ayala, F.J. (1994) *La teoría de la evolución*, Madrid, Ediciones Temas de hoy, 1994

Ayres, B. (1973). Effects of infant carrying practices on rhythm in music. *Ethos* (1): 381–404.

Bangert M. & Schlaug, G. (2006). Specialization of the specialized in features of external human brain morphology. *European Journal of Neuroscience*, 24: 1832–1834

Barkow, J., L. Cosmides, and J.Tooby, eds. (1992). The Adapted Mind. NY: Oxford UP

Bartra, Roger (2006/2012) *Antropología del cerebro*, México, Fondo de Cultura Económica

Bartra, Roger (2013) *Cerebro y Libertad. Ensayo sobre la moral, el juego y el determinismo*, México, Fondo de Cultura Económica

Blake, E. C., & Cross, I. (2015). The Acoustic and Auditory Contexts of Human Behavior. *Current Anthropology, 56* (1), 81-103

Blake, E., & Cross, I. (2008). Flint tools as portable sound- producing objects in the upper Paleolithic context: An experimental study. In P. Cunningham, J. Heeb, & R. Paardekooper (Eds.). *Experiencing archaeology by experiment* (pp. 1–19). Oxford, UK: Oxbow Books.

Blacking, J. (1974), *How Musical is Man?*, Washington: Washington University Press.

Becker, Judith (2011). Rhythmic Entrainment and Evolution. In *Music, Science, and Rhythmic Brain*, N.Y. Roudledge

Bermudez P., Lerch J.P., Evans A.C. & Zatorre R:J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex* 19 (7): 1583-1596.

Berthoz, A. (2000). The brain's sense of movement. Cambridge, MA: Harvard University Press.

Bibikov S. (1978). A stone age orchestra. In D Hunter, P Whitten (eds). *Readings in physical anthropology and archaeology.* London, UK: Harper and Row, pp. 134–148.

Boivin, Nicole (2011). *Material Cultures, Material Minds*. Cambridge University Press

Boyd, R. & Richerson, PJ (2005) *The Origin and Evolution of Cultures*. Oxford: Oxford University Press

Brown, Steven (2000a) "Evolutionary models of music: From Sexual Selection to group selection" en: F. Tonneau & N.S. Thomson, (eds) Perspectives in Ethology, XIII, New York, Plenum, 231-282

Brown, Steven (2000b)"The 'Musilanguage' Model of Music Evolution", en: Nils L. Wallin, Buörn Merker, and Steven Brown (eds.) The Origins of Music, Massachusetts, MIT Press, pp. 271-300

Breuer, B. (2011). The Birth of Musicology from the Spirit of Evolution: Ernst Haeckel's Entwicklungslehre as Central Component of Guido Adler's Methodology for Musicology (Unpublished PhD thesis in Philosophy). USA: University of Pittsburg.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist, 53*(5), 533-548.

Caldwell, D. (2009) Paleolithic whistles or figurines? A preliminary survey of prehistoric phalangeal figurines. *Rock Art Research* 26, pp. 65-82

Caporael L.R., Griesemer J.R. & Wimsatt, W.C. (2014). Developing Scaffolds. An Introduction. En Caporael, L.R., Griesemer J.R. y Wimsatt, W.C. (eds.) *Developing Scaffolds in Evolution, Culture and Cognition*, MIT Press, pp. 1-20

Castiello U, Giordano BL, Begliomini C, Ansuini C, Grassi M (2010) When ears drive hands: the influence of contact sound on reaching to grasp. PLoS One 5(8):e12240.

Changizi M (2011) Harnessed: How language and music mimicked nature and transformed ape to man. BenBella Books, Dallas, TX

Clayton M., Dueck B. & Leante, L (Edits.) (2013). *Experience and Meaning in Music Performance*, Oxford University Press

Clayton, Martin (2013) Entrainment, Ethnography and Musical Interaction. In Martin Clayton, Byron Dueck & Laura Leante (Edits.) *Experience and Meaning in Music Performance*, Oxford University Press

Clayton M., Sager R & Will, U. (2005) In time with the music: the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology*, 11: 3–142.

Clayton, Martin (2012) 'What is entrainment? Definition and applications in musical research', *Empirical Musicology Review*, 7 (1-2): 49-56

Clark, Andy & Chambers, David (1998)"The extended mind" en: Analysis 58:10-23. Disponible en: http://consc.net/papers/extended.html

Clarke, Erik (2005) *Ways of listening. An ecological Approach to Perception of Musical Meaning*, Oxford University Press, New York

Chase, P. & Nowell, A. (1998) Taphonomy of a suggested Middle Palaeolithic bone flute from Slovenia. *Current Anthropology*, 39, pp. 549-553

Chemero, Anthony (2009) *Radical Embodied Cognitive Sciences,* The MIT Press, Massachusetts

Conard, N. Malina, M. & Münzel, S. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature,* Vol. 460, pp. 737-740

Cook, N. (2001). Between Process and Product: Music and/as Performance. Music Theory Online (7):2.

Cook P, Rouse A, Wilson M, Reichmuth C (2013) A california sea lion (Zlophus Californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J Comp Psychol*.127(4):412-27.

Collins, J. E. (2018). Symbolic Arts and Rituals in the African Middle Stone-Age. *Utafiti, 13*(1), 1-22.

Coward, F & Gamble C. (2008) Big brains, small worlds: Material culture and the evolution of mind. In *Philosophical Transactions of the Royal Society of London*, Series B, Biological Sciences, 363: 1969-1979

Cross, I. (2012). Cognitive science and the cultural nature of music. Topics in Cognitive Science, 4, 668–677.

Cross, I., Zubrow, E., & Cowan, F. (2002). Musical behaviours and the archaeological record: A preliminary study. In J. Mathieu (Ed.), Experimental archaeology, British Archaeological Reports International Series 1035, (pp. 25– 34). Oxford: Archaeopress.

Cross, I. (2001) Music, Cognition, Culture and Evolution. *Annals of the New York Academy of Sciences*, Vol 930, pp 28-42.

Cross, I. (2007) Music and cognitive evolution. In Louise Barrett (ed) *Oup Handbook of Evolutionary Psychology*, University of Cambridge

Cross, I. (2018). The Nature of Music and Its Evolution. Susan Hallam, Ian Cross & Michael Thaut (eds). *The Oxford Handbook of Music Psychology,* Second Edition, Oxford University Press, pp. 3-17

Cross, I. (2011). The Nature of Music and its Evolution. In Hallam, S., Cross, I. y M. Thaut (eds.), *The Oxford Handbook of Music Psychology*, Oxford: Oxford University Press, pp. 3-13.

Cross, I. (2008). Musicality and the human capacity for culture. *Musicae Scientiae*, 12, 147–165.

Currie, A. & Killin, A. (2016). Musical pluralism and the science of music. *European Journal in Philosophy of Science*. 6(1): 9–30

Dalhaus, C. (1989). *The idea of absolute music*. Chicago, IL: University of Chicago Press. (Original work published 1978)

Dams, L. (1985) Paleolithic Lithophones: Descriptions and Comparisons. *Oxford Journal of Archaeology* 4, pp. 31-46

Dams, L. (1984) Preliminary Findings at the 'Organ Sanctuary' in the Cave of Nerja, Malaga, Spain. *Oxford Journal of Archaeology* 3, pp. 1-14

Darwin, C. (1871) Chapter XIX. Secondary Sexual Characters of Man. In *Descent of Man and Selection in Relation to Sex*, USA, Barnes & Noble Books, reed.(2004), pp. 505-529

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London. (J. Murray)

Dauvois, M. (1989) Son et musique paléolithiques. *Les Dossiers d'Archéologie* 142, pp. 2-11

Dauvois, M. (1999) Mesures acoustiques at témoins sonores osseux paléolithiques. In *Prehistorie d'os, recuil d' études sur l'industrie osseuse préhistorique offert á Mme Henriette Camps-Febrer.* Aix-en-Provence : Publications de l'Université de Provence

Dawkins, R (1976)The Selfish Gene, Oxford, Oxford University Press

Day RL, Laland KN. Odling-Smee FJ. (2003). Rethinking adaptation: the niche-construction perspective. *Perspect. Bio. Med*.46:80–95

d' Errico & Vila (1998) A Middle Paleolithic origins of music? Using cave-bear bone accumulations to asses the Divje babe I bone 'flute'. *Antiquity*, 72, pp. 65-79 d'Errico, F. et al (2003) Archaeological Evidence for the Emergence of Language, Symbolism, and Music – An Alternative Multidisciplinary Perspective. *Journal of World Prehistory*, Vol. 17, No. 1, pp. 1-70

d'Errico, F. et al (2001) An engraved bone fragment from c.70 000 yeras old Middle Stone Age levels at Blombos Cave, South Africa: Implications for the origins of symbolism and language. *Antiquity* 75, pp. 309-318

d' Errico & Vila (1997) Holes and grooves: the contribution of microscopy and taphonomy to the problem of art origins. *Journal of Human Evolution.* 33, pp. 1-31

Dissanayake, Ellen (2000) "Antecedents of the Temporal Arts in Early Mother-Infant" en: Nils L. Wallin, Buörn Merker, and Stever Brown (eds.) The Origins of Music, Massachusetts, MIT Press, pp.389-410

Donald, Merlin (2009) "The roots of art and religion in ancient material culture" en: Colin Renfrew & Iaian Morley (eds.) Becoming Human: Innovation in Prehistoric Material and Spiritual Culture, USA, Cambridge University Press, pp.95-102

Donald, M. (1993). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition.* Cambridge, Mass: Harvard Univ. Press

Dowling, W & Harwood, Dane L. (1986) Music Cognition, London, Academic Press.

Dupré, J & Nicholson, DJ (2018). A Manifesto for a Processual Philosophy of Biology. In Daniel J. Nicholson & John Dupré (edits). *Everything Flows. Towards a Processual Philosophy of Biology.* Oxford: Oxford University Press.

Elbert T., Pantev C., Wienbruch C., Rockstroh B & Taub E (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270: 305-307

Fahlander, F. & Kjellström A. (2010). Beyond Sight: Archaeologies of Sensory Perception. Fredrik Fahlander & Anna Kjellström (eds). *Making Sense of Things. Archaeologies of Sensory Perception.* Stockholm Studies in Archaeology 53: Stockholm University, pp. 1-13

Feld, Steven (1982). *Sound and Sentiment: Birds, Weeping, Poetics, and Songs in Kaluli Expression*, Filadelfia, reed. 1990

Feldman, M.W. and Cavalli-Sforza, L.L. (1989). On the theory of evolution under genetic and cultural transmission with application to the lactose absorption problem. M.W. Feldman (ed) *Mathematical evolutionary theory.* Princeton, N.J.: Princeton University Press, pp. 145-173.

Fitch, W.T (2018). Four principles of Biomusicology. Henkjan Honing (ed). *The Origins of Musicality*. Cambridge, Massachusetts: The MIT Press, pp. 23-48

Field TM, Woodson R, Greenberg R, & Cohen D. (1982). Discrimination and imitation of facial expression by neonates. Science. 8. 218(4568):179-81.

Flynn et al (2013) Developmental niche construction. *Developmental Science*, 16(2): 296-313

Foley, R.A. (2012) Music and Mosaics. In Nicholas Bannan (ed.) *Music, Language and Human Evolution, Oxford University Press*, pp. 31-57

Freeman, Walter (2000). A Neurobiological Role of Music in Social Bonding. In Nils L. Wallin, Buörn Merker, and Steven Brown (eds.) The Origins of Music, Massachusetts, MIT Press, pp.411-424

Fubini, E. (2008). *Estética de la música*. Madrid, Spain: A Machado Libros.

Fuentes, A. (2007). *Biological anthropology : concepts and connections.* New York: McGraw-Hill

Gallagher, S. (2012). In Defense of Phenomenological Approaches to Social Cognition: Interacting with the Critics. *Review of Philosophy and Psychology 3*(2): 187-212.

Gallagher, S., Martínez, S. F., & Gastelum, M. (2017). Action-space and time: Towards an enactive hermeneutics. In B. B. Janz (Ed.), *Place, space and hermeneutics* (pp. 83– 96). New York, NY: Springer.
Gallese V. & Lakoff G. (2005) The Brain´s concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22 (3/4), pp. 455-479
Gamble, C. et al (2010) The Social Brain and The Distributed Mind. In Dunbar, Robin et al. *Social Brain, Distributed Mind*. Nueva York: Oxford University Press
Gamble, C. (2012) When the Words Dry Up. In Nicholas Bannan (ed.) *Music, Language and Human Evolution*, United Kingdom, Oxford University Press
Gamble, C. (1999). *The Palaeolithic Societies of Europe.* Cambridge: Cambridge University Press.
García de León, A. (comp) (2006). *Fandango. El Ritual del Mundo Jarocho a través de los Siglos*. México City, Méxiico: CONACULTA.
Gibson, J. (1986) *The Ecological Approach to Visual Perception,* Hillsdale, Lawrence Erlbaum Associates
Giordano BL, Visell Y, Yao HY, Hayward V, Cooperstock JR, McAdams S (2012) Identification of walked-upon materials in auditory, kinesthetic, haptic, and audio-haptic conditions. *J Acoust Soc Am* 131(5):4002–4012.
Goldhahn, J. (2002). Roaring Rocks: An Audio-Visual. Perspective on Hunter-Gatherer Engravings in Northern Sweden and Scandinavia. *Norwegian Archaeological Review* 35(1): 29-61
Goldman, A. (1992). The Value of Music. *The Journal of Aesthetics and Art Criticism*, 50(1), 35–44
Gould & Lewontin (1979).The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme, Proceedings of the Royal Society of London, Series B, Vol. 205, No. 1161, pp. 581-598
Gould, S., & Vrba, E. (1982). Exaptation—a Missing Term in the Science of Form. *Paleobiology, 8*(1), 4-15
Godfrey-Smith P. (1996). *Complexity and the function of mind in nature*. New York, NY: Cambridge University Press
Godfrey-Smith P. (2001). Three Kinds of Adaptationism. In S. H. Orzack & E. Sober (eds.), *Adaptationism and Optimality*. Cambridge University Press, pp. 335-357.
Grahn JA, Brett M (2007) Rhythm and beat perception in motor areas of the brain. J Cogn *Neurosci* 19(5):893–906.
Graziano, A., & Johnson, J. K. (2006). The Influence of Scientific Research on Nineteenth-Century Musical Thought: The Work of Richard Wallaschek. *International Review of the Aesthetics and Sociology of Music, 37*(1), 17-32.
Harding, J. (1973) The bull-roarer in history and antiquity. *African Music* 5, pp. 40-42
Hegel, G. W. F. (1975). *Aesthetics: Lectures on fine arts*. Oxford, UK: Oxford University Press.
Henshilwood, C. et al (2001) Blombos Cave, Southern Cape, South Africa: preliminary report on the 1992-1999 excavations of the Middle Stone Age levels. *Journal of Archaeological Science* 28, pp. 421-448
Hernández Azuara, C. (2003) *Huapango. El son huasteco y sus instrumentos en los siglos XIX y XX*. México: CIESAS.
Hernández-Jaramillo, J.M. (2017) *De Jarabes, Puntos, Zapateos y Guajiras. Un sistema musical de transformaciones (Siglos XVIII-XXI).* Unpublished PhD Thesis in Ethnomusicology). Mexico City: National University of Mexico.

Higham, T. et al. (2012) Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: The radiocarbon chronology of GeißenKlösterle. *Journal of Human Evolution,* 62, pp. 664-676

Hiscock, P. (2014). Learning in Lithic Landscapes: A Reconsideration of the Hominid "Tool Making" Niche. Biological Theory 9 (1): 27-41.

Houston, S & Taube K. An Archaeology of the Senses: Perception and Cultural Expression in Ancient Mesoamerica. *Cambridge Archaeological Journal* 10 (2): 261-294

Huron (2001), "Is music an evolutionary adaptation?" In A*nnals of the New York Academy of Sciences*, Vol. 930, pp. 43-61

Husserl, E. (1998). *Thing and Space. Lectures of 1907*. Dordrecht: Kluwer Academic Publishers.

Hutchins, E. (1995). *Cognition in the Wild.* Cambridge MA: The MIT Press

Hutchins, E. (2006) The Distributed Cognition Perspective on Human Interaction. En. N.J. Enfield y S.C. Levinson (eds.). *Roots of Human Sociality. Culture, Cognition and Interaction*, New York: Berg, pp. 375-398

Hutchinson, S., Lee, L. H. L., Gaab, N., and Schlaug, G. (2003). Cerebellar volume of musicians. *Cereb. Cortex* 13: 943–9

Hutto, D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge, MA: The MIT Press.

Ingold, T. (2007). Materials against materiality. *Archaeological Dialogues*, 14, 1–16.

Ingold, T. (1993). The temporality of the landscape. *World Archaeology*, 25, 152–174.

Iriki, A., & Taoka, M. (2012). Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1585), 10-23

Jablonka, E. & Lamb M.J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge, Massachusetts:The MIT Press.

James, William (1890). The Principles of Psychology, New York: Dover Publications

Johnson, Mark (1987) *The Body in the Mind*, Chicago, University of Chicago Press

Johnson, Mark (2007) *The Meaning of the Body. Aesthetics of Human Understanding*. Chicago: The University of Chicago Press

Kant, I. (2000). *Critique of judgment*. New York, NY: Prometheus. (Original work published 1790)

Kapur A, Tzanetakis G, Virji-Babul N, Wang G, Cook PR (2005) A framework for sonification of vicon motion capture data. In: The 8th conference on digital audio effects, Madrid

Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-organization of Brain and Behavior*. Massachusetts: MIT Press

Kendal J, Tehrani JJ, Odling-Smee J. (2011). Human niche construction in interdisciplinary focus. *Philos Trans R Soc Lond B Biol Sci* ;366(1566):785–792.

Kendon, A. (1990) *Conducting Interaction. Patterns of Behavior in Focused Encounters*, Cambridge University Press

Kendon. A. (2009). Language's matrix. Gesture. Vol. 9 (3): 355–372

Killin, A. (2016a). Musicality and the evolution of mind, mimesis, and entrainment. *Biology & Philosophy*, 31, 421–434.

Killin, A. (2017). Plio-Pleistocene foundations of Hominin musicality: Coevolution of cognition, sociality and music. *Biological Theory*, 12, 222–235.

Killin, A. (2016b). Rethinking music's status as adaptation versus technology: a niche construction perspective. In *Ethnomusicology Forum*, Vol. 5, No.2, pp. 210-233

Kim RS, Seitz AR, Shams L (2008) Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One* 3(1):e1532.

Kivy, P. (1997). *Philosophies of Arts: An Essay in Differences. Cambridge*. New York: Cambridge University Press.

Krueger, J. (2011). Doing things with music. *Phenomenology and the Cognitive Sciences*, 10, 1–22.

Kunej & Turk (2000) New Perspectives on the Beginnings of Music: Archaeological and Musicological Analysis of a Middle Paleolithic Bone 'Flute' en: Nils Li. Wallin, Biörn

Merker y Steven Brown (edits). *The origins of music*, MIT Press, pp. 235-268

Kunst, Jaap. (1975/1950). *Musicologica*. Amsterdam: Royal Tropical Institute.

Lakoff, George & Mark Johnson (1980) *Metaphors We Live By.* Chicago, London, The University of Chicago Press

Lakoff, George & Mark Johnson (1980). P*hilosophy in the Flesh: the Embodied Mind & its Challenge to Western Thought*. New York: Basic Books

Laland K.N & O'Brien, MJ (2010). Archaeological Perspectives on Niche Construction *Theory*. *Journal of Archaeological Method and Theory*, 17(4): 303-322

Laland K.N & O'Brien, MJ (2011). Cultural Niche Construction: An Introduction. *Biol Theory.* DOI 10.1007/s13752-012-0026-6

Laland KN & Sterelny K. (2006). Perspective: seven reasons (not) to neglect niche construction. *Evolution* 60:1751–1762.

Laland, Kevin, Odling-Smee & Feldman, Marcus W. (2001). Niche Construction, Ecological Inheritance and Cycles of Contingency in Evolution" en: Oyama, Griffiths & Gray (eds.) Cycles of Contingency. Developmental Systems and Evolution, Massachusetts, The MIT Press, pp. 117-126

Laland K., Wilkins C. & Clayton N. (2016). The evolution of dance. *Current Biology* 26 (1): 5-9

Laland K.N., Uller T., Feldman M.W., Sterelny K., Müller G.B., Moczek A., Jabonkla E., & Odling-Smee J (2015) The extended evolutionary synthesis: its structure, assumptions and predictions. *Proceedings B* 282: 1-14.

Larsson, M. (2015). Too-use-associated sound in the evolution of language. *Animal Cognition*. 18: 993-1005

Larsson, M (2013). Self-generated sounds of locomotion and ventilation and the evolution of human rhythmic abilities. *Animal Cognition* 17(1): 1-14.

Larsson, M. (2015). Tool-use-associated sound in the evolution of language. *Animal Cognition* 18 (5): 993-1005.

Lawson & d'Errico (2002) Microscopic, experimental and theoretical reassessment of Upper Palaeolithic bird-bone pipes form Isturitz, France: ergonomics of design, systems of notation and the origins of musical traditions. In E. Hickman, A. Kilmer, y R. Eichnam (eds.) *Studien zur Misikarchäologie III*. Rahden: Verlag Marie Leidorf, pp.119-14

Leavens DA, Hopkins WD, Bard KA. (1996).Indexical and referential pointing in chimpanzees (*Pan troglodytes*) J*ournal of Comparative Psychology*. 110:346–353

Lee D.J, Chen Y. & Schlaug, G. (2003). Corpus callosum: musician and gender effects. *Neuroreport* 14 (2): 205-209.

Leroi-Gourhan, André (1964) Le Geste et la Parole, Paris, Editions Albin Michel

Levinson, S (2006) On the Human "Interaction Engine" In. N.J. Enfield y S.C. Levinson (Eds) *Roots of Human Sociality. Culture, Cognition and Interaction*, Oxford: BERG

Levinson, S. & Holler, J. (2014) The origins of human multi-modal communication. In *Philosophical Trans*. R. Soc. B 369: 1-9

Lewin R. & Foley R. (eds). (2004). *Principles of Human Evolution*. Blackwell Publishing Company

Lewis-Williams, J. D (2009). Of people and pictures: the nexus of Upper Paleolithic Religion, social discrimination, and art. In C Renfrew, & I. Morley I. (eds) *Becoming*

*Human: innovation in prehistoric material and spiritual culture*, Cambridge, Cambridge University Press, pp. 135-157

Lewis JE, Harmand S. (2016) An earlier origin for stone tool making: implications for cognitive evolution and the transition to Homo. *Phil. Trans. R. Soc. B* 371: 20150233.

Li X, Logan RJ, Pastore RE. (1991). Perception of acoustic source characteristics: Walking sounds. *Journal of the Acoustical Society of America*, 90(6):3036–3049.

Lilliestam, L. (1996). On playing by ear. *Popular Music*, 15, 195–216.

Livingstone, S. & Thompson, W.F. (2009). The emergence of music from Theory of Mind. Music Scientiae, Special Issue 2009/10, “Music and Evolution”, pp. 229-259

Lugarella, M. & Sporns, 0. (2006). Mapping Information Flow in Sensorimotor Networks. *PLoS Comput Biol* 2(10): e144. https://doi.org/10.1371/journal.pcbi.0020144

Lungarella M, Pegors T, Bulwinkle D & Sporns O (2005). Methods for quantifying the informational structure of sensory and motor data. *Neuroinform* 3: 243. https://doi.org/10.1385/NI:3:3:243

Martínez, Sergio (1998a). La síntesis de los conceptos de evolución y mecanismo en las explicaciones por selección natural. En Serio Martínez y Ana Barahona (comp.). *Historia y Explicación en Biología*, México, UNAM-Fondo de Cultura Económica, pp. 301-319

Martínez, Sergio (1998b).Sobre los conceptos de progreso y evolución en el siglo XIX. En Serio Martínez y Ana Barahona (comps). *Historia y Explicación en Biología*, México, UNAM-Fondo de Cultura Económica, pp. 155-167

Martínez, S., & Villanueva, L. A. (2018a). Las prácticas musicales como corporización de tecnologías básicas de la cogniciòn social. *Metatheoria. Revista de Filosofía e Historia de la Ciencia*, 8(2), 1–14.

Martínez, S. & Villanueva, A. (2018b). Musicality as material culture. *Adaptive Behavior*, *26*(5): 257–267

Maturana, H. & Varela, F. (1980). *Autopoiesis and Cognition. The realization of the living.* Boston, MA:Reidel

Matyja J. R. & Schiavio A. (2013) Enactive music cognition: Background and research themes. *Constructivist Foundations* 8(3): 351-357

Mayberry & Jaques (2000) Gesture production during stuttered speech: insights into the nature of gesture-speech integration. In McNeill (Ed.) *Language and Gesture*, Cambridge: Cambridge University Press, pp. 199-214

Mayr, E. (1962). Cause and effect in biology. *Science* 134: 1501-1506.

Mayr E. (1982). *The growth of biological thought: diversity, evolution and inheritance*. Cambridge, MA: Belknap Press

Mc Neil, D. (1992) *Hand and Mind: What Gestures Reveal About Thought,* London: University of Chicago Press

Mcbrearty, & Brooks A.S. (2000). The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution,* 39 (5): 453-563

McNeil, D. (2000) *Language and Gesture*, Cambridge: Cambridge University Press

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL, US: University of Chicago Press.

Merchant, H., Grahn, J., Trainor, L. J., Rohrmeier, M., & Fitch, T. (2018). Finding the beat: A neural perspective across humans and non-human primates. In H. Honing (Ed.), *The origins of musicality* (pp. 171–203). Cambridge, MA: The MIT Press.

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 7, 274.

Merker, Biörn (2000) "Synchronous Chorusing and Human Origins" en: Nils L. Wallin, Buörn Merker, and Stever Brown (eds.) The Origins of Music, Massachusetts, MIT Press, pp. 315-328.

Merlau-Ponty, M. (1962). *Phenomenology of Perception*. London: Routledge and Kegan Paul.

Merriam, Alan P. (1964) *The Anthropology of Music*, Indiana, Northwestern University Press, reed. 1980

Miller, Geoffrey (2000) Evolution of Human Music through Sexual Selection. In Nils L. Wallin, Buörn Merker, and Steven Brown (eds.) *The Origins of Music*, Massachusetts, MIT Press, pp.329-360

Miller, Geoffrey (2009) Spent. Sex, evolution and consuming behavior, New York, Penguin Group

Minetti, A. E. and Alexander, R. M. (1997). A theory of metabolic costs for bipedal gaits. *J. Theor. Biol.* 186,467 -476

Mithen, S. J. (1996) *The prehistory of the mind: a search for the origins of art, religion, and science*, London: Thames and Hudson

Mithen, Steven (2006) *The Singing Neanderthals. The Origins of Music, Language, Mind and Body*, Massachusetts: Harvard University Press

Mithen, Steven (2010) "Excavating the Prehistory Mind:The Brain as Cultural Artefact and Material Culture as Biological Extension" en: Dunbar, Robin et al. Social Brain, Distributed Mind. Nueva York: Oxford University Press

Moran, Nikki (2013) Music, Bodies and Relationships: An ethnographic contribution to embodied cognition studies. *Psychology of Music* 41(1): 5-17

Morley, Ian (2012a) A grand gesture: vocal and corporeal control in melody, rhythm, and emotion. In P. Rebuschat, M. Rohrmeir, J. A. Hawkins & I. Cross (Eds.) *Language and Music as Cognitive Systems,* Oxford University Press

Morley, I. (2006) *Evolutionary Origins and Archeology of Music*, Darwin College Research Report, Cambridge University (PhD Thesis of Philosophy)

Morley, I (2012b) Hominin Physiological Evolution and the Emergence of Musical Capacities. In Nicholas Bannan (ed.) *Music, Language and Human Evolution, Oxford University Press*, pp. 109-141

Morley, I. (2009) "Ritual and music: parallels and practice, and the Paleolithic" en: Colin Renfrew y Iaian Morley (eds.) Becoming Human: Innovation in Prehistoric Material and Spiritual Culture, USA, Cambridge University Press, pp. 159-175

Morley, I. (2013) *The prehistory of music. Human Evolution, Archaeology, and the Origins of Musicality*, Oxford University Press

Münzel, S et al (2002) The GeißenKlösterle flute-discovery, experiments, reconstructions. In E. Hickman, A. Kilmer, y R. Eichman (eds.) *Studien zur Misikarchäologie III*. Rahden: Verlag Marie Leidorf, pp. 107-118

Mugglestone, E. & Adler, G. (1981/1885). Guido Adler's The Scope, Method, and Aim of Musicology. *Yearbook for Traditional Music* (13): 1-21. (English Translation with an Historico-Analytical Commentary)

Müller, G.B. (2017). Why an extended evolutionary synthesis is necessary. *Interface Focus* (7): 1-11

Müller, G.B. (2019). Evo-devo´s challenges to the Modern Synthesis. In Giuseppe Fusco (ed). *Perspectives on Evolutionary and Developmental Biology Essays for Alessandro Minell*i. Padova: Padova University Press, pp. 29-39

Nava, F. (2010). Las (muchas) músicas de los pueblos y las (numerosas) sociedades indígenas. In Aurelio Tello (coord). *La música en México. Panorama del siglo XX*, México: Fondo de Cultura Económica, pp.29-103

Norell, M., Ji, Q., Gao, K., Yuan, C., Zhao, Y., & Wang, L. (2002). 'Modern' feathers on a non-avian dinosaur. *Nature, 416*, 36-37.

Nettl, B. (2000). An ethnomusicologist contemplates universal in musical sound and musical culture. NL. Wallin, Buörn Merker, and Steven Brown (eds.) *The Origins of Music*, Cambridge, Massachusetts: MIT Press, pp. 463-472.

Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me- up for infants' exploratory skills: Early simulated experiences reaching for objects using ''sticky mittens'' enhances young infants' object exploration skills. *Infant Behavior and Development*, 25, 279–295.

Nesbitt K (2003) Designing multi-sensory displays for abstract data. University of Sydney, Australia

Odling-Smee FJ, Laland KN, Feldman MW. (2003). *Niche construction: the neglected process in evolution. Monographs in population biology*, 37. Princeton, UK: Princeton University Press.

Orwin, M., Howes, C. & R. Kempson (2013), *Language, Music and Interaction*, UK: Milton Keynes.

Oyama, Griffiths & Gray (2001) Introduction: What is Development Systems Theory. In Oyama, Griffiths &Gray (eds.) *Cycles of Contingency. Developmental Systems and Evolution*, Massachusstts, Th MIT Press, pp. 1-11

Pastore RE, Flint JD, Gaston JR, Solomon MJ (2008) Auditory event perception: the source-perception loop for posture in human gait. *Percept Psychophys* 70(1):13–29

Panksepp, J. (2009) The emotional antecedents to the evolution of music and language. *Music Scientiae*, Special Issue 2009/10, "Music and Evolution", pp. 83-115

Patel, A.D. (2018). Music as a Transformative Technology of the Mind: An Update. In Henkjan Honing (ed). *The origins of Musicality*. Cambridge, Massachusetts: The MIT Press, pp. 113-126

Patel, A.D., Iversen, J.R., Bregman, M.R., and Schulz I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, 19 (10): 827-830

Patel, A. D. (2010) Music, biological evolution and the brain. In M. Bailar (Ed.), *Emerging Disciplines*. Houston, TX: Rice University Press, pp. 91-144

Peretz, I., Vucan, D., Largois, M.E., and Armony, J. (2018). Neural Overlap in Precessing Music and Speech. Henkyand Honing (ed). *The Origins of Musicality*. Cambridge, Massachusetts: The MIT Press, pp. 205-219

Patel, A.D. (2008). *Music, language, and the brain*. Oxford: Oxford University Press.

Peretz, Isabelle & Robert J. Zatorre (2003) *The cognitive neuroscience of music*, Oxford University Press

Peretz, I., Aubé, W., Armony, J. (2013). Toward a neurobiology of musical emotions. E. Altenmüller, S. Schmidt, and E. Zimmermann (eds*.) The evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man*. New York, NY: Oxford University Press, pp. 277-299

Peretz, I. (2003). Brain specialization for music: new evidence from congenital amusia. Isabelle Peretz & Robert J. Zatorre. *The cognitive neuroscience of music,* Oxford University Press, pp. 192-203

Peretz, I. (2011). Towards a Neurobiology of Musical Emotions. Patrik N. Juslin & John Sloboda (eds.), *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford University Press, pp. 99-126

Peretz, Isabelle & Pascale Lidji (2006). Une perspective biologique sur la nature de la musique. In *Revue de Neuropsychologie* 16 (4):335-386. Disponible en : http://www.brams.umontreal.ca/plab/downloads/Peretz__Lidji_2007.pdf

Peretz, I., Vuvan, D., Lagrois, M.E., and Armony, J. (2015). Neural overlap in processing music and speech. *Philosophical Transactions B*, 30 (1664): 1-8

Péretz, I., Cagnon, L., Macoir, J., and Hébert, S. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception.* 21 (3): 373-390.

Perlman, Marc. 2004. *Unplayed Melodies: Javanese Gamelan and the Genesis of Music Theory*. Berkeley: University of California Press.
Pinker, S. (1997) *How the Mind Works*. New York: W.W. Norton.
Pigliucci M. (2007). *Do we need an extended evolutionary synthesis*? Evolution 61, 2743–2749.
Pigliucci M, Müller GB (eds). (2010). *Evolution. The extended synthesis*. Cambridge, MA: The MIT Press.
Rahaim, M. (2012). *Musicking bodies. Gesture and Voice in Hindustani Music*. USA: Wesleyan University Press
Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour, 1*.
Rehding, A. (2000). The quest for the origins of music in Germany circa 1900. *Journal of the American Musicological Society, 53*(2), 345-385.
Regal, PJ. (1975). The evolutionary origin of feathers. *Q Rev Biol* 50:33–66
Renfrew, C. & Morley I. (eds) (2009) *Becoming Human: innovation in prehistoric material and spiritual culture*, Cambridge, Cambridge University Press
Renfrew, C. (2012). Towards a cognitive archaeology: material engagement and the early development of society. In: I. Hodder (ed), *Archaeological theory today*. Second edition,Polity Press, pp. 124- 145.
Reybrouck, Mark (2005) Body, Mind and Music: musical semantics between experiential cognition and cognitive economy. In TRANS-Revista Transcultural de Música, 9 (Artículo 13).
Reznikoff, Iégor & Michel Dauvois (1988) "La dimensión sonore des grottes ornées". En: Bulletin de la Société préhistorique française, Tome 5, N.8, pp. 238-246
Reznikoff, Iégor (2005) « On primitive elements of musical meaning ». En : JMM The Journal of Music and Meaning. Vol. 3
Reznikoff, Iégor (2008) «Sound resonance in prehistoric times : A study of Paleolithic painted caves and rocks ». En: The Journal of the Acoustical Society of America, vol. 123, issue 5, pp. 3137-4141
Rice, T (2014). *Ethnomusicology: A Very Short Introduction*. Oxford: Oxford University Press
Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. Ecological Psychology, 26, 325–352.
Roederer, Juan G. (1995) *Acústica y Psicoacústica de la Música*, Buenos Aires, Ricordi
Rousseau, J.J. (2000). *Essay on the Origin of Languages and Writings Related to Music*. Dartmouth College Press
Savage P.E., Brown S., Sakai E., & Currie T.E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences* 112 (29): 8987-8992.
Sachs, C. (1962). *The Wellsprings of Music*. Edited by Jaap Kunst. The Hague: M. Nijhoff.
Sánchez García, R. (2002). Diferencias formales entre la lírica de los sones huastecos y la de los sones jarochos. Revista de Literaturas Populares CENIDIM-INBA 2(1): 121-152.
Schiavio, A., van der Schyff, D., Kruse-Weber, S., & Tim- mers, R. (2017). When the sound becomes the goal. 4E cognition and teleomusicality in early infancy. Frontiers in Psychology, 8, 1585.
Seeger (1987) *Why Suyá Sing: A Musical Anthropology of Amazonian People*. USA, Library of Congress, reed. 2004
Scarre C & Lawson G. eds. (2006). *Archaeoacoustics*. Cambridge: McDonald Institute for Archaeological Research.

Shams L, Seitz AR (2008) Benefits of multisensory learning. Trends Cogn Sci 12(11):411–417.
Schelling, F. W. J. (1999). *Filosofía del arte*. Madrid, Spain: Tecnos.
Sheya, A., & Smith, L. B. (2010). Development through sensorimotor coordination. In J. Stewart, O. Gapenne, & E. A. Di Paolo (Eds.), *Enaction: Toward a new paradigm for cognitive sciences* (pp. 123–144). Cambridge, MA: The MIT Press.
Schmidt, R. A., & Lee, T. D. (2011). *Motor control and learning: A behavioral emphasis* (5th ed.). Champaign, IL, US: Human Kinetics.
Schopenhauer, A. (1969). *The world as will and representation* (Vol. 1). New York, NY: Dover Publications.
Seeger, A. (1987). *Why Suya sing: a musical anthropology of an amazonian people*. Urbana and Chicago: University of Illinois Press.
Sigrist R, Rauter G, Riener R, Wolf P (2012) Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychon Bull Rev*.
Scott-Phillips, T. C., Laland, K. N., Shuker, D. M., Dickins, T. E., & West, S. A. (2014). The niche construction perspective: a critical appraisal. *Evolution; international journal of organic evolution*, *68*(5), 1231–1243.
Sloboda, J. A. & O'Neill, S. A. (2001). “Functions of music in everyday life: An exploratory study using the experience sampling methodology” en: Musicae Scientiae, 5, pp. 9-32
Small, C. (1998). *Musicking: The meanings of performing and listening*. Middletown, CT: Wesleyan University Press.
Smith, L. B., & Pereira, A. F. (2009). Shape, action, symbolic play, and words: Overlapping loops of cause and consequence in developmental process. In: S. Johnson (Ed.), *Neo‑constructivist approach to early development*. (pp. 109–131). New York: Oxford University Press.
Spencer, H. (1904). *Essays. Scientific, Political and Speculative*, Vo. II, New York: D. Appleton and Company, Library Edition.
Stobar, H. (1996) The Llama’s flute: musical misunderstandings in the Andes. *Early Music* 24: 470-482
Sterelny, Kim (2012) *The Evolved Apprentice: How Evolution Made Humans Unique*, Cambridge MA: The MIT Press
Sterelny, K (2003). *Thought in a Hostile World: the Evolution of Human Cognition*. Oxford: Blackwell Publishing Ltd
Swift, G. N. (1990). South Indian ‘‘Gamaka’’ and the violin. *Asian Music*, 21(2), 71–89.
Thaut MH, McIntosh GC, Rice RR, Miller RA, Rathbun J, Brault JM.(1996). Rhythmic auditory stimulation in gait training for Parkinson's disease patients.*Mov Disord.* 11(2):193-200.
Tëmkin, I. (2004). Evolution of the Baltic psaltery: A case for phyloorganology?*The Galpin Society Journal*, 57, 219–230.
Thompson, E. (2007). *Mind in Life. Biology, Phenomenology, and Sciences of Mind*. Cambridge, MA: Harvard University Press
Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M. et al. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Gen*. (39): 31-40.
Tomasello, M. (2008) *Origins of Human Communication*, Cambridge, UK: MIT Press
Tomasello, Michael (1999) *The Cultural Origins of Human Cognition*, London, Harvard University Press
Tomlinson, G. (2015). *A million years of music: The emergence of human modernity*. Cambridge, UK: Cambridge University Press.

Tomlinson, G. (2001). Musicology, Anthropology, History. *Il Saggiatore musicale* 7(1), 21–37.

Trehub, S & Trainor L. J. (1998). Singing to infants: lullabies and play songs. *Adv. Infancy Res.* (12): 43–77.

Trehub S.E., Becker J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions B.* 370: 1-9

Trehub, S.E. (2001). Musical Predispositions in Infancy. *Ann. NY Acad. Sci*. 930: 1-16

Trehub, S.E. (2000) "Human Processing Predisposition and Musical Universal", en: Nils L. Wallin, Buörn Merker, and Steven Brown (eds.) The Origins of Music, Massachusetts, MIT Press, pp.427-448

Tuniz, C., Bernardini, F., Turk, I., Dimkaroski, L., Mancini, L. and Dreossi, D.(2012) Did Neanderthals Play Music? X-Ray Computed Micro-Tomography of the Divje Babe 'Flute'. *Archaeometry*, 54, 3, pp. 581-590

Turino, T. (2008). *Music as social life: The politics of participation*. London: University of Chicago Press

Turk et al (2006) Result of computer tomography of the oldest suspected flute from Divje babe I (Slovenia) and its chronological position within global palaeoclimatic and palaeoenvironmental change during Last Glacial. *L'Anthropologie*, 110, pp. 293-317

Tyrberg, T. (1998) *Pleistocene Birds of the Palearctic: A Catalogue*. Publications of the Nuttall Ornithological Club, No. 27, Cambridge MA

Ujhelyi, Molino & Brown (2000). Social Organization as factor on the Origins of Language and Music. In Nils L. Wallin, Buörn Merker, and Stever Brown (eds.). *The Origins of Music*, Massachusetts, MIT Press, pp. 125-134

Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge MA: MIT Press

Van der Schyff, D., & Schiavio, A. (2017). Evolutionary musicology meets embodied cognition: Biocultural coevolution and the enactive origins of human musicality. *Frontiers in Neuroscience*, 11, 519.

Villanueva, L. A. (2012) *El trio huasteco en la comunidad totonaca del Municipio de Huehuetla*, Puebla (Unpublished MA thesis in Ethnomusicology). Mexico City: National University of Mexico.

Waller, Steven J. (1987) « Rock Art Acoustics in the Past, Present and Future »En: Peggy Whitehead and Lawrence Lowendorf (Eds.)International Rock Art Congress (IRAC) Proceedings, Volume 1, AIRA Volume 26, pp.11-20

Waller, Steven J. (1993) « Sound reflections as an explanation for the content and context of rock art ». En : The Journal of Rock Art Research, Vol. 10, pp. 91-111

Whalen, A., Cownden D., Laland K (2015). The learning of action sequences through social transmission. Anim. Cogn. 18 (5): 1093-1103.

Wallaschek, R. (1893). *Primitive Music*. London: Longmans, Green & Co.

Wallin, Nils Lennart (1991) *Biomusicology: neurophysiological, neuropsychological, and evolutionary perspectives on the origins and purposes of music*. New York: Pendragon Press.

Wallin, Nils L., Buörn Merker, and Steven Brown (eds.) (2000) T*he Origins of Music Massachusetts*. MIT Press, pp.329-360

Weidman, A. (2012). The ethnographer as apprentice: Embodying sociomusical knowledge in South India. *Anthropology and Humanism*, 37, 214–235.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. Proceedings of the National Academy of Sciences of the United States of America, 106, 2468–2471.

Will, U. & Turow, G. (2011). Introduction to entrainment and cognitive ethnomusicology. In: G. Turow & J. Berger (Eds), *Music, Science and the Rhythmic Brain. Cultural and Clinical Implications*. New York: Routledge, pp. 3-30

Williams GC (1992) Gaia, nature worship, and biocentric fallacies. Q Rev Biol 67:479–486

Wilkins, J.F., & Godfrey-Smith, P. (2009). Adaptationism and the adaptive landscape. *Biology & Philosophy, 24*, 199-214.

Wimsatt, W. C. & James R. Griesemer (2007) Reproducing entrenchments to scaffold culture: The central role of development in cultural evolution. In Sansom, Roger & Robert Brandon (edits) *Integrating Evolution and Development: From Theory to Practice*. Cambridge: MIT Press

Wray. A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics* 21(4):463–489

Wynn, T. (2002). Archaeology and Cognitive Evolution. *Behavioral and Brain Sciences* 25(3): 389–438.

Wynn, T. (1991). Tools, Grammar, and the Archaeology of Cognition. *Cambridge Archaeological Journal* 1: 191-206

Wynn, T. & Coolidge, F.L. (2010). How Levallois reduction is similar to, and not similar to, playing chess. In April Nowell e Iain Davidson (eds). *Stone Tools and the Evolution of Human Cognition*, Colorado: University Press of Colorado, pp.83-103

Yamamoto S, Yamakoshi G, Humle T, Matsuzawa T (2008) Invention and modification of a new tool use behavior: ant-fishing in trees by a wild chimpanzee (*Pan troglodytes verus*) at Bossou, Guinea. Am J Primatol 70: 699–702.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8, 547–558.

Zubrow, E., Cross, I., & Cowan, F. (2001). Musical behaviour and the archaeology of the mind. *Archaeologia Polona*, 39, 111–126.