

Music in (en)action

Sense-making and neurophenomenology of
musical experience

Andrea Schiavio

May 2014

Submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy in the Department of Music

The University of Sheffield

*This work is dedicated to my parents
who always believed in me*

I hereby declare that this submission is my own work and it does not contain any material previously published or written by another person, except where explicitly acknowledged. In particular, this thesis contains entirely or parts of the following papers, essays and work of mine:

Schiavio, A. (2010). *Intenzionalità e cognizione motoria nella musica* (Unpublished MA dissertation). Università Degli Studi, Milan, Supervisor Prof. Sinigaglia, C. (in italian).

Schiavio, A. (2012). Constituting the musical object. A neurophenomenological perspective on musical research. In *Teorema*, 31(3), pp. 63-80.

Menin, D., & **Schiavio, A.** (2012). Rethinking musical affordance. In *AVANT. The journal of the philosophical-interdisciplinary vanguard*, 3(2), pp. 202-215.

Schiavio, A. (2013). Review: Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness, and Language. In *Philosophical Psychology*. Available at <http://www.tandfonline.com/doi/abs/10.1080/09515089.2013.852528?journalCode=cphp20#.Usl78mTuL6k>.

Schiavio, A., & Menin, D. (2013). Embodied music cognition and the mediation technology. A critical review. In *Psychology of Music*, 41(6), pp. 804-814.

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

Schiavio, A. (in preparation). Music, Evolution, Mirror Neurons.

Schiavio, A., & Gerson, S. (in preparation). An enactive perspective on infants' musicality.

Schiavio*, A., Gerson*, S., Timmers, R., Bekkering, H., & Hunnius, S. (in preparation). Sensorimotor integration and action-effect contingency in infants' engagement with a musical instrument. (*both first author).

Schiavio, A., & Høffding S. (in preparation). Playing together without communicating? A pre-reflective account of joint musical attention.

Schiavio, A., Menin, D., & Matyja, J. (in preparation). Music in the flesh. Sense-making and embodied simulation in musical understanding.

Schiavio, A., & Timmers, R. (in preparation). Cross-modal facilitation in learning atonal melodies.

Thanks to the editors, publishers and co-authors for permission to use the published material here.

Abstract

MUSIC IN (EN)ACTION SENSE-MAKING AND NEUROPHENOMENOLOGY OF MUSICAL EXPERIENCE

by

Andrea Schiavio

The aim of this work is to lay the basis of a *post-Cartesian* cognitive science of music. Traditional psychology of music often adopts a theoretical framework in line with the dualistic stance characterising the Cartesian approach, which implies a separation between mind and matter or, in its materialistic version, a separation between brain and body. I criticize such a paradigm on the basis of theoretical and empirical evidence, showing that alternative models of human musicality offer more plausible explanations without any dichotomy between objective/subjective and internal/external. The thesis that I will defend throughout this work holds that musical cognition is not something that occurs in our head. Rather, it is a process that extends beyond the boundaries of skull and skin, being constituted by the dynamic interplay between embodied agents and the environment in which they are embedded. I will defend such a claim through an interdisciplinary approach that lies at the intersection of different fields of research (cognitive neuroscience, philosophy of mind, phenomenology) and by providing an original interpretation of the *enactive* paradigm that emerged during the last decade of the Twentieth Century in the realm of cognitive science.

Acknowledgments

Special thanks are due to my family, for their love, generous care and for the inspiration that they keep on providing.

In the Music Department of The University of Sheffield I have always been able to ask for comments, opinions, and general feedback to the whole staff and never been disappointed. My experiences here have always succeeded to open my mind, exploring ideas, music and perspectives that played a crucial role for this dissertation. Nikki Dibben, Stephanie Pitts, George Nicholson, Dorothy Ker and Tim Shephard have enabled me to develop into a better thinker, writer, teacher and, in one word, person.

I am thankful to my colleagues and friends Yuko Arthurs, Julian Cespedes, Ioanna Filippidi, Ben Gaunt, Nicolas Farrugia, Giacomo Novembre, Damiano Menin, Simon Høffding and Jakub Ryszard Matyja. They have provided invaluable companionship throughout the project as well as fostering many ideas that we eventually developed and published.

I also wish to thank the various scholars I have contacted, met at conferences or other events: Corrado Sinigaglia, Eric Clarke, Fred Cummins, Evan Thompson, Dan Hutto, Ezequiel Di Paolo, Alva Noë, Giovanna Colombetti, Mark Reybrouck, Joel Krueger, Peter Keller, Sonja Kotz, Gary McPherson, Teed Rockwell, Lawrence Parsons, Gaye Soley and Katie Overy. Their willingness to comment upon my work, providing me with different opinions and argumentative strategies has been extremely helpful in many moments of my research. In particular I would like to thank Sabine Hunnius and Sarah Gerson, who hosted me in Nijmegen for some months at the Donders institute for Brain, Cognition and Behaviour. Working with them represented a real turning point for my work.

I express my infinite gratitude to my girlfriend Ayşegül Can for her longstanding love, support, encouragement and patience. *Senin sevgin olmadan yasayamam, aşkım!*

But most of all I am grateful to my supervisor, Renee Timmers, whose enthusiasm for my research came at the very beginning, and substantially boosted my own. Both as a person and a scholar, her open-minded approach combined with a rigorous attention for details helped me to face every challenge and obstacle with the necessary confidence and preparation. I am privileged to have been a beneficiary of her mentorship.

Table of Contents

1) Ghosts, dreams and nightmares	P. 8
2) <i>Diabolus in musica</i>	P. 23
3) The mark of embodiment	P. 42
4) Music beyond mirroring	P. 70
5) Exploring musicality	P. 99
6) From the cradle to the stage	P. 121
7) Conclusion. Towards a post-Cartesian cognitive science of music	P. 141
Bibliography	P. 147

Chapter 1

Ghosts, dreams and nightmares¹

Neurophysiologists, most of them, are still under the influence of dualism, however much they deny philosophizing. They still assume that the brain is the seat of the mind. To say, in modern parlance, that it is a computer with a program, either inherited or acquired, that plans a voluntary action and then commands the muscles to move is only a little better than Descartes' theory, for to say this is still to remain confined within the doctrine of responses

- James Gibson, *The ecological approach to visual perception*

Too often [...] the older dualism of soul and body has been replaced by that of the brain and the rest of the body

- John Dewey, *Democracy and education*

1.1 - Introduction. Reconceptualizing the mind

The thesis that I offer in this work holds that musical cognition is not something that occurs purely in our head. Rather, it's a process that reaches out beyond the boundaries of skull and skin, being constituted by the dynamic interplay between embodied agents and the environment in which they are embedded. I will defend this claim by means of the employment of an interdisciplinary approach that lies at the intersection of different fields of research (cognitive neuroscience, philosophy of mind, phenomenology, music cognition) and by providing an original interpretation of the *enactive* paradigm that emerged during the last decade of the Twentieth Century in the realm of cognitive science (Varela et al., 1991).

Generally, this dissertation can be seen as an attempt to provide a biologically plausible interpretation of the basis of human musicality. Dealing extensively with theoretical argumentations as well as empirical data I will critically analyse part of the current literature on music psychology and propose an alternative view on a variety of themes related to how human beings *experience* and

¹ This chapter contains material previously published as **Schiavio, A.** (2012). Constituting the musical object. A neurophenomenological perspective on musical research. In *Teorema*, 31(3), pp. 63-80. Moreover, it contains material from the following article in preparation: **Schiavio, A., & Gerson, S.** (in preparation). An enactive perspective on infants' musicality.

make sense of music. In these terms, my purpose is firstly to show how some leftovers from Cartesianism vitiate the current scientific study of musicality, and only secondarily do I try to convince the reader that what I address is the right way to overcome such a tradition. Finally, I lay out the philosophical and empirical implications of investigating music cognition through a different - enactively grounded - approach. To my mind, it is a prerogative of a modern science of music cognition to integrate its own advances with a continuous and rigorous confrontation of the results obtained by the new perspectives on the nature of the mind. During the last years, in fact, cognitive science has witnessed the foundation and development of different, yet intrinsically related, perspectives that aimed to reconceptualise our understanding of the mind and subjectivity. These approaches, labelled *embodiment*, *externalism*, *embedment* and *enactivism* are all committed to build up a *new science of mind* (Foolen et al., 2012), putting together empirical and theoretical evidences from many different disciplines. A *post-Cartesian cognitive science of music*, therefore, cannot avoid a deep confrontation with the main themes and perspectives that characterise such a new and challenging framework. Not only, in fact, will its critical analysis surely contribute to improving our understanding of human musicality, but the study of how a cognizer makes sense of music can also shed light on many aspects that traditional cognitive science might ignore. Perception, generally, is mostly treated in terms of *visual* perception, and a critical phenomenological analysis of other perceptual domains, enriched by the recent advances in cognitive science, is still missing.

The main aim of the scholars who modelled and developed the theoretical frameworks at the basis of the new science of mind is to propose a *decentralization of cognition* from one's head, by focusing on how the bodily power of action and its situated coupling with the environment play a constitutive role in mental processes. Many contemporary authors maintain a position for which 'the human mind is embodied in our entire organism and embedded in the world, and hence is not reducible to structures inside the head' (Thompson, 2005, p. 409). This shared standpoint radically challenges the paradigm that has dominated cognitive science from the second half of the last century: *cognitivism*. This latter view, broadly speaking, holds that an agent's mind can be seen as a computer, as it assumes that mental phenomena 'are effective or algorithmic processes in the sense of computability theory, i.e. processes that - according to Church's Thesis - compute partial recursive (or, equivalently, Turing-computable) functions' (Cordeschi & Frixione, 2007, p. 38). According to this standpoint, the realm of the mind begins and ends up inside the boundaries of the skull. In contrast, all these new approaches take a different direction, rejecting the mind-computer metaphor that characterises the cognitivist perspective as well as its *Cartesian* presuppositions. Proponents of the mainstream cognitive standpoint - as we will see - share with Descartes a particular view on the nature of mind. A view that is no longer sustainable: 'The Cartesian tradition is mistaken in supposing that the mind is an inner entity of any kind, whether mind-stuff, brain states, or whatever. Ontologically, mind is much more a matter of what we do within environmental and social possibilities and bounds. Twentieth-century anti-Cartesianism thus draws much of mind out, and in particular outside the skull' (van Gelder, 1995, p. 380). In order to get a very basic and superficial idea of these *anti-Cartesian* views, we could introduce the following, intuitive, assumptions

- Cognition does not depend solely on brain processes, but results from structures widely distributed across the whole body of a living system (the mind is *embodied*).

- Cognition arises from interactions with the (social and physical) environment, being a category actively immersed in the world (the mind is *embedded*).
- Cognition can reach beyond the boundaries of skull and skin, integrating resources internal and external to the animal (the mind is *extended*).
- Cognition is sense-making, being an emergent, skilful know-how, constituted by embedded and embodied forms of interactions between a self-organized living system and its environment. Through this dynamic interplay, the creature *enacts* or *brings forth* its own domain of meaning (the mind is *enacted*).

It is quite often assumed that these paradigms could theoretically be endorsed all together as a whole, hence defining a ‘4 Es’ approach (Ward & Stapleton, 2012). According to this standpoint, saying that the mind is **Enacted**, would automatically imply that the mind is also **Embodied**, **Embedded** and (at least potentially) **Extended**. I agree only in part. The position that I will defend in this thesis draws from all of these paradigms, taking from each of them some of their main aspects and standpoints. However, I will not embrace all of them at the same time. In fact, the variety of externalist proposals (Hurley, 2010), the different interpretations of embodied approaches (Gallagher, 2011) and the impossibility to define a unique interpretation of the enactive perspective (i.e. Hutto & Myin, 2013), would substantially compromise such a unitary view. In particular, the differences between externalism and enactivism, as Di Paolo (2009) points out, seem to rule out many levels of possible interaction. Nevertheless, it is surely correct that there is some continuity among each standpoint, if we are to consider only the basic above-mentioned assumptions. Simply put, when considering the body of an agent as an inseparable category from the mind in order to understand cognition (embodiment), we cannot but think of this body as a situated entity (embedment). But being a body-situated-in-the-world means being part of a dynamic interplay with the environment (enaction). And, finally, this co-determinative processes between a living system (with its recurrent patterns of sensorimotor coupling) and its world (intended as its relational domain) would imply that internal and external resources are integrated in a way that would enable a subject to accomplish her goals thanks to a variety of external scaffoldings, *extending* the boundaries of skull and skin (externalism).

If this line of reasoning were as straightforward and flowing as this oversimplification portrays, the task of this work would be much easier. Unfortunately, the story is not that trivial, considering that not all of these categories, as previously mentioned, can be easily intermixed. Such a plan would inevitably fall short due to the distinctive and peculiar aspects of each theoretical framework. Some of the arguments used by proponents of the extended mind thesis, as we will see, stand in open contrast with a truly enactive characterisation of cognition, and might eventually collapse into a dualistic account that considers body and mind as separate and distinct entities (see Thompson & Stapleton, 2009). Moreover, some types of embodied approaches seem to be constituted on a very *internalistic* vision of cognitive processes, which are seen as widely distributed across the whole body but not necessarily coupled with external resources (i.e. Adams & Aizawa, 2010). The debate on mirror neurons, their interpretation and functional role in understanding the actions, emotions

and intentions of other individuals will exemplify this problem in the following chapters. Given these considerations, this work will mostly focus on enaction. This theoretical framework is in fact crucial for coherently maintaining a clear anti-dualistic position without any risk to be committed to a mere internalistic stance. After a brief overview of the enactive standpoint, this chapter will focus on the clarification of the conceptual topography of the Cartesian cognitive science, its philosophical foundations and its relevance for the development of the current leading perspectives in music psychology.

1.2 - A world of meaning

Enactivism is a term used by Varela, Thompson and Rosch (1991) to define a new theoretical and empirical paradigm in the realm of cognitive science, based on the codetermination between embodied agents and their world. This approach considers cognition as an emergent phenomenon constituted by embedded and embodied forms of interactions between a self-organized living system and its environment. According to Thompson and Stapleton (2009), one of the main goals of the enactive perspective is to explore how a living system is organized in order to reach *autonomy*, namely the ability to sustain itself under precarious conditions (Maturana & Varela, 1980; De Jaegher & Di Paolo, 2007). Autonomy is essentially what allows living systems to regulate their interactions with the environment, as ‘an autonomous system is a system composed of processes that generate and sustain that system as a unity and thereby also define an environment for the system’ (Thompson and Stapleton, 2009, p. 2). In order to achieve an ideal condition of (autonomous) viability, an organism transforms the world into a place of salience, where the meanings enacted rely on the degree of complexity of the agent (Colombetti & Torrance, 2009, p. 507). In other terms, a living organism, rather than being a mere spectator of passively witnessed events, makes sense of the world by *enacting* or *bringing forth* his or her own perspective on the basis of his or her adaptive autonomy. This ascription of meaning regards the animal as a whole organism - thus also in terms of thermodynamic requirements and not only considering intellectual processes - and is mainly defined through the actions performed in the environment, which intrinsically entwine the living organism with the world in a normative way (Varela 1997; Di Paolo 2005).

The organism and environment are not actually separately determined. The environment is not a structure imposed on living beings from outside but is in fact a creation of those beings. The environment is not an autonomous process but a reflection of the biology of the species (Lewontin, 1983, p. 99).

As Sheets-Johnstone (2012) puts it, the agents’ *animation* makes the world intelligible for them, and this principle constitutes the meaningful core of the animal’s situatedness in the world. In fact, ‘animate beings do not simply move in an epiphenomenal sense: they are born to move; they are moved to move; they move *meaningfully*’ (ibid., p. 34). The centrality of action for sense-making emphasises the view that considers our experiences of the world as *pragmatic* in nature - being

mediated by goal-directed sensorimotor patterns.

The enactive approach considers the *experienced* world as determined by the dynamic coupling between an animal's physiology, its sensorimotor organization and the environment in which it is situated, thus addressing perception and action not as two separate categories, but rather, as a unitary inseparable entity. If cognition and perception are dynamic sensorimotor activities, and the environment is not conceived of as the result of neural computations, then an agent's world is a result of the *enacted* experience of the organism's bodily engagement within it and the taken-for-granted division between internal (neural) states and a pre-given external world therefore should be ruled out. As I see it, the enactivist perspective represents a challenge for the traditional cognitive science of music, as it could provide a radically different model to understand the relationship that ties a musical subject and a musical object. After all, every musical experience occurs within this relationship. I can imagine music, I can listen to it, analyse it, perform it or learn it. Outside of this relationship there is no musical experience. The problem therefore is to define the *terms* of this relationship. Is it a causal relationship, where, for example, musical objects and musical subjects are independent categories? Or is the relationship itself a codetermination between the two, where subjects and objects are inseparable and thus impossible to be studied distinctly? How exactly are these two categories related? And how could we define "musical subjects" and "musical objects"? Before answering these questions and proposing an enactive approach to musical experience, however, I will provide a detailed analysis of the previous models employed to understand human cognition, its characterisation, and its influence on current music psychology.

1.3 - A dream and a nightmare

When describing the notion of "Cartesian Anxiety" firstly proposed by Bernstein (1983) and then reintroduced by Varela and colleagues (1991), Steve Torrance uses the following, striking, metaphor: 'This anxiety generates a dream and a nightmare - the absolutist dream of achieving a guarantee of objective truth in our internal representations, and the nihilist nightmare that such a guarantee is forever beyond us' (Torrance, 2005, p. 3). This *anxiety*, broadly speaking, consists in assuming that "to know *x*" means "to have a representation of *x*". Torrance perfectly portrays the consequences of this equivalence: in order to provide a plausible description of a given phenomenon, cognitive science should give a reliable account of our internal representations, as they portray the concrete and objective properties of the world that we experience. On the other hand, the focus on mental objects leads *de facto* to the investigation of a realm that is not physical and therefore not objectively measurable. This contrast generates an unsolvable tension: no account of human cognition has ever provided a plausible and falsifiable explanation of how a physical entity could generate a non-physical representation (Nagel, 1974). As David Chalmers puts it:

It is undeniable that some organisms are subjects of experience. But the question of how it is that these systems are subjects of experience is perplexing. Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an

emotion? It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does (Chalmers, 1995, p. 203).

This is what Chalmers himself labels as the “hard problem of consciousness”, that is, simply put, the problem of experience. There is no doubt, in fact that ‘lived, *first-hand*, experience is a proper field of phenomena, irreducible to anything else’ (Varela, 1996, p. 347) and that the possibility of bridging the *explanatory gap* (Levine, 1983; Searle, 1992) between the qualitative character of subjectivity and its purely physical and functional correlates has enamoured many scholars. Many contemporary philosophers and neuroscientists, for example, maintain an *identity theory* in contrast to dualism. For them, a subjective feeling *is* neural activity (i.e. Churchland, 1986).

However, as we will see, the identity between mind and brain is not the only way to reject the Cartesian dualism. This is the reason why solving this explanatory gap is not the goal of this dissertation. Rather, what I am actually interested in, is to provide a coherent critique of the *dualistic* foundations of such an issue, advocating an account that does not admit any contrast between *external* and *internal*, *subjective* and *objective*. Any attempt to bridge the gap, as I see it, would only perpetuate its aprioristic assumptions, thus giving rise to a circular scenario. A perfect example of this fallacy can be seen in the book *Embodied Music Cognition and the Mediation Technology* (Leman, 2007) where the author considers the body as the *mediator* between musical experience (mind) and sound energy (matter), thus remaining committed to the Cartesian perspective that he himself explicitly rejects (ibid., p. 13). Here, like in other cases, the hypostatization of the starting dichotomy (between mind and matter) necessitates an *ad hoc* solution (body-as-mediator) in order to mediate between the two different categories. I will come back to Leman’s interpretation of embodiment and music cognition in Chapter 4. Before doing that, however, a clarification seems necessary. Given that the Cartesian legacy is, in fact, ‘both widespread and tenacious, not only as an explicit doctrine but, more significantly, in the clandestine influence it has on explicit doctrines of the mind’ (Rowlands, 1999, p. 3), it seems reasonable to have a closer look into its principles and main implications, as it still constitutes a fundamental source of inspiration for many of the contemporary accounts of musical cognition.

1.4 - The dogma of the ghost in the machine

The view endorsed by Descartes is generally called “Substance Dualism” as it proposes a radical dichotomy between two different substances: *mind*, which has no physical extension and is defined by the property of “thinking”, and *matter*, whose property is “extension”. A *substance* can be defined not as a collection of properties but, rather, as the actual entity that possesses these properties. Therefore the human mind is not a mere collection of beliefs, mental states and so forth but rather it is *that thing whose property is to think*. While the human mind is a thinking entity, governed by the principles of reason, matter (the physical body, the world) is ruled by mechanistic principles (Rowlands, 1999). Descartes maintains a profound asymmetry between the ways of

knowing and understanding the human mind and the world. While the mind is in fact an internal category, the latter is an external one. As Gilbert Ryle puts it:

There is a doctrine about the nature and place of the mind, which is prevalent among theorists, to which most philosophers, psychologists and religious teachers subscribe with minor reservations. Although they admit certain theoretical difficulties in it, they tend to assume that these can be overcome without serious modifications being made to the architecture of the theory. [The doctrine states that] with the doubtful exceptions of the mentally-incompetent and infants-in-arms, every human being has both a body and a mind. [...] The body and the mind are ordinarily harnessed together, but after the death of the body the mind may continue to exist and function (Ryle, 1949, p.11).

A classic line of reasoning for proponents of this form of dualism is based on the well-known principle called *Leibniz's law* or the *indiscernibility of identicals* (Braddon-Mitchell & Jackson, 1997, p. 5). If we assume that $x = y$, then we have to admit that every property P of x is also a property P of y , and every property P of y is also a property P of x .

$$x = y \rightarrow \forall P(Px \leftrightarrow Py)$$

Probably the most famous argument proposed by Descartes to defend his thesis is the one that employs this law in relation with the notion of *doubt*. Descartes (argument A) holds that (i) you can doubt about the existence of your body, but (ii) you cannot doubt about the existence of your mind. It follows that since mind and body display *different properties* then they must be distinct entities. The two premises i and ii, however, raise several problems when defining *being such that I can doubt its existence* as an actual property of the two objects - mind and body. How can *being an object of doubt* be identified with a real property of a given substance? Let's consider the following analogous argument (B). I can have doubts about Clark Kent being Superman, but this will not change the fact that Clark Kent *is* Superman. To clarify this point, let's consider the two arguments A and B as follows:

A)

- (i) I cannot doubt that I am my mind
 - (ii) I can doubt that I am my body
- Therefore \rightarrow Mind and body are two different entities

B)

- (i) I cannot doubt that Clark Kent is Clark Kent
- (ii) I can doubt that Clark Kent is Superman

Therefore → Clark Kent and Superman are two different entities

As we can notice from this traditional (i.e. Goldberg & Pessin, 1997) counterexample, the conclusion of B is clearly mistaken. It seems, therefore, that we do not have any reason to consider the *doubt argument* as a valid candidate to demonstrate the separation between mind and body. Being a *propositional attitude*, doubt cannot be used properly through Leibniz's law. But what happens when, for example, using the notion of *extension*? Let's now imagine that *x* is the *sensation of sadness* I feel while listening to John Dowland's tune 'Flow my tears' while *y* is the *sound that comes out from the speakers* of my laptop. While *y* is measurable and has a location in space - defined by the waves of pressure that propagate through a given medium - *x* cannot be objectively measurable and does not have a physical, spatial existence. Therefore *x* and *y* do not share the same properties, which means that my feeling is something different from any physical state. Since the *mind* has no extension and it does not share the same attributes with *matter* they cannot be the same thing. However, if we do accept dualism, the problem of how non-physical minds and extended bodies interact will probably never be solved. As Jerry Fodor observes:

The chief drawback of dualism is its failure to account adequately for mental causation. If the mind is nonphysical, it has no position in physical space. How, then, can a mental cause give rise to a behavioural effect that has a position in space? To put it another way, how can the nonphysical give rise to the physical without violating the laws of the conservation of mass, of energy and of momentum? (Fodor, 1981 [1994], p. 25).

Since the body takes up room and it is measurable it must be a separate and distinct substance from the mind (Johnson, 1987). But then, how can an abstract, non-physical, mind cause any kind of change in a body? And how can a body, which is observable, extended, affect something as immaterial as the mind? In his *Meditations on First Philosophy* (1641 [1980]), Descartes investigates the causal relationship between the two. In the sixth meditation, for example, he allows that what causes bodily sensations, like disgust or pleasure, is to be found in the precise structural organization of the nervous system. In particular the nerves transmit to the *pineal gland* the relevant information, which are processed and transformed into the actual feeling. In other words, we could argue that for Descartes the body acts on the mind through the movements of 'spirits' that eventually reach the brain.

However, it remains unclear *how* the body machinery could generate something subjectively unique, and therefore immaterial, like a feeling. As the above-mentioned 'hard problem of consciousness' shows, an explanation of how the relationship between physical and mental entities takes place is still nowadays considered a core issue for the study of human cognition. Cartesian dualism, we can say with Rowlands (2003), suffers from a major problem when trying to *reincorporate* the non-physical mind into the body. How is it possible that a body displays at the same time, two separate substances? Descartes describes this interaction in these terms:

I am present to my body not merely in the way a seaman is present to his ship, but... I am tightly joined and, so to speak, mingled together with it, so much so that I make up one single thing with it (Descartes, 1641 [1980], p. 94).

As Dorothee Legrand noticed, ‘despite his radical dualism, Descartes himself thus fails to disembody the mind of the factual subject and his view rather calls for further investigation of the bodily-self’ (Legrand, 2010, p. 182). Indeed, paraphrasing the British philosopher Gilbert Ryle (1949), the Cartesian mind is not only “a ghost”, but it actually is “a ghost in the machine”.

As Thompson (2007) states, it is purely accidental that *my* mind is embodied in *my* body. At least theoretically, a thinking substance can be genuinely conceived as existing apart from the body, and thus my mind could be easily united with someone else’s body. The *ghost* could inhabit other *machines* without necessarily cause any metaphysical problem. However, we *do* have some problems. The gap still remains unbridgeable, and the *ghost* evoked by the Cartesian stance would eventually transform every *dream* of a fully developed model of the human mind into the *nightmare* of not being able to grasp the ultimate objective reality of such a disembodied substance. Let’s now reconsider the two strong arguments that stem from the Cartesian view of the mind:

- The mind is an immaterial substance, *separated* from the body
- The mind lives inside a body, being a ghost *inside* the skull.

But another strong dichotomy implicitly lies in these assumptions, if we are to take them seriously:

- While human experience is an *internal* property, the world in which our bodies are situated is nothing but an *external* category

The doctrine of dualism encourages a view that radically divides organisms and their environments, thus creating a separation between *internal* and *external*, *subjective* and *objective*. While modern cognitive scientists, psychologists, neuroscientists and philosophers have heavily criticized the dichotomy between mind and matter, the assumption that cognition lies *inside* the head has been silently accepted. Considering the mind-matter dichotomy, Ryle - among others - famously alleged that treating the mind like a non-physical object is a *category mistake*, because things or facts of one kind would be presented as if they belonged to another. In fact, ‘the logical type or category to which a concept belongs is the set of ways in which it is logically legitimate to operate with it’ (1949, p. 8). Probably the most famous of the examples provided by Ryle is the one about a student who comes to Oxford and, while viewing the library and the colleges, asks “but where is the university?”. The category-mistake lies in the visitor’s assumption that a University would fall into the category of ‘buildings’ rather than, say, ‘institutions’. Hence, according to Ryle, Descartes is wrong in attributing the existence of something called ‘mind’ over and above someone’s *behavioural dispositions*.

Ryle's solution is in line with the behaviourist perspective of providing a non-dualistic psychology that focuses on the observable behaviour of living systems (Baum 1994). Hence, on the one hand we have a doctrine like Cartesianism, which tries to explain human behaviour relying on hidden, non-physical causes - an immaterial mind that moves a material body. On the other hand, the observable criteria studied by the behaviourists do not allow occult or hidden mental states. According to a behaviourist, in fact, a mental state of a subject is defined by her behavioural disposition. In other words, I understand that my friend is angry because she shouts at me and not because I inferentially attribute to her the mental state of being angry. All I need to do is to observe her behaviour. 'Behaviourism had allowed no reference to internal states of the organism; explanations of behaviour had to be formulated in terms of sensory stimuli and behavioural conditioning' (Thompson, 2007, p. 4).

This position, however, has been heavily attacked, mainly for being too narrow. As Jerry Fodor points out 'once it has been made clear that the choice between dualism and behaviourism is not exhaustive, a major motivation for the defence of behaviourism is removed: we are not required to be behaviourists simply in order to avoid being dualists' (Fodor, 1968, pp. 58-59). The idea of reconciling mind and matter, to attribute a legitimate role to internal mental states, avoiding *de facto* the rise of behaviourism, thus motivated the emergence of the so-called *cognitive* approach. Proponents of this theoretical framework are firmly committed to the idea that, in order to shed light on the nature of the mind, the only possibility is to unravel the structures *inside* the skull - in terms of subjective mental representations or cytoarchitectonic networks. However, as previously observed, the separation between mind and matter is not the only dichotomy that needs to be overcome.

As the next chapter will show, the attempt to bridge the explanatory gap between mind and matter, as offered by the cognitive approach or by any other form of reductionism, would substantially leave unaltered its dualistic presuppositions. The dualities between *internal* and *external*, *subjective* and *objective*, *inputs* and *outputs* are still controversial issues that perpetuate the Cartesian anxiety. As we will see, the birth and the development of cognitivism, has serious implications for the science of music cognition as it largely inspired the main trends, assumptions and methodologies of musical research.

1.5 - Perpetuating the dichotomies

Cognitivism is the term that labels the anti-behaviouristic movement in cognitive science, which came into being around the half of the last century (Fodor, 1968). While behaviourism tried to shed light on the nature of behaviour by referring to law-like relationships between sensory inputs and behavioural outputs, the cognitive paradigm offered a position that sees 'the mind as a device [i.e. a computer] to manipulate symbols is thus concerned with the formal rules and processes by which the symbols appropriately represent the world' (Thelen, 2000, p. 4).

Having essentially dominated the field for more than forty years, its importance and relevance for current research in music cognition and, more in general, for psychology and cognitive science, has anything but disappeared. As we will see, in fact - despite the emergence of challenging theoretical frameworks - the cognitive perspective is still widespread. Before embarking on this, a critical overview of concepts, methodologies and implications underlining this standpoint will be

provided. According to Mark Rowlands, ‘traditional attempts to study the mind are based on the idea that whatsoever is true of mental processes - perceiving, remembering, thinking, reasoning, and so on - they exist in brains’ (Rowlands, 2010, p. 2). Cognitivism is, in effect, the paradigmatic example of such commitment and Andy Clark (1997, p. 83) strategically summarizes its main assumptions into the following five points:

- Memory is a *retrieval* from a stored symbolic database
- Problem solving is seen as a logical *inference*
- Cognition is *centralized*
- The environment is (just) a *problem* domain
- The body is an *input* device

One by one, these points led to the idea proposed by Jerry Fodor of a *methodological solipsism* (Fodor, 1980a). Mental processes, as he claims, can be understood only on the basis of a subject’s individual mental properties, without any involvement of *external* factors, such as the environment. As Wilson puts it, ‘methodological solipsism in psychology is the view that psychological states should be construed without reference to anything beyond the boundary of the individual who has those states’ (Wilson, 2004, p. 77). The mind-as-a-computer metaphor implies that the key to understand human cognition lies inside the boundaries of the skull. Therefore, what is required to fulfil such a goal is precisely an *isolationist* analysis, a cognitive mapping of an individual mental states, or representations.

But this *dualistic* view of the world ‘presupposes that mental individuation can indeed be explained without recourse to worldly conditions. It presupposes just the existence of thinkers and their thoughts’ (Schantz, 2004, p. 15). Oversimplifying, cognitivism transformed the dualistic asymmetry between mind and matter into a materialistic form, where the brain and the mind form an interactive processing system separated from the rest of the body and from the world. The mind is essentially the program that the brain runs. As Hilary Putnam admits, in fact

If one assumes that the mind is an organ, and one goes on to identify the mind with the brain, it will then become irresistible to (1) think of some of the “representations” as analogous to the classical theorist’s “impressions” (the cerebral computer makes inferences from at least some of the “representations”, the outputs of the perceptual processes, just as the mind makes inferences from impressions, on the classical story), and (2) to think that those “representations” are linked to objects in the organism’s environment only causally, and not cognitively (just as impressions were linked to ‘external objects’ only causally, and not cognitively) (Putnam, 1981, pp. 9-10).

If the mind were a sort of computer, then the brain substrates would represent the “hardware”. On the other hand, mental processes would constitute the ‘programs’. The difference between “software” and “hardware” is well exemplified by the research programs that are supposed to shed

light on their nature and functional properties. While *cognitive psychology* is devoted to investigating and identifying the cognitive entities, *cognitive neuroscience* studies how these work and where they are implemented in the brain, with the ultimate aim to provide an extensive and precise cytoarchitectonic mapping of neural networks. As Noë noticed, in fact, ‘Psychology is interested in what the brain does, but at higher levels of abstractions than that of neuroscience’ (Noë, 2004, p. 25). The cognitivist perspective opened a new gap between computational cognition and subjective mental phenomena (Thompson, 2007, p. 6), that is, as Jackendoff observes, a *mind-mind* problem:

The upshot is that psychology has not two domains to worry about, the brain and the mind, but three: the brain, the computational mind, and the phenomenological mind. Consequently, Descartes’ formulation of the mind-body problem is split into two separate issues. “The phenomenological mind problem” [...] is, How can a brain have experience? “The computational mind problem” is, How can a brain accomplish reasoning? In addition we have the “mind-mind problem”, namely, What is the relation between computational states and experience? (Jackendoff, 1987, p. 20, quoted in Varela et al., 1991, p. 52).

The mind-mind problem, as clearly emerges², is a variant of ‘the hard problem of consciousness’ that we previously encountered. This implies that cognitivism, far from concluding the discussion on the nature of mind and cognition, expands the gap between mind and body into a new series of problems that call for further solution. The legacy that links Cartesianism to cognitivism lies in the fact that the strong dichotomies created by the dualistic stance are perpetuated by cognitivist materialism. Several solutions have been posited to overcome these difficulties (e.g. Dreyfus, 1995).

One of the most famous critiques of the computational mind is probably represented by the controversial (see Dennett, 1980, 1987; Block, 1980; Fodor, 1980b; Crane, 1991; Cole 1991) mental experiment of “the Chinese room” offered and revisited several times by John Searle (1980). Searle wants to demonstrate that a computer cannot *think* or that, in other terms, symbolic manipulation deals only with syntax and does not have anything to do with *understanding* or with semantics. His attack was mainly against the proponents of the so-called *strong Artificial Intelligence*, who claimed that a coherently programmed computer is not only a simulation, or a model, of the mind, but can actually be a mind. Therefore, the argument also goes against the mind-as-a-computer metaphor proposed by the cognitive standpoint. If we assume, along with the cognitive framework, that the mind is software and that the brain is hardware, then it would certainly be true that a machine, or an actual computer, can have mental states as long as it runs the right kind of software. The Chinese Room argument goes as follows:

Imagine a native English speaker who knows no Chinese locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating

² This analogy has been proposed by Evan Thompson (2007, p. 7).

the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in the room to pass the Turing Test for understanding Chinese but he does not understand a word of Chinese (Searle, 1999, p. 115).

The point Searle tries to stress, hence, is that *without understanding*, it would be impossible to posit that the machine is *thinking* - or having a mind. As an example, let's consider the software *Finale*[®] or *Sibelius*[®]: although they manage to process many information about music, would it be possible to say that these programs *know* music? Searle's reply would be that symbolic manipulation systems are completely unable to generate any kind of mental phenomena, including our theoretical knowledge of music. I fundamentally agree, here. Acknowledging the mistake of having two parallel octaves while composing a fugue in the style of J.S. Bach does not have anything to do with the composer's *understanding* of musical styles. I will return to the problem of *musical understanding* in Chapter 3. For now, I briefly want to focus on the solution adopted by Searle.

While rejecting any form of mental-physical dualism, as well as the computational theory of the mind proposed by the cognitive standpoint, Searle embraces what he labels as *biological naturalism* (Searle, 1992; 2004). Mind and brain are essentially the same thing. Therefore, it comes as no surprise that its methodological paradigm essentially demands Neural Correlates of Consciousness. There are two ways to achieve such a goal: (i) to look for *building blocks* in the human brain that would underlie a particular, subjective, conscious state - like the feeling evoked by the sound of middle C - or (ii) to look for a whole conscious field of *unified subjectivity* (Searle, 2004, p. 116). The first strategy - that Searle himself ultimately refuses (Searle, 2000a, p. 572) - underlies a one-to-one *identity relationship* between mental and neural states. This is what can be called the "matching content doctrine" (Noë & Thompson, 2004a; 2004b). To provide an example of such a view, let's consider David Chalmers' position (Chalmers, 2000). He assumes that neural and experiential content are exactly the same thing. By doing that, however he exposes himself to a *category mistake*. As Thompson notes, *experience* is intentional (world-presenting), holistic (constituted by interrelated perceptions, intentions, emotions and actions) and intransitively self-aware (has a non-reflective subjective character). Neural content as standardly described has none of these features (Thompson, 2007, p. 350). The second approach embraced by Searle (unified subjectivity) maintains that any given conscious state is not a mere aggregation of building blocks, but rather a modulation of already existing conscious states. For example, the feeling of middle C does not depend on the intermodal connections between building blocks but, rather, on the modification of a pre-existing *conscious field* in the background.

One way or another, Searle's solution is committed to the idea that the relationship between mind and body 'has a rather simple solution. Here it is: conscious states are caused by lower-level neurobiological processes in the brain and are themselves higher-level features of the brain' (Searle, 2002, p.9). As Chemero puts it: 'the most common strategy employed here is to argue that cognitive science is reducible to cognitive neuroscience' (Chemero, 2009, p. 167). In other words, conscious experience is completely reducible to the brain, and all we need to do in order to obtain an advance

for the “hard problem of consciousness” is, literally, to look inside the boundaries of the skull. Let’s consider the following passage:

The brain is all we have for the purpose of representing the world to ourselves and everything we can use must be inside the brain. Each of our beliefs must be possible for a being who is a brain in a vat because each of us is precisely a brain in a vat. The vat is the skull and the ‘messages’ coming in are coming in by way of impacts on the nervous system (Searle, 1983, p. 230).

The expression “brain-in-a-vat” refers to the famous thought-experiment proposed by Hilary Putnam (1981), which basically mirrors the traditional skeptical argument used by Descartes about the impossibility of telling whether there is an evil entity that feeds us with false information about the world through our senses. In other words, a brain could potentially be fooled into anything if only it were fed with appropriate stimuli. Cosmelli and Thompson put it in this way:

Suppose that a team of neurosurgeons and bioengineers were able to remove your brain from your body, suspend it in a life-sustaining vat of liquid nutrients, and connect its neurons and nerve terminals by wires to a supercomputer that would stimulate it with electrical impulses exactly like those it normally receives when embodied. According to this brain-in-a-vat thought experiment, your envatted brain and your embodied brain would have subjectively indistinguishable mental lives. For all you know - so one argument goes - you could be such a brain in a vat right now (Cosmelli & Thompson, 2009, p. 362).

Evan Thompson (2007) proposes a *null hypothesis* for this thought experiment. His strategy consists in individuating a particular set of specific conditions that the vat would have. As he points out, in fact

The vat setup would have to be capable of (i) keeping the brain alive and up and running, (ii) duplicating all exogenous stimulation and (iii) compensating in exactly the right way for all endogenous (and self-organizing) activity. Such a setup would almost certainly have to duplicate many of the chemical, biochemical, and sensorimotor properties of the body, probably even the body’s sensorimotor coupling with the world. In other words, the null hypothesis is that any vat that meets the requirements of this scenario will be a surrogate body (Thompson, 2007, pp. 240-241).

I think that two main conclusions can be drawn from this brief discussion. On the one hand, the strategy of looking for specific mental or brain states in order to explain cognition appears to be problematic for several reasons. If we assume that the mind is a computer then we cannot grasp the

causal relationship that connects symbolic manipulation to reasoning, as the Chinese room experiment shows. On the other hand, a reductionist approach only seems to perpetuate such a problematic and ungraspable relationship. As many authors point out, indeed, even if we know what causes or correlates to a specific conscious state - like the sensation of middle C - we could still dispute whether experiences are nothing *more* than their causes or their correlates (i.e. Hutto & Myin 2013). But there is, as I said, another conclusion. The role of the *body* clearly represents a challenge for any form of *reductionism* or *methodological solipsism*. The consideration that the brain *is not the sole cause of experience* - as it emerges from the discussion about the brain-in-a-vat - calls for a significant paradigm shift from the traditional views of the mind. The *envatted brain* experiment shows that a brain cannot be seen as *identical* to the mind because experience emerges 'not from neural activity alone but from the interaction between the brain and the device' (Rockwell, 2005, p. 68); a *device* that inevitably will be a surrogate of the body and thus, immersed in the world.

As I briefly mentioned in the introduction, the new theoretical frameworks developed in the last few years by philosophers, psychologists, neuroscientists and cognitive scientists dramatically reconsider the role of the *body* and of the *environment* for cognition, and radically challenge the view that mental representations are the only possible tool for reasoning and understanding the world. The *brain-body-world nexus*, as Rockwell (2005) puts it, implies not only that the mind *is not in* the brain, but also that the mind *is not the* brain. The reconceptualization of categories like *perception*, *emotion*, *action* and *cognition*, provided by these new standpoints, would enormously contribute to the process of developing a new perspective on musical experience. As I see it, indeed, most of the current research in music cognition relies on computational or reductionist accounts that, both theoretically and empirically, cannot provide a biologically and phenomenologically plausible account of musical experience. The next chapter is devoted to convincing you of this.

Chapter 2

Diabolus in Musica³

The reason for the failure of rational culture, as we said, lies not in the essence of rationalism itself but solely in its being rendered superficial, in its entanglement in naturalism and objectivism

- Edmund Husserl, *The Vienna Lecture*

Cognitivism, like life and pasta, comes in a bewildering variety of forms

- Andy Clark, *Microcognition*

2.1 - Beyond the modularity of the musical brain

The devil is in the details, they say. This chapter is dedicated to exploring some assumptions at the basis of traditional research on human musicality. In particular I want to show that the explicit or implicit use of the cognitive standpoint cannot allow a coherent model of music cognition. Traditionally, Noam Chomsky (1957; 1966) and Jerry Fodor (1983) are considered the prototypical proponents of the cognitivist framework, as they explicitly supported the dichotomies between *inner* mental states and *external* observable behaviours and between *modular*, fixed brain functions and unspecific, *dynamical* learning mechanisms (Brincker, 2010, p. 5).

For *modularism*, I refer to the theory proposed by Jerry Fodor (1983) about a cognitive architecture set out in specialized vertical structures - *modules* - underlying the mental ability to transform the input into representations afterwards offered to central areas for more complex elaborations (see Calabretta et al., 2003). As Calabretta and Parisi put it ‘modular systems can be defined as systems made up of structurally and/or functionally distinct parts. While non-modular systems are internally homogeneous, modular systems are segmented into modules, i.e., portions of a system having a structure and/or function different from the structure or function of other portions of the system’ (Calabretta & Parisi, 2005, p.3). According to Prinz (2006a), Fodor defines brain modules by their special properties. ‘He says that modular systems are: (1) Localized: modules are realized in dedicated neural architecture; (2) Subject to characteristic breakdowns: modules can be

³ This chapter contains material previously published as **Schiavio, A.** (2012). Constituting the musical object. A neurophenomenological perspective on musical research. In *Teorema*, 31(3), pp. 63-80. Moreover, it contains material from the following article in preparation: **Schiavio, A., & Gerson, S.** (in preparation). An enactive perspective on infants’ musicality.

selectively impaired; (3) Mandatory: modules operate in an automatic way; (4) Fast: modules generate outputs quickly; (5) Shallow: modules have relatively simple outputs (e.g., not judgments); (6) Ontogenetically determined: modules develop in a characteristic pace and sequence; (7) Domain specific: modules cope with a restricted class of inputs; (8) Inaccessible: higher levels of processing have limited access to the representations within a module; (9) Informationally encapsulated: modules cannot be guided by information at higher levels of processing' (Prinz, 2006a, pp. 22-23).

In his analysis, Prinz offers a series of valid arguments against modularity. However, my purpose here is not to consider each property one by one. Rather, I aim to show that the radicalization of such a model, in particular with regard to points 6 and 7, would substantially compromise any plausible account of music experience. In general, William Uttal (2003), arguing that an efficacious taxonomy of all mental processes has yet to be developed, rejects the idea that higher cognitive processes can be localised to particular regions (or networks) in the brain referring to the studies of localization as a modern variant of phrenology. Moreover, any sort of reductionism cannot account for a phenomenologically relevant theory of experience that takes seriously the intimate relationship between brain and body. However, the modular perspective did influence many scholars involved in music-related issues from the early eighties to the following decades (i.e. Peretz & Morais, 1989; Piccirilli et al., 2000) often through a comparison between music and language (see McMullen & Saffran, 2004, for a review).

The classic method of reasoning can be described as follows: (i) the musical signal is an external stream of information that comes from the environment; (ii) it is processed by some specific brain mechanisms, and finally (iii) it results in a mental representation. The inferential fragmentation of this process via perception, neural integration and representation is in league with the traditional view that considers someone's musical experience as an inner, computationally implemented, event. This assumption often led researchers towards a model of the musical mind that is, at least, disputable. The structural organization that *objectively* constitutes the physical signal (pitches, rhythm, dynamics, harmony and so forth) - which is reflected by the brain's modular architecture - is *subjectively* experienced *via* representations. In other words, the assumption underlining this commitment is that *music is not experienced as a whole* but, rather, through the inferential integration of different domains such as pitch, rhythm, harmony, as processed by their relevant modules. Daniel Levitin (2006) for example, insists that we can identify some categories (once again, pitch, timbre, rhythm and so forth) that can be individually varied without affecting other musical dimensions. In his view, indeed, the brain extracts from the musical stimuli each of these attributes autonomously. Several studies, however, make doubtful that (for instance) *melodic interval perception* is codified *per se*, showing that estimates of interval size are actually affected by temporal structure (Russo & Thompson, 2005a), as well as register, melodic direction (Russo & Thompson, 2005b) and listener expertise, giving rise under certain timbre manipulations to an interval size illusion.

According to Ian Cross, moreover, 'what we know of music in neurobiological and neuroscientific terms is constrained by a conception of music that is narrowly shaped by historical and cultural notions of what constitutes "music"' (Cross, 2010, p. 2). As I see it, this problem concerns not only the classical theoretical problems of the definition of music (see paragraph 2.5 for discussion), but also refers to the modular perspective. Proponents of this argument cite specific music-related deficits in patients that are consistent with a modular perspective of the mind. In a well-known paper, for example, Isabelle Peretz (1993) describes the case of patient G.L., whose ability to elaborate sounds in a *tonal* context was compromised after bilateral temporal lobe damage

due to strokes. Interestingly, G.L. did not display any other problem in processing temporal structures or melodic contours: the only deficit was in discriminating the coherence of a given pitch with the harmonic context. G.L. was diagnosed with *amusia without aphasia* as s/he scored in the normal range of standard aphasia tests. However, getting back to Cross' consideration, we should ask whether this dysfunction can be considered *biologically musical*. In fact, tonal music is only a small part of the complex phenomenon of music, being a Western feature that emerged in the Fifteenth Century. Is it then possible to posit species-specific brain architectures phylogenetically developed to process distinct aspects related to peculiar, Western-based, musical organizations? If the neural basis of musicality were a pre-wired biological system, this would imply that musical knowledge is pre-determined, being encapsulated in domain-specific brain modules from birth. But, as Honing and Ploeger (2012) pointed out, 'while there are several aspects of musicality that emerge spontaneously, early in life, and with minimal exposure to music, and as such suggesting heritability (cf. Lewontin, 1998), none of them need to be specific to music' (ibid., p. 5). Moreover, a reflection on the processes underlying human development shows that developmental plasticity leads to knock-on effects in different domains (Karmiloff-Smith & Thomas, 2003, p. 648).

Therefore, the localisation of a specific musical rule in one area of the brain cannot be accepted as evidence for the modular view of music since a specific function in the brain may result from many different developmental processes (advances in sensorimotor abilities, language development, etc.) not involving 'genetic specification of epistemological content' (Tomasello, 1999, p. 203). A modular interpretation of the mind/brain, moreover, seems to be hardly sustainable considering that the history of *Homo Sapiens* 'is simply not sufficient, under any plausible scenario, for genetic variation and natural selection to have created many different and independent human cognitive modules' (ibid., p. 55).

2.2 - Statistical properties, expectations and mental representations

The modular account of the musical mind represents an important outcome of the cognitive neuroscience of music. The focus on specific brain mechanisms underlying our ability to be engaged with a musical stimulus has been largely accepted during the last decades, leading to a research program whose ultimate aim is to realize a map of what is going on in the brain during any sort of musical experience. As emerged, however, this methodology reflects a leftover from the Cartesian stance, though reconsidered on the basis of a materialistic commitment: if the brain's physiological architecture is the sole source of relevant musical information, then (i) the external world - the music that comes from the environment - remains an independent category from the subjective experience that occurs in someone's brain, and (ii) it still remains unclear how exactly neural tissues or mental computations would substantially enable us to have *that* particular musical experience. Once again the dream of achieving an objective perspective on the nature of musical experience meets a ghostly nightmare.

The Cartesian anxiety is, in fact not cured: even if particular neural processes present a clear correlation with specific musical structures we still miss a complete picture of the phenomenon. Correlation, after all, is not causation. The radical form of *internalistic materialism* portrayed by the modular approach to human musicality precludes every consideration about what happens *outside* the skull because of the commitment to the idea that the mind is imprisoned inside the head. On a

related note, if we are to take seriously the claim that mind and brain form a unique information-processing mechanism we are still unable to understand how exactly a subjective feeling can be generated. No matter how deep our understanding of neurons and their synaptic connections might be: the gap between these physical properties and the subjective feeling that they generate is still open.

I see David Huron's work *Sweet Anticipation* (2006) as a complementary attempt to solve this puzzle. Despite the noble intentions, however, Huron's work seems to be committed not only to the materialistic constraint that sees music as having *objective* properties which can be autonomously coded by the human brain, but also to the idea that categories like musical expectations are basically "mental *or* corporeal belief" (Huron, 2006, p. 41) splitting *de facto* body and mind/brain into two distinct systems. This standpoint in fact - as any internalist approach devoted to bridging the explanatory gap - perpetuates the *Cartesian anxiety*. Huron's story is that statistical properties of tonal music, to which listeners have been familiarised with by means of exposure, would elicit particular *qualia* (or subjective feelings) via relevant expectations. As it appears from the book, his argumentation can be summarized as follows:

- i) Tonal music exhibits statistical properties on different levels (pitch, scale degrees etc.)
 - ii) Listeners display sensitivity to the probabilities of different musical events and patterns
 - iii) Listeners use these probabilities to generate predictive expectations about the future
 - iv) Expectations imply mental representations
 - v) Mental representations are physical brain states
 - vi) The violation or confirmation of such expectations generates specific music-related qualia
- Therefore → Qualia derive from brain states elicited by expectations and representations of a given musical parameter.

The argument, however, seems to be fragile for several reasons. Firstly, the premises (iv) and (v) do not really persuade me. They do not logically follow from the other premise (iii), and thus reflect an aprioristic assumption about the involvement of mental phenomena. How exactly, do expectations imply mental representations? And why? Huron adopts a sort of biological naturalism, which consists in the idea of an identity between conscious experience and neural substrates. Consider the following passage:

Auditory images are not organized in the brain like phonograph recordings. Instead, brains *interpret, distill, and represent* sounds. [...] Expectations imply some sort of mental representation. The *what, when, and where* of expectation exist as mental codes. These mental codes are not disembodied abstractions. They exist as real biological patterns that have taken up residence somewhere inside people's heads (Huron 2006, p. 101).

When Huron says that "mental codes are not disembodied abstractions", he does not mean that they are constituted by wider processes distributed over the body in its relationship with the

environment. This would be, in fact, a traditional embodied claim (i.e. Shapiro, 2011). Huron rather claims that they are “embodied” within the brain. But how this assumption would justify the involvement of mental representations? After all, for the sake of the argument we could maintain that the neural activity in the brain directly generates the relevant subjective quality without any inferential, representative, subordination. My point is that Huron tends to overintellectualize a process that potentially could not involve high-level processing. I do not see any contradiction, for example, in postulating *expectations* without *explicit representations*. The data from the following experiment might provide a useful insight into this issue. Sebylle Herholtz and colleagues (2008) asked two groups of musicians and non-musicians to listen to and familiarise themselves with a series of thirteen notes, all with the same duration.



Fig. 1
One of the melodies used in the experiment

After the familiarisation phase, they manipulated the same stimuli by cutting the five notes before the ending tone. Then they asked the participants to imagine the missing part of the melody.

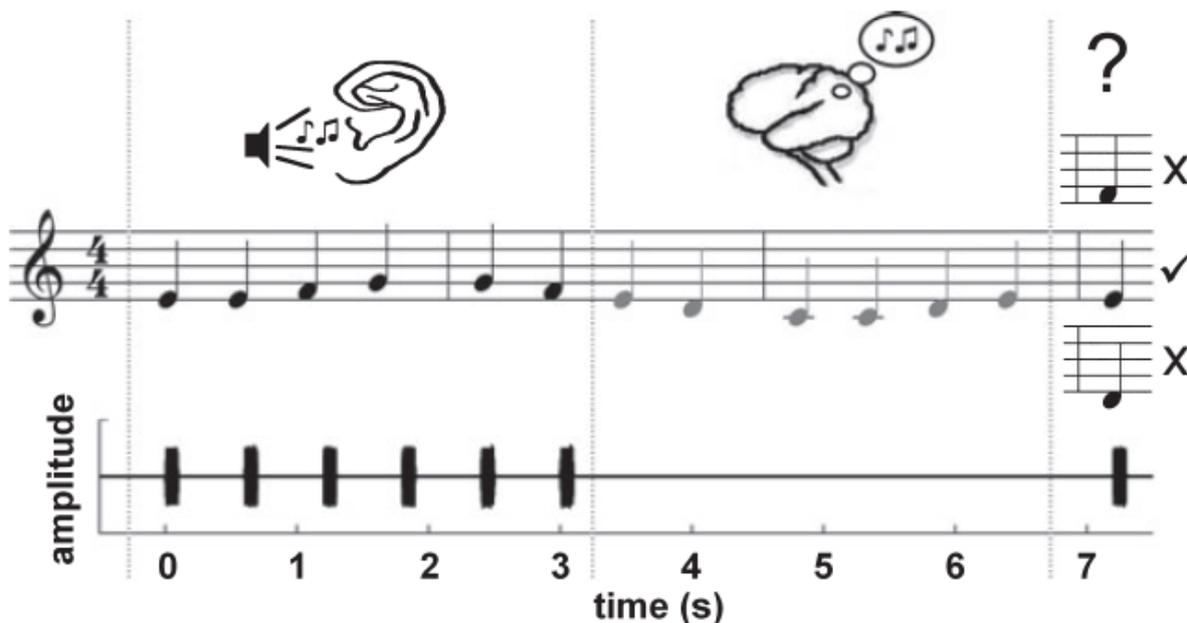


Fig. 2
Imaginary task (from Herholtz et al., 2008). Reprinted with permission of the publisher. License Number: 3333561359397.

After different listening trials, however, the researchers substituted the final note with a *different pitch*. The idea of the experiment was, therefore, to generate an unexpected event and to see what happens in the brain of the listeners, thanks to the employment of MEG techniques. The researchers found that in the musicians' group, this experience provoked pre-attentive responses, (measured through the imagery mismatch negativity or iMMN). The time between the unexpected auditory event and the recorded brain signal is *too short* (175 ms) for the subjects to recruit contentful representations, suggesting that expert musicians process music without any explicit representational activity. From these results, it emerges how the familiarisation with a musical pattern can improve the *pre-attentive* processing of the given auditory feedback (see also Koelsch et al., 1999; Sittiprapaporn et al., 2003) before any kind of cognitive subordination. With regard to this point, someone could argue that Huron's view involves different (high-level) kinds of expectations, therefore requiring an adequate cognitive load (in terms of representations). However, it seems quite clear that Huron bases his hypothesis on the parallelism between qualia and expectations, in which a single difference in pitch would generate different qualia. If some qualia can emerge without explicit mental representations (without semantic content) it would be contradictory to expect different qualia to exhibit representational (or semantic) content. Otherwise the latter would fall into another category, rather than qualia, displaying additional properties that would differ from their proper characterisation. In other words, if Huron were right, even the violation of a single pitch's expectation would give rise to a mental representation above other sensorimotorically grounded modalities.

In this sense, the data from Herholtz and co-workers' experiment cannot fit into the computational model of musical understanding, where perceptual experience is coded through top-down and bottom-up processes, maintaining *mental representation* as the outcoming category that stems from their integration. On the contrary, Herholtz and colleagues show that the violation of expectations of a given melodic fragment has little to do with disembodied phenomena like mental objects. The principal and primal source of understanding, as will be shown later, is the body of the subjects, in its dynamic interplay with the brain and the intended musical object,

The idea of avoiding representational processes has been explored in the last few years by different theoretical frameworks. A strong anti-representational account for cognition is for example the *dynamic system theory* (Beer, 1990; Thelen & Smith, 1994). Broadly speaking, this approach avoids any recruitment of centralized and isolated representational mechanisms. Rather, proponents of this standpoint try to explain how a living system can experience the world in terms of its embodied actions and its embodied situatedness. Instead of computational architectures underlying meaningless symbols manipulation, they assume that multiple systems implement categories such as action and perception 'where each system is capable of residing in one of infinitely many continuous states' (Barsalou, 2008, p. 621). This framework, therefore, tries to avoid the notion of mental representation by considering the environment as a place of salience, with which the animal interacts through sensorimotor, affective and meaningful patterns. In fact, 'over learning, states of these systems become coupled to reflect patterns of interaction with each other and with the environment effective in achieving goals' (ibid.). But reconsidering the role of the interaction between embodied agents and the world in non-representationalist terms would also mean that *expectations* - being forms of interaction with the world - should somehow display a different format. If we identify expectations with sensorimotor patterns of interaction with the world, then it seems that some of the processes that the brain is doing, cannot be decoupled from the environment. Hence, as Noë puts it, perceptual experience becomes a 'temporally extended

process of exploration of the environment on the part of an embodied animal' (Noë, 2000, p. 128). To provide an example, let me introduce the following vignette: I am walking through Rome with my girlfriend when we suddenly hear some music coming from the left side of Piazza Navona. After we hear the sound we turn our heads towards the side of the square where a violinist is improvising a variation on a Caprice by Paganini. As Noë's standpoint implies, my understanding of where the violinist is situated is not a *conceptual* experience (I do not need to elaborate a specific mental state to understand where the violinist is playing). Rather, my experience of the violinist standing of the left side of Piazza Navona consists in my sensorimotor expertise of how my meaningful movements through the environment would allow me to reach her there.

It takes no thought or intellectual skill to know that to bring the item off to the left better into view, you must turn your head to the left [...] [or] when you hear a sound as being on the left you don't need to think about which way to turn your head in order to orientate toward the sound. [...] You do need to think about how to maneuver a couch to squeeze it through a small passage. But you do not need, in the same way, to think about how to maneuver your body to squeeze it through the doorway. Just perceiving the doorway as having certain spatial qualities is perceiving it as enabling, requiring, or permitting certain kinds of movements with respect to it (Noë, 2004, p. 89).

To use the terminology employed by Huron, my *where-expectation* does not require any sort of conceptual content. Recent studies on cross-modal mapping of musical properties, moreover, show that pitch-space can correspond cross-modally to physical space (Lidji et al., 2007; Eitan & Timmers, 2010) thus revealing a level of abstraction that is not entirely decoupled from the embodied situatedness in the physical world. From this standpoint, any isolationist analysis of single sounds' attributes would ultimately result in the inability to grasp the complexity of the relationship that links embodied agents and musical objects. Huron, as I see it, perpetuates the Cartesian anxiety by neglecting the role of processes that could be easily considered as *sensorimotorically grounded*.

As Dan Hutto recently noted, with regard to the cognitive approach: 'this intellectualist way of understanding the basic nature of minds taps into a long tradition stretching back at least as far as Plato; it was revived by Descartes in the modern era, and regained ascendancy, most recently, through the work of Chomsky during the most recent cognitive revolution' (Hutto, 2013, p. 252). Given these considerations, I take the liberty to partially modify Huron's story, by getting rid of premises (iv) and (v), as well as part of the conclusion, without substantially compromising the general scheme:

- i) Tonal music exhibits statistical properties on different levels (pitch, scale degrees etc.)
- ii) Listeners display sensitivity to the probabilities of different musical events and patterns
- iii) Listeners use these probabilities to generate predictive expectations about the future
- ~~iv) Expectations imply mental representations~~
- ~~v) Mental representations are physical brain states~~
- vi) The violation or confirmation of such expectations generates specific music-related qualia

Therefore → Qualia derive from brain states elicited by expectations ~~and representations~~ of a given musical parameter.

By avoiding fully blown mental representation, Huron's position still maintains the identity between qualia and brain states. This standpoint presents several problems and it is based on questionable assumptions. The next section offers a critique of such a position.

2.3 - Musical qualia

For *qualia*, Huron refers to the term firstly employed by Clarence Irving Lewis in his 'Mind and World Order' (1929), where he writes:

There are recognizable qualitative characters of the given, which may be repeated in different experiences, and are thus a sort of universals; I call these "qualia". But although such qualia are universals, in the sense of being recognized from one to another experience, they must be distinguished from the properties of objects. Confusion of these two is characteristic of many historical conceptions, as well as of current essence-theories. The quale is directly intuited, given, and is not the subject of any possible error because it is purely subjective (Lewis, 1929, p. 121).

In other words, a *quale*, is the subjective experience associated with a given sensory event, a certain feature of the bodily sensation (Jackson, 1982). The theory of qualia, therefore, holds that 'for something to look red to someone is for it to give rise to an experience with a certain qualitative or sensational property. Its looking red consists in the fact that it gives rise to that qualitative state in a person' (Noë, 2004, p. 133). The sensation of a middle C, or the feeling of closure evoked by a cadence in tonal music seem to be very familiar but still cannot be found outside our mind, as properties of the objects.



Fig. 3
Traditional tonal cadential pattern

If we ask a scientist about the sense of tension elicited by the cadential chords in the traditional harmonic progression I (in second inversion) V, I, her reply would be in terms of sound waves propagated through the medium of air. Sound travels through disturbances in given medium (air for example) pressure, being constituted by waves of different amplitude and frequency.

But *where* is the tension provoked by this passage if it is not present in the outside world? Where is the *sensation* that makes me feel the need to solve this melodic pattern to the tonic? I believe, for reasons that will be clear at the end of this chapter (see in particular the next section), those questions to be *wrong* questions. As we will see, they only perpetuate an aprioristic Cartesian presupposition, and hence can be classified as *category mistakes*. However most philosophers and neuroscientists do take these questions seriously and, like Huron, suggest that the answer has to be found inside the brain. In particular, a standard answer would be as follows:

- i) A musical stimulus reaches the sensory input
 - ii) The sensory input sends the signal to the auditory cortex
 - iii) A particular *neural state* would be set up
- Therefore → This neural state corresponds to my subjective feelings (*qualia*)

As previously discussed, between (iii) and the conclusion of the argument lies the *explanatory gap* of the ‘hard problem of consciousness’. The two solutions usually adopted to explain how (iii) and (iv) relate are *dualism* (qualia are immaterial entities generated by a physical substance) or *reductionism* (qualia are to be identified with a particular brain state). In both cases, however, there is a problem. The first approach will always face the ontological issue of *causation* while the second usually sees the brain as a *machine* isolated from the rest of the nervous system. Given his commitment to identifying mental/brain states to qualia, Huron’s general discussion implies that the musical mind is functionally independent from the *dynamic interactions* between brain, body and world. Many arguments, however, have tried to reject the identity theory between neural states and qualia. In particular, it has been posited that two functionally identical living systems might display complete different *qualia*.

The *inverted spectrum* argument originally conceived by John Locke (1689), proposes a scenario where two different persons (name them John and Mike) are qualitatively inverted with respect to the colours they experience in the world - although their brain and visual system work in the very same way. Both John and Mike refer to green apples as “green”. When John sees green apples, his experience is about what everyone would consider as “green”. However, when Mike sees green apples he has, for unknown reasons, a reddish experience. What follows then, is that it is possible to imagine that *qualia* (the sensation of the greenishness of green) have a different relationship with a given brain state. We can in fact *conceive* a change in the property of how a given object looks to us without any physical basis (Tye, 1995). As Noë puts it, the inverted spectrum advocates a position where ‘two individuals who are identical in all behavioural dispositions (including their sensorimotor skills and discriminatory capacities) could differ in what it is like for them to experience something red looking’ (Noë, 2004, p. 124).

Another argument that would undermine the identity theory endorsed by Huron is the *Zombie argument* developed, among others, by David Chalmers (1996). This mental experiment relies on a

line of reasoning that has been previously employed by Descartes⁴ (1637 [1960]) and assumes the logical possibility of an imaginary world dominated by creatures exactly like normal human beings, who, however, lack consciousness. If, in this imaginary world, bodies with no consciousness (zombies) were physiologically indistinguishable from human beings, then, even its logical conceivability would be a sound argument to posit that qualia are non-physical entities. To have a clearer picture of the argument we can use the following scheme:

- i) I can conceive a world of zombies
 - ii) If I can conceive something, that *something* is possible
 - iii) Zombies are possible
- Therefore → mental states can be distinct from the body's physical state

The rise of physicalist models to explain the nature of human cognition made clear that everything that happens in the world must have a physical causation. As discussed before, however, it is quite hard to make conscious processes fit into this reductionist view. While materialists would claim that mental states are basically a matter of neural processes, scholars who refuse this view basically collapse into dualist-like assumptions, being committed to the idea of non-physical categories. David Chalmers falls in between these two categorizations, considering that his approach could be defined as a “naturalistic dualism⁵” (Chalmers, 1996). The main problem with the zombie argument, as well as the inverted spectrum argument, however, lies in the second premise. Does conceivability really entail possibility? Daniel Dennett, for example, fiercely rejects the assumption that conceivability is sufficient for logical possibility and then offers several arguments against the existence of qualia themselves (1988; 1991). I think it is worth drawing a brief comparison between Dennett and Huron, here. Dennett's strategy consists firstly in individuating four main properties defining qualia (1988, p. 385). In his view, qualia are:

- Ineffable - it is impossible to communicate a *quale* to other persons. The only way to apprehend *qualia* is by direct experience.
- Intrinsic - they are inside the organism and therefore non-relational.
- Private - they are subjective, so any interpersonal comparison would be impossible.
- Directly or immediately apprehensible in consciousness - there is no intermediate level or representational inference.

⁴ For Descartes non-human creatures like animals are essentially *automata* whose behaviour is ruled only by the principle of mechanics. Given that animals do not display language creativity and appropriate non-verbal behaviour - which are two main typical human behaviours - he concluded that no machine could realistically behave like a human being. What would, then, clearly differentiate human from non-human creatures? As previously discussed, what it is required is precisely a non-physical category like the human mind. Apparently, from this conclusion, it follows that an imaginary world populated by bodies lacking thinking substances it is not conceivable.

⁵ Chalmers admits that the brain generates mental states, but at the same time he maintains that mental states are ontologically separated from brain activity - which is the reason why he talks about *explanatory gap*.

We can immediately see a radical incompatibility between the above-mentioned properties and the way Huron treats *qualia*. Let's for example consider the following quote, where the latter illustrates his *qualia survey*:

I asked ten experienced Western-enculturated musicians to describe the different scale degrees for the major key. Each musician received the following instruction: For each of the following scale degrees describe as best you can the distinctive quality or character of that tone. Describe how the tone makes you feel in as much detail as possible. Imagine the tones for the major key only. Please think of pitches rather than chords. Asking musicians to perform a task like this makes a lot of assumptions. The task assumes that isolated tones are able to evoke *qualia* independent of some musical context. It assumes that musicians can introspect about such matters without actually hearing the sounds. The task assumes that language can be used to express the ineffable qualities people experience. The survey is also limited to the major scale with no consideration of the various minor scales (Huron, 2006, p. 144).

Huron aprioristically assumes that qualia are *independent* from the context, they can be rather easily *expressed through language* and that *introspection* could be enough to provoke the same qualia that the subjects would experience while hearing actual sounds. I am not saying that Huron is wrong in assuming that qualia display those properties - after all they only contradict Dennett's position. But the fact that all of these assumptions are not rigorously motivated raises several problems for the architecture of the whole argument. Huron, furthermore, agrees with Dennett, about the second point. Qualia, for both of them, are intrinsic to the animal. To be fair, Huron also recognizes the role of the socio-cultural environment in determining the results of the computations (Huron 2006, p. 109) that eventually generate *qualia* but, as this computation implies, he holds that qualia are not *directly* experienced.

But let's go back to Dennett's argument. After having individuated these properties, he asserts that although *qualia* stand for *the intuitive way in which things seem to us*, conscious experience has no special properties to make sense of them. To defend this claim he provides fifteen arguments, or mental experiments, named "intuition pumps" in order to show that, after all, the word "qualia" is only an excess term in the vocabulary. If we were able to associate the term itself to some referents (that we already know precisely), then there would be no necessity to postulate the existence of qualia. Our experience of the colour green will be described as being "green" and using an entity like a *quale* to make sense of it only generates an unnecessary contradiction. If indeed qualia define with another word something already well understood, like the colour "green" then there is no need to postulate their existence. There is no mechanism in the brain that would justify this claim. He therefore concludes:

So when we look one last time at our original characterisation of qualia, as ineffable, intrinsic, private, directly apprehensible properties of experience, we find that there is nothing to fill the bill. [...] So contrary to what seems obvious at first blush, there simply are no qualia at all (Dennett, 1988 p. 409).

Thus, Dennett ends up rejecting the notion of *qualia*, denying that there is a hard problem of consciousness at all (see also O'Regan & Noë, 2001b; Hutto, 2006, Cohen & Dennett, 2011). Let's briefly recapitulate, now. Huron posits that qualia exist, being intrinsic, representational and comparable among different people because easily communicable through language. Dennett, on the other hand, claims that the notion of qualia itself is fundamentally confusing, and therefore the *hard problem of consciousness, as well as qualia*, are simply non-problems. Who is right then?

2.4 - Towards an enactive solution

Besides the conclusion and some important differences in defining *qualia*, both Dennett and Huron share a common view. Simply put, they assume that everything that is related to the mind - *qualia* or not *qualia*, *hard problem* or not *hard problem* - must occur inside the boundaries of the skull. In fact, as previously pointed out, the only property of qualia they agree on is their being *intrinsic* - located inside the living system. I think that this is an assumption as well, *pace* Dennett. Unlike Huron, Dennett denies the existences of entities like *qualia* but - like Huron - does not deny the realm of subjectivity. And they both *reduce* it to brain mechanisms.

According to Evan Thompson, however, 'the standard formulation of the hard problem is embedded in the Cartesian framework of the "mental" versus the "physical" [...]. The guiding issue is no longer the contrived one of whether a subjectivist concept of consciousness can be derived from an objectivist concept of the body. Rather, the guiding issue is to understand the emergence of living subjectivity from living being, where living being is understood as already possessed of an interiority that escapes the objectivist picture of nature' (Thompson, 2007, p. 236). Following this analysis, we could distinguish between the *morphology* of the body and its dynamical, sensorimotor *interplay* with the environment. This paradigm shift from the standard view presents a different way to think about the explanatory gap. In this case there is no dualism between mental and physical entities, but rather, there is continuity between two different typologies of embodiment (Hanna & Thompson, 2003). The traditional formulation of the hard problem of consciousness calls for a solution in terms of accepting the separation between mental and physical (dualism) or in terms of reducing the mental into the brain (materialism). If a zombie is functionally and corporeally identical to me, then it is not even conceivable to imagine a scenario in which they lack conscious experience. Subjectivity is animation, but animation is also coupling with the environment. You cannot separate these two categories without applying a rough *reification* (see also Sheets-Johnstone, 2012). Note that this move enables inference of a *quasi*-similar conclusion to the one provided by Dennett, though without reducing subjectivity to brain mechanisms.

While Dennett refuses to solve the hard problem because of the controversial usage of *qualia*, many proponents of the enactive paradigm refuse to solve the hard problem because both dualism and internalism (or reductionism, or physicalism) are not valid alternatives.

Hutto and Myin are quite straightforward about this point. In their "recipe" (Hutto & Myin 2013, p. 168) to overcome the hard problem of consciousness, *step two* asks the reader to 'deny that there is a relation between the phenomenal and physical that needs explaining. Consequently, solving the Hard Problem isn't just hard; it is impossible' (ibid., p. 169). Without the metaphysical distinction

between mental and physical phenomena, we are left with two different modalities of *being-in-the-world* (Heidegger, 1927) - the body's physiology and its embodied subjectivity - though conceived as fundamentally inseparable. In other words, the sensorimotor primacy posited by the *enactive perspective* moves away from the notion of qualia, considering phenomenological consciousness as a world-involving skilful activity (Noë, 2009).

The questions "where is the mind?" or "*where* is the tension provoked by a dominant chord during a cadence if it is not present in the outside world?" do not make any sense, because phenomenal experience involves processes that are part of the dynamic interplay that links an animal with the environment, throughout sensorimotor patterns of meaningful (i.e. goal-directed) actions and consequently not reducible to mechanisms inside the skull. The tension of the dominant chord therefore has no location, being part of the mutual determination involving musical subjects and objects. Pelinski (2005) provides a useful insight into this problem, when considering musical contexts: 'it would be a mistake to interpret the insistence on the embodiment of our musical experiences as a naïve attempt to substitute reason with the body, or intersubjective rationality with subjective experience. Neither is the body in the mind nor is the mind in the body: both phenomena are imbricated in musical experience to such an extent that it seems meaningless - and unnecessary - to create "clear and distinct" representations of one or the other'. If we are therefore to overcome the Cartesian leftovers of many traditional ways to think about music - for example in terms of *cognitive ability* (Sloboda, 1985) - the first step will be to consider a living system in its *totality*, and not as an intellectual web of representations inside its skull, nor as a functionally body-independent, brain-centred organism.

2.5 - A challenge in music research

Before exploring the implications of a truly enactive approach to musical experience, a word on the methodology. It is commonly assumed that the first step in conducting research on music might be providing a clear and precise idea of the investigated object. But, regrettably, a simple and unproblematic definition which sums up in a bare statement all the different features and nuances of past and present music doesn't exist yet (Merriam, 1964; Molino, 1975; Cook, 1990; Nettl, 2005) for 'there is no limit to the number or the genre of variables that might intervene in a definition of the musical' (Nattiez, 1987, p. 42). Furthermore, this strategy would also be pointless as not always what *we* - Western listeners, scholars or musicians - consider music is intended as such (Bohman, 2002; Delalande, 2009). The complexity of musical practices can indeed be described as a set without well-defined boundaries (Giannattasio, 1998) impossible to reduce into a predefined and abstracted notion of music.

However, this crucial position has not always been accompanied by an adequate epistemological caution, and many contributors in current music research still continue to pursue the goal of investigating music given a predefined idea or looking explicitly for a strict definition, two strategies which may lead to rough argumentations and inadequate logical passages. Usually, in fact, the classic line of reasoning aims to define music from one of its physical properties (e.g. McDermott & Hauser, 2005, p. 30) or from one of its specific functions (e.g. Schubert, 2009-2010, p. 76). Criticizing the first speculative paradigm, Cook states that although 'there have been many attempts to define what music is in terms of the specific attributes of musical sounds [...] it is not

possible to arrive at a satisfactory definition of music simply in terms of sound' (Cook, 1990, p. 10-11). Anyhow, Robert J. Zatorre, trying 'to make sense of a complex phenomenon such as music' (Zatorre, 2001, p. 232) justifies his decision to explore the processing of pitches asserting that 'pitch appears to be a central aspect of all music' (ibid.) apparently assuming the problematic (Piana, 2007) distinction between *musical* and *non musical* sounds (see Pierce, 1983) and generalizing from an aprioristic notion of music, despite a great number of possible counterexamples: for example Nattiez (quoted in Giannattasio, 2005, p. 989) shows how the Inuit population use the term *nipi* not only for all the facts that we intend as musical, but also for noise and language. Similarly, Japanese *shakuhaci* music and *sanjo* music of Korea 'fluctuate constantly around the notional pitches in terms of which the music is organized' (Cook, 1990, p. 10) and some African music (see for example Arom, 2000) - i.e. percussive music - does not require fixed pitches at all. Finally, also in our familiar Western music, we can see an emancipation of the so-called non-musical sounds perhaps since Luigi Russolo's manifesto *L'arte dei rumori*⁶ (1913) and, then, in many contemporary musical techniques, from microtonality to electronic music (Holmes, 1985; Griffiths, 1995). Zatorre's approach will, at most, shed light on how the human brain processes pitches through a precise description of its functional organization, a different goal from understanding the *complex phenomenon of music*.

Similarly, an approach based on the study of a single function cannot explain the various manifestations of musical practices. Let's consider for instance the argument scheme used by the evolutionary psychologist Geoffrey Miller to demonstrate that 'music is functionally analogous to sexually selected acoustic displays in other species' (Miller, 2000, p. 338). Forcing the classical Darwinian approach (Darwin, 1871; Kivy, 1959) the author assumes that music and dance are basically two sets of indicators of the executer's state of health, force and coordination, suggesting that the function of rhythm would be to show the cerebral capacity to sequence complex movements (see also Mithen, 2005), while melodic creativity reveals the singer's inventiveness and intelligence. The article stresses the possibility that chant has to be supposed dangerous as it could make the singer observable (audible) by predators. Furthermore, considering Jimi Hendrix's sexual behaviours⁷ as a current example of how music has conserved its original sexual-selection origins, Miller adds that 'our ancestral hominid-Hendrixes could never say, 'OK, our music's good enough, we can stop now', because they were competing with all the hominid-Eric-Claptons, hominid-Jerry-Garcias, and hominid-John-Lennons. The aesthetic and emotional power of music is exactly what we would expect from sexual selection's arms race to impress minds like ours' (Miller, 2000, p. 331). The author's argument, however, seems to be based on a predefined idea of music derived from only one of its presumed phylogenetic functions (Merriam, 1964) and, not without a hint of irony, Fitch (2006) analysing Miller's points, wrote: 'for every Bach with many children there may be a Beethoven who died childless, and for every popular conductor or lead guitarist there may be a lonely oboist or bassist' (ibid. p. 201). According to David Huron (2003) in fact, there are at least eight theories about the birth of music:

⁶ Translated as *The art of noises*, the book contains the main ideas of his *futurist* aesthetics, namely that the industrial revolution had given modern men a greater capacity to appreciate more complex sounds than the traditional ones used in music. That's the reason why he construed the *intonarumori* (noisemaker) a noise-generating device to be used in concerts and performances.

⁷ 'He did have sexual liaisons with hundreds of groupies [...]. As Darwin realized, music's aesthetics and emotional power, far from indicating a transcendental origin, points to a sexual selection origin where too much is never enough' (Miller, 2000, p. 331).

- *Mate selection*: music is an indicator of the executer's state of health, force and coordination, playing a key role in sexual reproduction processes.
- *Social cohesion*: the establishment of music is based on its power to gather people, a highly important feature considering the history of humanity itself.
- *Group effort*: through synchronisation, musical practices helped the groups' coordination, facilitating collective working practices.
- *Perceptual development*: music is an activity aimed at developing peculiar perceptual skills
- *Motor skill development*: singing and making music represent the concrete basis of some motor abilities. From this perspective, singing can be considered an antecedent of speaking.
- *Conflict reduction*: in contrast with language, which may create problems related to the arguments debated, singing together might have provided a safer social activity.
- *Safe time passing*: music can be considered as safe activity, where nothing harmful can happen.
- *Transgenerational communication*: music favoured the transmission of oral information across countries and time.

Besides these theories, moreover, we can add the *homeostatic* hypothesis provided by Niels Wallin (2000) and the polemical counter-proposal of Steven Pinker (1997) who considers music to be *auditory cheesecake* (p. 595). Given the relevance of each proposal, it seems hardly accomplishable to provide an exclusive, defining, account of music by focusing on just one of them.

Outside the field of the phylogeny of music, ethnomusicology and comparative musicology often make comparisons between different notions of music investigating their social and cultural aspects in local and global contexts (e.g. Blacking, 1973; Feld, 1982; Rowell, 1992) and, even if defining music 'is not the ultimate aim of the ethnomusicologist' (Nettl, 2005, p. 25) one of the discipline's main tasks seems to be 'studying the definitions provided by the world's musical cultures in order to shed light on the way of conceiving music' (ibid.). Though not explicitly dealing with any sort of definition, the methodology used by most of music analysis shares an implicit assumption with the ethnomusicological approach: music can be considered as an *already structured material* shaped in well defined parameters (pitches, rhythm, dynamics etc.) since its *constitution* can be studied only within a sociological and cultural perspective (in the case of ethnomusicology) or through its inner rules (e.g. tonality, modality, etc.) from musical analysis.

On the other hand, implicitly considering the musical signal as an unidirectional, external, stream of information coming from the environment, many studies on the physiological basis of music

perception aim to localize and describe domain-specific, innate and independent brain centres which represent a high level interactive system of music processing. Indeed ‘it is common in this literature to read suggestions that a certain cognitive characteristic (e.g. pitch perception) is governed by neural tissue at a certain location (e.g. primary auditory cortex)’ (Tan et al., 2010, p. 54) and, for instance, Peretz and Coltheart (2003) admit that ‘musical abilities are now studied as part of a distinct mental module with its own procedures and knowledge bases that are associated with dedicated and separate neural substrates’ (ibid., p. 688). As Prinz puts it, this paradigm assumes that ‘like a Swiss Army Knife, the mind is an assembly of specialized tools, each of which has been designed for some particular purpose’ (Prinz, 2006a, p. 22). As we saw, however, this model is flawed and based on aprioristic assumptions derived from the Cartesian tradition.

2.6 - The spectre of objectivism

In his book *Ways of listening: An Ecological Approach to the Perception of Musical Meaning* (2005), Eric Clarke explicitly acknowledges the need to overcome the dichotomies between the *subjects* and *objects* of a musical context, *passive* and *active* listening, *autonomy* and *heteronomy* of musical experience. However, as we saw in the previous paragraph, most of the research in music cognition is still dominated by radical dualistic commitments. In particular, it relies on the assumption that music is something *external* - outside from our mind - while music cognition is something *internal* - because the mind is supposedly located inside our head.

While musical analysis, for example, studies the *objective* properties of music, the modular account of music cognition reflects the opposite view, considering the *subjective* individual properties (mental states, brain states, or modules), admitting a strong separation between *musical subjects* and *musical objects*. But being committed to these radical divisions would leave the Cartesian presuppositions unaltered. David Huron’s methodology exemplifies this separation very well. On the one side we would have a deep analysis of the statistical properties of music, while on the other side there would be brain states, only *causally* related to music through *qualia*. In light of these considerations, therefore, we could make an approximate distinction between two different ways to investigate music and musicality referring to: (i) an *objectivist* and (ii) a *subjectivist* perspective.

Both these approaches take for granted the relation between music and the perceiver/executer focusing alternatively on the *subjective* and the *objective*⁸ side, considered as two different aspects of musicality. But in the concrete musical activity the subject and the object are two inseparable features of any musical experience (Cook, 1990). The above mentioned approaches are two different sides of the same coin, and represent what Husserl (1901; 1907; 1936) called *objectivism*: the scientific and naturalistic claim that reality is objective, and that sense data correspond with it, excluding from the research the first-person viewpoint. Noticing that the natural attitude is an

⁸ With regard to this point, we could also use the classical phenomenological distinction between *noetic* and *noematic*. Using Husserl’s terminology, *noematic* refers to that which is experienced (it’s that through which the object is grasped) while *noetic* is ‘the concretely complete intensive mental process approached in such a way that its noetic components are clearly emphasised’ (Moran, 2000, p. 156). In an intentional relationship between a subject and an object, the *noetic* is real and fundamental (as acts of consciousness), while the *noematic* is dependent and unreal because that which is perceived is constituted, thus far as it is perceived, by the subject’s intentional acts.

attitude that obscures itself and remains unknown to itself, Husserl invites us to take off our usual blinker to look from a new perspective at the complexity of the world and its various representations going ‘back to the things themselves’ (Husserl, 1901, p. 7; 1912-1929 [first book]).

To show the ambivalence of the objective and the subjective side in the realm of musical experience we might consider the article by Fred Lerdahl (2009) *Genesis and architecture of the GTTM project*, written on the occasion of the twenty-fifth anniversary of the publication of *a Generative Theory of Tonal Music* (Lerdahl & Jackendoff 1983) which retrospectively shows the birth of one of the most discussed and influential theories in music psychology (Menin, 2009, p. 15). The author, finding in the Chomskian linguistic theory the crucial influence for structuring the lines of research developed by him and Jackendoff, specifies the reasons of their interest:

Our interest was not in a literal transfer of linguistic to musical concepts, as Leonard Bernstein (1976) attempted. Rather, it was Chomsky’s way of framing issues that attracted us: the supposition of specialized mental capacities, the belief that they could be studied rigorously by investigating the structure of their outputs, the distinction between an idealized capacity and its external and often accidental manifestations, the idea of a limited set of principles or rules that could generate a potentially infinite set of outputs, and the possibility that some of these principles might be unvarying beneath a capacity’s many different cultural manifestations (Lerdahl, 2009, p. 187).

In other words, the two scholars were enamoured by the possibility to use a *subjective* methodology in the study of music. However, after a short section about Heinrich Schenker’s attempt to define a fundamental structure (*Ursatz*) at the basis of the tonal music’s complexity, they criticized⁹ such a setting of the problem, orienting their research to an apparently opposite side, though conceived in line with the assumptions and interests of the Chomskian approach:

Rather than begin with a putative ideal structure and generate musical surfaces, we would begin with musical surfaces and generate their structural descriptions [...]. Three methodological perspectives borrowed from generative linguistics helped launch the enterprise. First, we would assume as given the musical surface - essentially quantized

⁹ ‘For a number of reasons we soon gave up this Chomskian-Schenkerian approach. First, we could not justify the *Ursatz*. Although this *a priori* construct was understandably central to Schenker, a thinker steeped in 19th-century German philosophical idealism, its status made little sense to a modern, scientifically inclined American. Nor could schema theory in cognitive psychology (Neisser, 1967) defend this construct, for the *Ursatz* is too remote from a musical surface to be picked up and organized by a listener who is not already predisposed to find it. Second, the *Ursatz* is inapplicable to music of other times and cultures. We sought a theoretical framework that could accommodate diverse musical idioms. Third, the non-rhythmic character of the *Ursatz* presented a formal and musical problem. How was rhythm to be introduced into the derivation, and why should it have inferior status? Fourth, even supposing that the *Ursatz* or some comparable structure could be justified as a foundation, there would be very many possible ways of generating a given musical surface. What would determine an acceptable derivation? How could derivations be constrained? Fifth, it was not interesting to begin our work by considering abstract background musical structures and presumed transformations; the exercise felt too abstract. Sixth and most importantly, what of psychological interest would there be even if we managed to build a system that generated this or that piece from an *Ursatz*-like foundation? What mattered to us was not the output per se but the structure attributed to the output. It was not clear how generating a piece could reveal much about mental structures and their principles of organization’ (Lerdahl, 2009, p. 188).

pitches and rhythms with dynamic and timbral attributes - without worrying about the complex perceptual mechanisms that construct the surface from the audio signal. Second, our quest for cognitive principles would proceed from our own musical intuitions. Only later would we seek experimental corroboration. Third, we would build a final-state rather than processing theory, on the view that it was advantageous to specify the mental structures in question before trying to articulate how they operated in real time (Lerdahl, 2009, pp. 188-189).

Although this approach has been developed from the original theoretical paradigm it presents a radical form of *objectivism*, showing the common presupposition of both methodologies, namely that it's productive and useful to investigate *distinctly* objective and subjective aspects of musical phenomena, implicitly assuming that the specific features of music are substantially avulsed from the concrete practices through which it constitutes itself. But, as Evan Thompson states, 'to reduce conscious experience to external structure and function would be to make consciousness disappear [...]; to reduce external structure and function to internal consciousness would be to make external things disappear' (Thompson, 2007, p. 225).

Given those remarks we might consider an alternative approach for studying this still mysterious musical *object*: an *enactive* way, which considers the perceived and the perceiver as two inseparable aspects of musical experience for 'in this transaction between the subject of sensation and the sensible it cannot be held that one acts while the other suffers the action, or that one confers significance on the other. Apart from the probing of my eye or my hand, and before my body synchronises with it, the sensible is nothing but a vague beckoning' (Merleau-Ponty, 1945 [1962], p. 214). It is no coincidence that I continuously refer to Maurice Merleau-Ponty and Edmund Husserl in this paragraph. *Enactivism* shares indeed many aspects of its characterisation with the philosophical movement called *phenomenology*. Pure phenomenological research seeks essentially (i) to describe rather than explain and (ii) to start from a perspective free from hypotheses or preconceptions (Husserl, 1901). In particular, as the next chapter will show, the enactive notion of *sense-making* can be derived from the phenomenological notion of *intentionality*.

A genuine investigation of music, therefore, cannot use a merely analytical approach that does not take into consideration the subject, nor a cognitivist perspective, which tries to explain our musical behaviour in light of aprioristically defined mental processes. As emerged, after all, these standpoints share the same assumption that music and mind are two distinct categories and therefore require two different methodologies to be investigated. In contrast to this view, enactivism is focused on the codetermination between embodied agents and their environment, ruling out any possible distinction between mind and matter, brain and mind, brain and body, internal and external, subjective and objective.

Before embarking on the exploration of such an approach, let's conclude with a brief summary of the main differences between enactivism and cognitivism. Daniel Dennett's critical review (1993) of Varela and colleagues' seminal work of enactivism - *The Embodied Mind* (1991) - provides a clear and brief summary of the main assumptions at the basis of the enactive approach. As he points out, the fundamental differences are encapsulated in answers to three questions (Dennett 1993, p. 121):

- *Question 1: What is cognition?*

Cognitivist Answer: Information processing as symbolic computation - rule-based manipulation of symbols.

Enactivist Answer: Enaction. A history of structural coupling that brings forth a world.

- *Question 2: How does it work?*

Cognitivist Answer: Through any device that can support and manipulate discrete functional elements - the symbols. The system interacts only with the form of the symbols (their physical attributes), not their meaning.

Enactivist Answer: Through a network consisting of multiple levels of interconnected, sensorimotor subnetworks.

- *Question 3: How do I know when a cognitive system is functioning adequately?*

Cognitivist Answer: When the symbols appropriately represent some aspect of the real world, and the information processing leads to a successful solution to the problem given to the system.

Enactivist Answer: When it becomes part of an ongoing existing world (as the young of every species do) or shapes a new one (as happens in evolutionary history).

Chapter 3

The mark of embodiment¹⁰

Nothing determines me from outside, not because nothing acts upon me, but, on the contrary, because I am from the start outside myself and open to the world

- Maurice Merleau-Ponty, *Phenomenology of Perception*

To say that I have entered into the world, “come to the world”, or that there is a world or that I have a body is one and the same thing

- Jean Paul Sartre, *Being and Nothingness*

3.1 - From dualism to Neurophenomenology

In the previous two chapters I have argued that traditional cognitive science and most of modern music psychology are still under the influence of Descartes' philosophy. As emerged, this commitment gave rise to problematic models of music cognition, given the intrinsic dichotomies that characterise any form of dualism. In particular, the separation between musical objects (considered as *external streams of information*) and musical subjects (conceived as *disembodied minds*), reflects the aprioristic assumptions that (i) music is something already defined and fully constituted in the external world and that (ii) music is a cognitive representation, underlined by specialized mental or brain states. In both cases, what underlies each corresponding research methodology is the *reification* of the musical object, always considered as something *fully given* - in the external world or inside the head. For this reason, I have also argued that no definition of music is necessary (or sufficient) for a genuine investigation of related issues. Starting an investigation on *music cognition* with an already given definition of *music* would only vitiate any possible methodological paradigm. As I have tried to show, many scholars in music psychology are taking for granted *assumptions* about the nature of music and the nature of cognition that are (i) not corroborated by concrete musical practices and (ii) not sufficiently defended from a theoretical and empirical standpoint. Is there any alternative?

¹⁰ This chapter contains material previously published as **Schiavio, A.** (2012). Constituting the musical object. A neurophenomenological perspective on musical research. In *Teorema*, 31(3), pp. 63-80, and as Matyja, J., & **Schiavio, A.** (2013). Enactive music cognition. Background and research themes. In *Constructivist Foundations*, 8(3), pp. 351-357. Moreover, it contains material from the following articles in preparation: **Schiavio, A.**, & Gerson, S. (in preparation). An enactive perspective on infants' musicality, and **Schiavio, A.**, Menin, D., & Matyja, J. (in preparation). Music in the flesh. Sense-making and embodied simulation in musical understanding.

The enactive approach, along with its intrinsic *phenomenological* and *embodied* roots, seems to be a perspective able to treat music as an object *in constitution* - an object that is never externally given or fully immanent to one's mind - and to provide a biologically plausible interpretation of music cognition. The point here is that subjects and objects of musical experience are two inseparable categories that actively and mutually participate in the process of perception and sense-making. In order to justify this position, this chapter will focus on the roots of the enactive approach, integrating theoretical and empirical arguments, thus drawing from the working hypothesis of what Francisco Varela called *neurophenomenology*¹¹. According to Thompson (2007), neurophenomenological research aims to integrate (i) the phenomenological analysis of consciousness, (ii) the dynamic system theory and (iii) the empirical investigation of biological systems. Accordingly, my plan for this chapter is therefore the following: I will

- (i) Introduce the crucial notion of *musical intentionality*, emerging from a purely phenomenological analysis
- (ii) Show how the *dynamic system theory* is used in *embodied* cognitive science
- (iii) Analyse the *empirical* discovery of the human mirror mechanism as well as its implications and interpretations
- (iv) Redefine musical intentionality on the basis of (ii) and (iii)

The notion of musical intentionality, or *musical sense-making*, represents an important step forward for musical research as it defines musical subjects and musical objects in a completely different way from the traditional cognitive science of music. In this entire chapter I will maintain a working methodology that always looks at approaches, models and concepts with an enactive eye. Rather than just presenting a given standpoint I will try to propose a dialogue between these approaches and the enactive perspective.

3.2 - Intentionality and Musical Intentionality

The doctrine of intentionality provided by Brentano (1874) reintroduced in modern philosophy the discussion on the Latin term *intentio* used among the scholastic thinkers (see Chisholm 1967) to indicate *what is before the mind in thought*. This term 'literally means a tension or stretching (from the verb *intendere*, to stretch)' (Crane, 2001, p. 9) and actually derives from the Aristotelian word *noema* (Knudsen, 1982). Brentano's main goal, in using this terminology, was to make a clear distinction between *physical* and *psychical* phenomena, arguing that intentionality is the main characteristic of all acts of consciousness:

¹¹ This research methodology holds that 'phenomenological accounts of the structure of experience and their counter parts in cognitive science relate to each other through reciprocal constraints. The key point here is that by emphasizing a co-determination of both accounts one can explore the bridges, challenges, insights and contradictions between them. This means that both domains of phenomena have equal status in demanding a full attention and respect for their specificity' (Varela 1996, p. 343).

Every mental phenomenon is characterised by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction towards an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not all do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on. This intentional in-existence is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it. We could, therefore, define mental phenomena by saying that they are those phenomena which contain an object intentionally within themselves (Brentano 1874, [1995], pp. 88-89).

However, many scholars considered Brentano's thesis about the object-directedness of *all* the acts of consciousness as *too strong* (McIntyre & Woodroof Smith, 1989). Considering consciousness as intentional insofar as it refers to, or is directed at, an object, Husserl (1912-1929 [first book]; 1931) argued that consciousness may have intentional and non-intentional phases, and intentionality is what gives to it its objective meaning. Indeed, for example, 'moods such as depression or euphoria are not always "of" or "about" something. Moreover, as Husserl notes, sensations such as pain or dizziness are not obviously representational or "directed towards" some object' (McIntyre & Woodroof Smith, 1989, p. 149). But what happens when thinking about non-existent objects (i.e. Pegasus)? In this case there is no actual object, but only *the act of thought with a particular intentional content* - a particular *meaning* (Crane, 2006).

Thus, what distinguishes intentional from non-intentional experiences is the former's having intentional content. Husserl, whose work can be seen as the first proper phenomenological investigation of intentionality (Gallagher & Zahavi, 2008) holds that the intentional directedness does not consist in the relation to special (mental) objects towards which one is directed, but rather is conceived as the possession by mental acts of a *noematic structure*.

The intentional openness of consciousness is an integral part of its being, not something that has to be added from without. Thus, intentionality does not presuppose the existence of two different entities - consciousness and the object. All that is needed for intentionality to occur is the existence of an experience with an appropriate internal structure of object-directedness (Zahavi, 2003, p. 21).

There is intentionality only when there is a duality between the *noetic* and the *noematic* side as 'intentionality is the name for a certain "achievement" or "accomplishment": that of the consciousness of identity from within the "Heraclitean flux" of flowing subjective life. Any object is a "pole of identity" within such a flux' (Smith, 2003, p. 68). Another problem regards the ontological status of the intentional objects of the experience. According to Husserl (1907; 1912-1929 [first book]; see also Ghigi, 2007) objects in perception are always transcendent because they are *experienced as* perspectively given (what is given to consciousness is not the whole object but

only a particular aspect). Only my consciousness can make sense of that particular transcendental object, and experience it as a unity:

Reflective experience teaches us that there is no progressively perceived thing, nor any element perceived as a determination within it, that does not appear, during perception, in multiplicities of different appearances, even though it is given and grasped as continuously one and the same thing. But in normal ongoing perception, only this unity, only the thing itself stands in the comprehending gaze while functioning processes of lived experience remain extra-thematic, ungrasped and latent. Perception is not a simple empty “having” of perceived thing, but rather a flowing lived experience of subjective appearances synthetically uniting themselves in a consciousness of the self-same entity existing in this way or that (Husserl, 1925-1962 [1977], quoted in Smith, 2003, p. 67).

As a subject I cannot experience the world. Rather, what I can experience is the intentional meaning of the world, while the world itself is the object intended, which is transcendental by definition. Husserl argued that to discover the intimate essence of an object without taking it for granted, as we commonly do in our everyday life, the unique useful approach consists in avoiding the natural attitude of the naïve observer with the suspension of belief in the existence or non-existence of phenomena. Husserl (1912-1929) uses the term *epoché* (a word derived from the sceptics’ tradition, which means *cessation*) to describe this process of suspension of judgements (Moran, 2000, p. 148).

Suspending the empirical subjectivity, our consciousness (now considered as *pure Ego*) can define the pure essence of a psychological phenomenon. The pure Ego, finally free from the natural attitude has the Cartesian *cogito* as its principle and this “I think” can direct its acts (*cogitationes*) immanently - when the objects are within the Ego - or transcendently - in the realms outside my Ego. In the natural attitude, as a naïve observer, when I look at an external object like a tree, I consider it as a transcendent object of the world (I cannot see, for example, its hidden side), but under the bracketing of existence all the beliefs about its actual existence are excluded. The intentional object can be in fact considered as a *determinable x in a noematic sense* (Husserl 1912-1929 [first book]), for the noema is what relates my thought to the intended object.

The tree *simpliciter*, the physical thing belonging to nature, is nothing less than this *perceived tree as perceived* which, as perceptual sense, inseparably belongs to the perception. The tree *simpliciter* can burn up, be resolved into its chemical elements, etc. But the sense - the sense of *this* perception, something belonging necessarily to its essence - cannot burn up; it has no chemical elements, no forces, no real properties (Husserl, 1912-1929 [first book], quoted in Moran, 2000, p. 157).

Listening or recalling in memory the incipit of Mozart’s *Symphony 40* represent two different modalities of experiencing that particular melody: but there is something that makes those pieces of music the *same* piece for me, even if under different conditions. What makes this possible is the

noema. There is a series of sounds with silences (intentional object); there is my listening to this set of sounds and silences (intentional experienced): The *effective* content of the experienced (sounds and silences per se) cannot be sufficient to revealing its *intentional* content, namely its meaning from my own perspective. Indeed, I have to *give sense* to this object if I want this melody to constitute that particular melody, the one I hear in this way. To have sense or to have something in mind is the main feature of every consciousness, which is never a general experience but an experience with a sense, a noetic experience (Husserl, 1912-1929 [first book]).



Fig. 4
Incipit from Mozart's Symphony 40

Let's think for example about Pablo Casals who, at 93, told an interviewer that he played one of the Bach's 48 preludes and fugues every morning, having done so also for the previous 85 years. Asked if he had not grown tired of them after all this time, he said that actually he found fresh beauties in them each time he played one, for every performance was like an act of discovery (Sacks, 2007). Using a phenomenological terminology we could say that his directedness towards the musical object continuously *constitutes* it through every intentional act. The transcendental nature of the object each time reveals a different intentional content, for 'the objects of which we are "conscious", are not simply in consciousness as in a box, so that they can merely be found in it and snatched at in it; [...] they are first constituted as being, what they are for us, and as what they count as for us, in varying forms of objective intention' (Husserl, 1901; quoted in Zahavi, 2003, p. 27). As Thompson puts it

When we ask the constitutional question of how objects are disclosed to us, then any object, including any scientific object, must be regarded in its correlation to the mental activity that intends it. This transcendental orientation in no way denies the existence of a real physical world, but rather rejects an objectivist conception of our relation to it. The world is never given to us as a brute fact detachable from our conceptual framework. Rather, it shows up in all the describable ways it does thanks to the structure of our subjectivity and our intentional activities (Thompson, 2007, 82).

This is the reason why Pelinski (2005) states that intentionality 'is fundamental for a musical aesthetics conceived from a phenomenological perspective: a piece of music doesn't concretize its potentialities as a meaningful musical event if it doesn't become the object of an intentional perception'. So, given these considerations we could propose a first characterisation of musical intentionality by saying that *musical intentionality comprehends all those intentional acts of my*

consciousness directed towards a musical object, which receives its configuration through its intentional constitution. In other words, through an intentional relation, a subject's musical cognition constitutes the musical object being intentionally directed towards it. The most important feature of musical intentionality, therefore, is the ability of consciousness to *give sense* to the musical object, through its intentional, object-directed acts. But what does it mean, "to give sense"? And, above all, is this ability limited to our high level cognition? In his masterpiece *Phenomenology of perception* (1945) Merleau-Ponty argues that the real nature of perception is not consciousness, but the body, intended not as the simple object existing into the physical world, rather considered a *lived* and *living* body. The Husserlian *Cogito*, the principle of pure Ego, becomes now "I can" instead of "I think".

Bodily experience forces us to acknowledge an imposition of meaning, which is not the work of a universal-constituting consciousness, a meaning which clings to certain contents. My body is the meaningful core which behaves like a general function, and which, nevertheless, exists and is susceptible to disease (ibid., [1962], p. 46).

Every consciousness is no more consciousness of something; rather it is *perception of something*, a perception that is strictly linked to our body, assumed as a permanent condition of experience in its constant openness to the world. We are not transcendental subjects in the world. We are present in a *world that immediately makes sense*. Being in the world is already being open to meaning-ascription, as *to perceive is to give sense* (Merleau-Ponty, 1945). The world in which we are situated discloses itself without forcing any explicit thought about it. We are entwined with the world, which simultaneously is entwined with us. As emerges, the distinction between *inner* and *outer* seems to be unnecessary to understand the modalities in which an agent is situated in the world. The analysis of the body of a subject, being the category that allows our being-in-the-world, is therefore essential for understanding the relationship between mind and world, mind and body, and my mind and the minds of the others.

In saying that the mind is intentional, phenomenologists imply that the mind is relational. "Being-in-the-world" (Heidegger) and the "lived body-environment" (Merleau-Ponty) are different ways of articulating this kind of relation. The spatial containment language of internal/external or inside/outside (which frames the internalist/externalist debate) is inappropriate and misleading for understanding the peculiar sort of relationality belonging to intentionality, the lived body, or being-in-the-world. As Heidegger says, a living being is "in" its world in a completely different sense from that of water being in a glass (Thompson & Stapleton 2009, p. 27).

Narrowing the research field to music cognition, we might assume that the genuine collaboration between phenomenology and cognitive neuroscience can develop our current knowledge of music and musicality considering that 'as an intentional object of perceptual experience, music doesn't symbolize; it doesn't reflect reality: it is reality' (Pelinski, 2005). In light of these considerations we

could propose a new formulation for musical intentionality by focusing on the co-determination between the two poles of identity (musical subjects and musical objects): *musical intentionality is constituted by the dynamic interplay between embodied agents and their (musical) environment*. In fact,

In perception we do not think the object and we do not think ourselves thinking it, we are given over to the object and we merge into this body which is better informed than we are about the world. (Merleau-Ponty, 1945 [1962], p. 238)

As Carman points out, ‘Merleau-Ponty bases his entire phenomenological project on an account of bodily intentionality and the challenge it poses to any adequate concept of mind. Embodiment thus has a philosophical significance for Merleau-Ponty that it could not have for Husserl’ (Carman, 1999, p. 206). The next section, therefore, will focus on this perspective, clarifying the meaning of terms such as “embodied agents” and “embodiment”, as well as analysing the centrality of the bodily power of action for sense-making by means of employment of the dynamic system theory.

3.3 - Embodied and Situated approaches

The word ‘embodiment’ defines the body of a living system as a constitutive category of its perceiving, knowing and doing (Gallagher, 2005). In other words, cognition (including *off-line* cognition) depends on processes that are intrinsically connected to the organism’s body, thus being widely distributed beyond the boundaries of the brain (Hutchins, 2010). This is the central idea of the so-called *embodied approach* (Shapiro, 2010; Gallagher, 2011), a theoretical paradigm that is usually seen as having roots in the phenomenological philosophy of Maurice Merleau-Ponty (1945) and in the ecological psychology of affordances provided by James Gibson (1979). Consequently, “embodied cognition” refers to many different ideas concerning - among the others - the role of bodily actions for sensorimotor representations (Goldman, 2012), and for intersubjectivity (Gallese 2010), how bodily schemas account for mental skills (Lakoff & Johnson, 1999), the offloading of computational processing into a wider bodily structure outside of the skull (Wheeler, 2010; Clark, 2008), and the role of situated sensorimotor coupling with the environment in order to bring forth a living system’s domain (Varela et al., 1991).

As reported by Shaun Gallagher and Dan Zahavi (2008), the focus on the agent’s body reflects a position free not only from the Cartesian dualism, but also from any form of materialistic reductionism. In fact, the embodied perspective does not simply put the mind into the body - implicitly maintaining a distinction between mind and body as two different categories. Rather, this approach aims to rethink the ordinary notions of body and mind. *Corporeality*, as Merleau-Ponty (1945) points out, is a notion that goes beyond the mere distinction between physiology and psychology. My body does not divide my subjectivity from the world. Rather, it is a *transparent tool* that allows me to be (and to act) in a world that discloses itself through it, through my perceptive, bodily-based, experiential properties. As Romdenh-Romluc puts it:

Although I can perceive my body in the same way that I perceive other things - I can see it, touch it, taste it, etc. and, in these cases, it is the object at which some state of awareness is directed - there is another form of bodily awareness. It is what can be called an “adverbial” form of awareness. Rather than being aware of my body, I am simply aware in my body (Romdenh-Romluc, 2011, p. 105).

In the context of cognitive science, the embodied approach is part of what Andy Clark (1997) labelled as *embodied-embedded* perspectives (see also Wheeler, 2005), which heavily contributed to the fall of the cognitivist orthodoxy in the 1990s. The focus on the agent’s embodiment, however, cannot be conceived as a *solipsistic* condition for cognition: a subject is always a subject in the world. In fact ‘cognition is a situated activity which spans a systemic totality consisting of an agent’s brain, body, and world’ (Froese 2009, p. 17). Gibson’s ecological approach to perception is usually seen as one of the most important contributions with regard to the situatedness of agents in the world. As Chemero (2009) suggests, Gibson’s standpoint has three major tenets, where (iii) follows from (i) and (ii)

- (i) Perception is direct (no mental representations)
- (ii) Perception is primarily for the guidance of actions (it is not for neutral information gathering)
- (iii) Perception is of *affordances* - that are directly perceivable, environmental opportunities for behaviour

In Chapter 4 I will discuss the problem of affordances with regards to musical experience, analysing the notion of *musical affordances* and proposing an alternative characterisation for them, different from the ones provided by music psychologists. For now, instead, we should focus on how Gibson’s approach has been influential for the embodied-embedded theories of cognition. When discussing enactivism, for example, its influence is clear: the brains of living creatures do not represent objects of the world as proponents of classic cognitivism typically assume. Rather, each ecologically-situated creature *enacts*, *brings forth* or *co-constitutes* a world through evolutionarily-selected sensorimotor and goal-directed patterns of meaningful actions.

The enactive approach not only holds that bodily activity is crucial for cognition but also aims to provide a biological account for autonomous agency (Weber & Varela, 2002). The complementary focus on two types of subjectivity (biological and phenomenological, as emerged in the discussion on qualia) is what allows the enactive paradigm to be distinguished from other competing approaches in phenomenology, cognitive science or philosophy of mind (see Thompson, 2007). The rejection of representationalism, the overcoming of the distinction between inner and outer, and the radical codetermination between living systems and the world are all central commitments for the enactive research and are indeed deeply influenced by Gibsonian psychology.

However, although the notion of *affordance* itself is radically linked to the agent’s sensorimotor patterns, it might eventually ‘lead to a research strategy in which one attempts to build an ecological theory of perception entirely from the side of the environment. Such a research strategy ignores not

only the structural unity of the animal but also the codetermination of animal and environment' (Varela et al., 1991, pp. 204-205). As Susan Hurley puts it:

Perhaps the received tradition has focussed too much on the internal aspects of perception and ignored the external aspects. But we can correct this bias and take on board the role of movement in making information available, without going to the opposite extreme of denying that the brain processes information at all [...]. The right response to Gibson is ecumenical: both movement through real environments by whole organisms and brain activity play essential roles in extracting information from the environment and enabling a creature to have a perceptual perspective (Hurley, 2001, p. 20).

In fact, 'organism and environment enfold into each other and unfold from one another in the fundamental circularity that is life itself' (Varela et al., 1991, pp. 171-172). This interplay between living systems and world is the central pillar of the enactive perspective and thus can be seen as an original contribution in the broader domain of the *embodied-embedded* approaches.

The overall concern of the enactive approach to perception is not to determine how some perceiver-independent world is to be recovered; it is, rather, to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually guided in a perceiver-dependent world. Consequently, cognition is no longer seen as problem solving on the basis of representations; instead, cognition in its most encompassing sense consists in the enactment or bringing forth of a world by a viable history of structural coupling (Varela et al., 1991, p. 205).

Embodied cognition is *necessarily* situated, and describing the interplay between cognition, body and environment should require 'an explanatory tool that can span the agent-environment border' (Chemero, 2009, p. 25). Many proponents of embodied cognition, therefore, consider cognitive systems in terms of dynamical systems where 'a dynamical system is a set of quantitative variables changing continually, concurrently, and interdependently over time in accordance with dynamical laws that can, in principle, be described by some set of equations' (ibid.) Following Chemero (ibid.) we can say that the interaction with the environment can be best described by dynamical systems as they have parameters on 'each side of the skin' (p. 26). For example, we can use the equations used by Beer (1995a) to describe how embodied agents and environment are coupled:

$$X_A = A(X_A; S(X_E))$$

$$X_E = E(X_E; M(X_A))$$

A and E are continuous-time dynamical systems, modelling the organism and its environment, respectively, and $S(X_E)$ and $M(X_A)$ are coupling functions from environmental variables to organismic parameters and from organismic parameters to environmental parameters, respectively. It is only for convenience (and from habit) that we think of the organism and environment as separate; in fact, they are best thought of as forming just one nondecomposable system, U. Rather than describing the way external (and internal) factors cause changes in the organism's behaviour, such a model would explain the way U, the system as a whole, unfolds over time (Chemero, 2009, p. 26).

This is a classic description of how body and environment interact in terms of dynamical systems, as many proponents of embodied cognition assume. This point is to some extents in common among enactivism and traditional embodied approaches. There are, however, two main differences between these approaches: (i) Enaction holds a stronger antirepresentationalist position from traditional embodied cognition and (ii) Enaction allows mental processes to be distributed beyond the boundaries of skull and skin, while traditional embodied paradigms (too) often focus only on the bodily power of action, thus considering the environment only *causally*. The next paragraphs will discuss these two points in more detail.

3.4 - A radical interpretation of embodiment

As previously discussed, enactivism borrows ideas from different approaches, nevertheless providing a unique interpretation to cognition in terms of embodied action, sense-making, and the ability to enact a world. Enactivism can be seen as a radicalization of traditional embodied approaches as (i) it provides a stronger rejection of mental representations and (ii) goes beyond the idea of solely inner processing to account for cognition. If it is true that many embodied theories are *also* situated, it is also true that a too strong accent on the bodily basis of cognition might obscure the dynamic co-determination processes between animal and environment, providing a fully internalist model.

In regards to (i), probably the most famous example is provided by what Andy Clark (1997) defines as *action-oriented representations* (AOR). These kinds of representations are unlike the *propositional representations*¹² advocated by proponents of Computationalism; rather, as Wheeler (2005; 2008) puts it, AOR are (i) action-specific - they represent the world in terms of action-possibilities, (ii) egocentric - the co-ordinated system represented is always in first-person, and (iii) context-dependent - they define the contexts of activity on the basis of particular operating principles (see Hutto & Myin, 2013, p. 52). In other words, AOR can be described as inner states that at the same time inform the subject of the world in which she is embedded and prescribe possible modalities of actively interacting with it. Clark concludes: 'it may [...] be wise to consider

¹² 'Minds are [...] processors of information. Cognitive devices [used for] receiving, storing, retrieving, modifying and transmitting information' (Branquinho 2001, pp. xii-xiii). In order to facilitate these processes the mind uses mental representations. 'Mental representations might be thought of as images, schemas, symbols, models, icons, sentences and so on' (ibid. p. xiv).

the intelligent system as a spatio-temporally extended process not limited by the tenuous envelope of skin and skull [...]. Cognitive science [...] can no longer afford the individualistic, isolationist biases that characterised its early decades' (Clark, 1997, p. 221). If Clark is right in emphasising the active situatedness of agents in the world, his proposal is still committed to the idea of inner processing via representations and computational processes (see also Clark, 2009). The enactive perspective, in contrast, provides a radical rejection of mental representations - whether they are action-oriented representations or traditional, propositional, content-bearing, representations. As Di Paolo and colleagues put it:

Organisms do not passively receive information from their environments, which they then translate into internal representations. Natural cognitive systems are simply not in the business of accessing their world in order to build accurate pictures of it. They participate in the generation of meaning through their bodies and action often engaging in transformational and not merely informational interactions; *they enact a world*. Enactivism thus differs from other nonrepresentational views such as Gibsonian ecological psychology on this point [...] For the enactivist, sense is not an invariant present in the environment that must be retrieved by direct (or indirect) means. Invariants are instead the outcome of the dialogue between the active principle of organisms in action and the dynamics of the environment (Di Paolo et al., 2010, p. 39).

What does it mean that cognition is embodied action? And how is this bodily activity characterised? According to Thompson and Varela (2001) there are three dimensions of embodiment - bodily self-regulation, sensorimotor coupling and intersubjective interaction.

1) The first dimension of embodiment refers to bodily self-regulation, usually defined by the term "autopoiesis", that is the idea at the basis of the so-called *life-mind continuity thesis* (i.e. Thompson, 2007; Godfrey-Smith, 1996). This thesis holds that a living system 'is a process with the particular property of *engendering itself* indefinitely' (Stewart, 2010, p. 2). *Autopoiesis*, implies autonomy: living systems are not simply coupled with other creatures, they are not passively responding to an external stimulus, they rather enact a world, self-regulating the processes required to co-constitute the environment (Di Paolo, 2005). According to Ruiz-Mirazo and Moreno, in fact 'basic autonomy is the capacity of a system to manage the flow of matter and energy through it so that it can, at the same time, regulate, modify, and control: (i) internal self-constructive processes and (ii) processes of exchange with the environment' (Ruiz-Mirazo & Moreno, 2004, p. 240 [quoted in Thompson & Stapleton, 2009]). A creature's sense-making, therefore, has its roots in these circulatory capacities to maintain itself under precarious conditions (Varela, 1979; Thompson, 2007) thereby generating its autonomous identity.

A classic example of an autopoietic system is the *prokaryotic cell*. As Varela argues (1987), the processes of self regulation, in this case, aims to maintain the cellular organization as a unity by keeping its thermodynamic imbalance with the environment. Organisms, even bacterial cells, are maintained within a material border (membrane, skin and so forth), and the exterior - what lies beyond these boundaries - does not provide any input to the autopoietic system (see Weber, 2001). In fact, any autopoietic organism self-produces all that is needed for its maintenance. The functional

relationship with the environment, however, is not broken off. The exchange of energetic processes with the environment can be realized by the living cellular system, for example, through the regulation of the permeability of the membrane. Di Paolo (2005) defines this relation with the environment as “interactional asymmetry”. As Varela puts it, the self-regulating processes to keep the agent’s conservation as auto-sufficient establish the dialectic between agent and environment:

Whence the intriguing paradoxicality proper to an autonomous identity: the living system must distinguish itself from its environment, while at the same time maintaining its coupling; this linkage cannot be detached since it is against this very environment from which the organism arises, comes forth (Varela, 1991, p. 85).

The external world, hence, is constituted by means of the degree of complexity of the agent, that is, its inner and biological structure. Also the nervous system, for example is a closed system. As Weber notices, ‘it brings forth only its own inner states, which can be stimulated by the environment but may be not influenced causally in an ambiguous way. The nervous system does not receive information. It rather creates a world by defining which configurations of the milieu are stimuli’ (Weber, 2001, p. 15). An Autopoietic system, therefore, gives rise to “intentionality” (Varela, 1992) or, as Evan Thompson puts it, ‘living is a process of sense-making, of bringing forth significance and value’ (Thompson 2007, 158). In the case of the living cell, for example, the meanings enacted are, basically, “nutrition” and “maintaining my own integrity”. Simply put, *life means to make sense*. An agent’s self-constitution always is the constitution of a world of meanings.

Autonomy is the property of a system that self-produces and strives to maintain its identity as the system that it is. The nature of this identity is conceived of as a dynamic network of precarious processes where each process is enabled by other processes in the network and also contributes to enable other processes in the network [...]. The enactivist calls this identity autonomous because the system is constrained but not fully determined by external factors; instead it follows its own intrinsic norms. Linked to this idea is the notion of sense-making, which refers to a system’s ability to regulate its states or interactions with the world adaptively (Di Paolo, 2005). This is inextricably linked to autonomy insofar as the regulation happens with respect to the implications for the continuation of the system’s autonomous identity. For enactivism a system is cognitive when it acts in terms of value or concern with regards to its own existence (Kyselo & Di Paolo, 2013, pp. 7-8).

A detailed analysis of this dimension of embodiment, with regard to musical experience, will be provided in Chapter 5, when considering the spontaneous and self-organized exploratory musical behaviours of infants. As we will see, these kind of motor activities are supposed to be at the basis of human musicality.

2) The second dimension of embodiment is sensorimotor coupling. Susan Hurley (1998)

referred to the standard view of cognition as a “sandwich”, where the meat (cognition) is segregated between two slices of bread (perception and action). Enactivists, on the contrary, assume that perception and action are radically entwined. Autonomous agency requires that a living system generates by itself its own actions into the environment, therefore self-determining its normative domain. If an embodied agent uses sensations to guide actions in order to maintain autopoiesis, then this means that *cognition* and *life* are the same category. This is precisely the *life-mind continuity thesis* just sketched above. Cognition is life and life is embodied action. Actions, in fact, are not isolated events but rather they are present in the world, modifying it by bringing forth the living system’s affective, relational, meaningful domains. By modifying the world (or the relation with the world) through embodied actions, an agent also modifies the subsequent sensory return. This pragmatic form of knowledge depends on the vocabulary of possible actions that the creature can perform in the world. This point is decisive for the discussion of Mirror Neurons in the following paragraphs and will result as essential for the thesis that I defend in this work.

The centrality of action for sense-making is exemplified by the well known notion of *sensorimotor contingencies* provided by O’Regan and Noë (2001a). In their words sensorimotor contingencies are ‘the structure of the rules governing the sensory changes produced by various motor actions’ (O’Regan & Noë, 2001a, p. 941). Embodied action, therefore, can be seen as a mutual interplay between the environment and an agent, where the latter acquires the concepts that lead to the mastery of a particular (i.e. musical) perceptual experience, through its engagement with the world, developing a corresponding *sensorimotor knowledge* (Noë, 2004). This is a pragmatic, practical, form of knowledge - that is, what Ryle (1949) named *know how* knowledge, in contrast with a *know that* knowledge (see also Stanley & Williamson, 2001).

It is possible, however, to distinguish between a strong and a weak interpretation of this sensorimotor claim (Shapiro, 2011; Kyselo & Di Paolo, 2013): on the one hand, the exercise of sensorimotor knowledge requires an actual movement in order to reveal sensorimotor contingencies (*strong* interpretation). On the other hand there is a *weak* interpretation of embodiment, according to which ‘the exercise of sensorimotor knowledge consists only in the potential to perform those actions that define a class of sensorimotor contingencies’ (Shapiro, 2011, p. 168).

In the context of music cognition, Bennett Hogg (2011), who is also a practicing improviser and composer, defines music as *intertextual* because sounds do not carry meanings in and of themselves, but rather are the sites of: ‘complex and mediated sets of relationships between physical sounds, perceptual systems, personal associations, culturally significant gestures, bodily and emotional responses, observed actions and reactions, and culturally learned expectations.’ (Hogg 2011, p. 89). Hogg claims that from this perspective, it would be easy to see that if being conscious of music is something we *do* rather than something that is *happening* to us, it takes place within both the cultural and the physiological constraints of our bodies: music consciousness is an *embodied* and *encultured* activity. The same applies to musical improvisation. According to Hogg, it needs to be defined in terms of the *play* across our memory, embodiment and situated consciousness, and includes multisensory experiences and actions that lead to our perceptions of sounds.

Hogg’s claim resembles the position of Vijay Iyer (2004), an established jazz pianist, who maintains that music perception (and music cognition itself) is *actively* constructed by the listener, rather than passively transferred from performer to listener within the given culture and context. Interestingly, this claim parallels von Glasersfeld’s (1988) principles, according to which knowledge is actively built up by the cognizing subject instead of being passively received by

senses or means of communication. In his papers, Iyer argues that discernment of pulse and meter from a piece of music depends on a person's culturally contingent listening strategies. Iyer thus argues that certain varieties of *microrhythmic* variation (within the rhythmic expressions of our embodied involvement with music) are shown to display and carry encoded sonic traces of the culturally situated music-making bodies (Iyer, 2004). Sensorimotor coupling therefore, is the radical integration of sensorial and motor aspects for cognition. What makes this possible, is the body:

The body enables the realization of a flexible being in the world. In this, enactivism is very close to pragmatic notions of action. This implies that a body that does not engage with the world, whether by moving itself, or interacting with objects, would in time very likely cease to be the means by which we are in the world as skillful agents. It is very likely the case in general that in order to maintain bodily capacities they need to be regularly put to practice. The enactive body is therefore a body-in-action. The enactivist does not limit the role played by the body in sustaining specific cognitive skills to only a historical role (Kyselo & Di Paolo, 2013, p. 11).

The realm of music, as well, is a context where meanings are not already present in their structural or formal properties. Rather they are *enacted* by each musical subject, through different performances, different *ways of listening* and different interpretations. A musical subject is not a passive listener, as we saw. She participates in every musical experience actively, as suggested by the following empirical findings. The activation during listening tasks of neural circuits involved in motor activity and the planning of motor sequences (Carroll-Phelan & Hampson, 1996) not only allows 'hypothesis about the induction of a sense of beat or pulse in the listener' (Iyer 2002, p. 392) but also reflects the ecological dynamics of music perception, which is never a solipsistic or isolated event. Rather it is an event immersed in worldly conditions, where the ability to act is radically intermixed with sense-making. Listening to music, as will be shown in detail later, is an active, skilful, sensorimotor, exercise (Krueger, 2009).

3) The third type of embodiment (intersubjective interaction) is best understood by focusing on classical theories in social cognition. Traditional views in cognitive neuroscience and philosophy of mind address the ability to understand *others* by means of the employment of a "folk" psychology by the observer. The latter eventually 'reads the minds' of others ascribing to them beliefs, intentions and other mental states (Davies, 1994). This claim represents the central pillar of the so-called Theory-Theory (TT) and of the Rationality-Teleological Theory (RTT).

According to TT, people understand others' behaviour by developing a particular, common-sense, theory of mind consisting 'of a set of causal/explanatory laws that relate external stimuli to certain inner states (e.g. perceptions), certain inner states (e.g. desires and beliefs) to other inner states (e.g. decisions), and certain inner states (e.g. decisions) to behaviour' (Gallese & Goldman, 1998, p. 496) (see also Stich & Nichols, 1992). The intentional stance (RTT) proposed by Dennett (1987) assumes that what underlies mentalizing is a set of rational principles that the mind-reader uses to determine which mental state would be adopted by the others, seen as rational agents.

These positions stand in contrast with the so-called Simulation Theory (ST), which, in its most popular and basic version, holds that in order to make sense of the behaviours and beliefs of other

individuals, an agent ascribes to others mental states by simulating them internally in his/her cognitive system (Gordon, 1986; Currie & Ravenscroft, 2002; Gallese & Goldman, 1998).

Another kind of simulation-like approach is the so-called Embodied Simulation Theory (ES). The theory holds that the basic skills of social cognition (understanding others' sensations and emotions) don't require any kind of folk psychology (Gallese, 2001; 2005). This position has been interpreted as a low-level form of mental simulation (Goldman, 2006), based on the unmediated - thus below the threshold of consciousness - processes underlying mirror-neuron activity (Gallese & Lakoff, 2005 p. 5, see also: Gallese et al., 1996; Rizzolatti et al., 1996). This is in contrast to a high-level activity, associated with the attribution of complex mental states (e.g. propositional attitudes), accessible to consciousness (Goldman, 2006, p. 147). The enactive approach, however, rejects all these interpretations. Rather it tries to explain intersubjectivity in terms of *interaction theory* (Gallagher, 2012) or *participatory sense-making* (De Jaegher, 2006). The main point here is that enactivism aims to provide a biologically plausible investigation of intersubjectivity not in terms of inner mechanisms (building a theory or simulating) but, rather, considering the active processes of mutual interaction among different agents.

These three dimensions of embodiment, as clearly emerges, are all radically intermixed, rejecting not only any form of representationalism, but also every form of inner processing - intended as *sufficient* for cognition. A standard argument for proponents of the internalist perspective, in fact, posits that the brain is the cause of conscious experience, intentionality, or, in one word, mind. As Noë (2009) observes, in contrast, we have no empirical evidence that our experience depends on what happens inside our head. We live in (or as) a body and these two categories (brain and body) are always in a deep, indissoluble, relation. However, this relation by itself is not sufficient to provide a biologically plausible notion of cognition. Cognition is always both embodied and situated and it follows that inner processes are only part of the story. As Di Paolo argues,

Cognition is sense-making in interaction: the regulation of coupling with respect to norms established by the self-constituted identity that gives rise to such regulation in order to conserve itself. This identity may be that of the living organism, but also other identities based on other forms of organizationally closed networks of processes, such as sociolinguistic selves, organized bundles of habits, etc. Some of these identities are already constituted by processes that extend beyond the skull. But in any case, cognition is always a process that occurs in a relational domain. Unlike many other processes (e.g. getting wet in the rain) its cognitive character is given normatively and asymmetrically by the self-constituted identity that seeks to preserve its mode of life in such engagements. As relational in this strict sense, cognition has no location. It simply makes no sense to point to chunks of matter and space and speak of containment within a cognitive system. Inspect a baby all you want and you'll never find out whether she's a twin (Di Paolo, 2009, p. 19).

An argument usually employed in modern neuroscience, to defend the opposite thesis (that is, inner processes *are sufficient* to generate conscious experience) is based on oneiric experience:

dreams, as it is argued, are the paradigmatic example of how the brain is able to *create* (i.e. represent) a world. In normal everyday life, brains use mental representations to make us understand and interact with the environment but during sleep they generate dreams. In both cases it is a disembodied activity. I have already discussed the brain-in-a-vat thought experiment to claim that a brain cannot generate any kind of meaningful experience by itself. *Dreaming*, in contrast, can be seen as an example of how an agent, without any actual dynamic interplay with its own body and with the world, could actually have an experience that solely depends on what is going on inside her head. However, a deeper analysis of these issues (Noë, 2009; Rockwell, 2005) shows that the experience of dreaming cannot be similar to the experience that we have when we are awake. In dreaming, perception is not as stable as it is in reality. If you are reading a newspaper in a dream and - still while dreaming - you look somewhere else before looking again at the words, the text changes. As Noë (2009) observes, this happens because in my normal everyday activity I do not have to exercise any complex skill to make sense of a detail. Details, the words of the article I am reading, *are simply there*. The objects of the world are not, in reality, a construct of our creative imagination, as happens during dreams. When I am awake I do not need any effort to stabilize the details. Then, why in dreams do the details show up in a different way? Probably, as Noë concludes, it is because when we dream we are detached from the world in which we are usually embedded. The nature of the dream-like experience is therefore vitiated precisely by the fact that *the world* is absent. And with no world, there is no relevant experiential content.

We do not have any sound argumentation to maintain that only the brain is what allows conscious experience. All this argument can show, perhaps, is that during sleep a subclass of experience is generated by what happens inside us. Once again the Cartesian view that imprisons cognition inside the skin, or the skull, lacks of any valid candidate to corroborate its working hypothesis. The brain, only the brain by itself, cannot be sufficient to generate a conscious experience. Of course, there are different variables in this equation. For example, as in traditional embodiment, brain processes are seen as off-loaded into the body, thus beyond the boundaries of the skull. However, this move is not sufficient by itself in order to provide a more realistic picture of cognition. In both cases the bonds between agents and environment are well defined and maintained, thus reflecting the separation between inner and outer that enaction explicitly refuses. This division, when not explicitly overcome, is risky. Assuming that cognition is an inner property of the animal, can indeed lead to positions advocating an identity between brain and mind. As Rockwell (2005) points out, this idea was so common in William James' times that many scientists believed that 'thought was phosphorus', given the large amount of this element in the brain. In contrast, considering mind and consciousness as functional, rather than purely biological categories, allows a different interpretation. Cognition is part of a system, and not an inner property of a living creature. The next section, thus, will focus on this interpretation of embodiment through inner processing, introducing and exploring the conceptual geography that relates to Mirror and Canonical Neurons.

3.5 - The Human Mirror Mechanism

The discovery of *canonical* neurons, which fire when someone observes, without performing any movement, objects whose size and shape is congruent with the type of hand shape coded by the

neuron (Rizzolatti et al., 1988; Rizzolatti & Sinigaglia, 2008), and *mirror* neurons, which become active both when performing a motor action and when observing or hearing a similar action made by another individual (Di Pellegrino et al., 1992; Sakata et al., 1995; Gallese et al., 1996; Rizzolatti et al., 1996; Kohler et al., 2002) has represented a turning point in cognitive neuroscience and beyond. Thanks to this discovery, many ideas at the basis of the classic cognitivist position have been subjected to a gradual renewal on the basis of empirical evidence and not only through theoretical arguments.

During the 1980s, in Italy, the team of neuroscientists led by Giacomo Rizzolatti was studying a particular area of the macaque monkey's brain, the so-called Area F5. This area, as Iacoboni (2009) explains, is located in the ventral premotor cortex, and deals with planning, selecting and executing actions (ib, p. 9). In particular, the neurons of F5 code particular motor actions, such as grasping, tearing or bringing objects-food-to the mouth. The aim of Parma lab's investigation was basically to understand the neurophysiological mechanism at the basis of the hand's motor behaviour of the monkey. A better understanding of these mechanisms, in fact, would have been beneficial in order to improve rehabilitation paradigms for humans who, after a brain damage, lost some degree of control of their hand.

More generally, however, their investigation also aimed to overcome the traditional view of the pre-motor cortex' Brodmann's area 6. Traditionally, this area was conceived as anatomically uniform with the precise role of planning and executing actions (Rizzolatti & Gentilucci, 1988). However, 'as opposed to this traditional picture they had found by way of research on macaque monkeys that there were very interesting differences in anatomy, connectivity as well as functional properties between the medial areas now labelled F2 & F3 and the areas F4 and F5 in the lateral part of premotor cortex' (Brincker, 2010, p. 19).

Through laborious experimentation, the Rizzolatti team had acquired an impressive understanding of the actions of these motor cells during various "grasping" exercises with the monkeys. (They are called motor cells because they are the first in the sequence that controls the muscles that move the body.) Then one day, about twenty years ago, the neurophysiologist Vittorio Gallese was moving around the lab during a lull in the day's experiment. A monkey was sitting quietly in the chair, waiting for her next assignment. Suddenly, just as Vittorio reached for something-he does not remember what-he heard a burst of activity from the computer that was connected to the electrodes that had been surgically implanted in the monkey's brain. To the inexperienced ear, this activity would have sounded like static; to the ear of an expert neuroscientist, it signalled a discharge from the pertinent cell in area F5. Vittorio immediately thought the reaction was strange. The monkey was just sitting quietly, not intending to grasp anything, yet this neuron affiliated with the grasping action had fired nevertheless. Or so goes one story about the first recorded observation of a mirror neuron. Another involves one of Vittorio's colleagues, Leo Fogassi, who picked up a peanut and triggered an excited response in F5. Yet another credits Vittorio Gallese and some ice cream. There are others, all plausible, none confirmed. (Iacoboni, 2009, pp. 10-11).

More or less, this was how mirror neurons were discovered in the monkey. These neurons, in fact, fired not only when the monkey was moving its hand but also when the animal was simply watching another agent performing a task that employed the same action. In addition to these *mirror neurons*, another set of neurons was dubbed *canonical neurons*, as they discharge when the monkey only observes an *object* without performing any movement, as well as when it grasps that object. The selectivity of these neurons, appreciable through the congruence between the codified motor features and objects' visual properties, argues for their pivotal role in the process of transforming the visual information of objects into the appropriate motor acts (Jeannerod et al., 1995).

Brain imaging techniques have revealed that a similar neural activity is present also in humans. Fadiga and colleagues (1995) used TMS to show that motor evoked potential of the human motor cortex were increased during simple observation tasks, reflecting the muscle activity relevant for the actual performance of the observed motor behaviour. These findings, therefore, suggest that in humans there is a mechanism that is homologous to the one previously found in monkeys. In particular, the areas associated with mirroring processes are the inferior frontal gyrus, the lower part of the precentral gyrus and the temporal, occipital and parietal visual areas (Rizzolatti & Craighero, 2004). A monkey or a human, therefore, while observing some other individual doing a grasping action, basically recruits the same motor plans in its brain that would be necessary to perform the same action themselves (*mirror neurons*). Similarly, the same neural activation occurs whenever a monkey or a human simply *observe* an object (*canonical neurons*).

The only difference between humans and monkeys, with regards to mirror-like activities, is that in humans there is also activation when observing a meaningless movement - like pantomime, waving the hands, raising the arm and so forth (Gallese et al., 1996; Umiltà et al., 2001). This can be explained by considering the fact that mirror neurons do not fire when observing a motor behaviour that is not part of the motor vocabulary of the observer. Monkeys, in fact do not pantomime (Iacoboni, 2009). The idea of a *motor vocabulary* is therefore fundamental, as it represents one of the key concepts related to mirror neurons activation. In Rizzolatti and Luppino's words:

This motor vocabulary is constituted of "words", each of which is represented by a population of F5 neurons. Some words code the general goal of an action (e.g. grasping, holding). Others code how, within a general goal, a specific action must be executed. These words select specific "motor prototypes" such as, for example, the configuration of fingers necessary for the precision grip. Finally, other words specify the temporal aspects of the action to be executed (e.g. opening of the hand). Thus, each action is represented by specific populations of neurons at different degrees of abstraction (Rizzolatti & Luppino, 2001, p. 891).

An elegant fMRI experiment carried out by Buccino and coworkers (2004) can provide further clarification with regards to this point. In their experimental setup, the team of researchers asked a number of participants to watch a silent video, where a man, a monkey and a dog were performing (i) ingestive and (ii) communicative actions. In the first condition, the action was *biting food* and all three animals were doing the same. In the second condition, however, the motor behaviours

associated with communication are very different from one species to another. Therefore, the actions employed for the stimuli were motor behaviours relevant to each species such as *talking* (humans), *lip smacking* (monkeys) and *barking* (dogs). The results of the experiment showed that for condition (i), there was a clear overlapping of the brain areas that became active during the observation of the videos. Indeed ‘watching the three videos produced the activation of two sites (a rostral and a caudal) in the inferior parietal lobule as well as the posterior part of the inferior frontal gyrus and the adjacent precentral gyrus’ (Rizzolatti & Sinigaglia, 2008, p. 132). On the other hand, the results of the second condition were significantly different. The mirror-like activation, in fact, was much weaker when the participants were observing acts like lips smacking and disappeared when watching the dog barking.

The mirror mechanism, it can thus be argued, facilitates the understanding of a goal-directed action that is present in the motor repertoire of the subject without involving sophisticated mental representations, providing the subject with a particular, intrinsically motor, modality to make sense of the witnessed action.

This idea is best understood thinking that the way in which a subject understands the meaning of barking is different from the way in which she understands, say, grasping a mug. This “primarily motor” modality to understand actions and intentions reflects quite clearly the phenomenological importance of the body for sense-making. As Merleau-Ponty’s analysis pointed out in the previous section, agents are immersed in the dynamics of action, and the bodily openness to the world is what allows us to ascribe meaning, to understand the others’ behaviours, emotions and intentions. Another property of mirror neurons is their sensitivity for action goal-directedness. Many neurons with mirror-properties are elicited not by the precise movement performed by another individual but, rather, by the goal of the given action. What matters for these premotor neurons, therefore, is not the mere kinetic property of the motor acts, such as contractions of single groups of muscles (see Rizzolatti et al., 1988). ‘Virtually all mirror neurons display this behaviour: they respond only if the action is directed at an object. They therefore respond to the action of grasping and not simply the hand movements involved in grasping’ (Keysers, 2007, p. 3).

Mirror neurons are in fact insensitive to different kinematic strategies and to the different agents involved in the execution of the action. These neurons code the *goal* of the action and not the movements necessary to perform it. As long as the goal of a given action is different, these neurons are selective enough to distinguish between different kinematic schemas (see Fogassi et al., 2005). In the brain, therefore *perception* and *action* are not separated entities, somehow encapsulated in autonomous, independent, aprioristically defined, brain modules. The relationship between perception and action can be intended more as *perception-for-action*. As Rizzolatti and Sinigaglia put it

The same rigid boundary between perceptual, cognitive and motor processes ends up being largely artificial: not only perception is immersed in the dynamics of the action, being more articulate and complex than previously thought, but the brain that acts is primarily a brain that understands. This is [...] a pragmatic, pre-conceptual and pre-linguistic, understanding, and yet no less important, since it rests on many of our much-celebrated cognitive abilities. (Rizzolatti & Sinigaglia, 2008, p.3).

Mirror neurons, as previously said, are different from canonical neurons: the latter, in fact, discharge simply at the sight of a given object, being influenced by the size and shape of it. Mirror neurons firing, on the contrary, relies on the observation of goal-directed actions involving a body part (hand or mouth)-object interaction and it is uninfluenced by the spatial location of the witnessed motor behaviour (ibid.). The distance of the given object from the subjective space, on the contrary, modulates canonical neurons. The following two quotes encapsulate the standard interpretation of how the mirror mechanism is usually intended and summarize the discussion so far:

The Mirror Mechanism, given the present state of knowledge, maps the sensory representation of the action, emotion or sensation of another onto the perceiver's own motor, visceromotor or somatosensory representation of that action, emotion or sensation. This mapping enables one to perceive the action, emotion or sensation of another as if she were performing that action or experiencing that emotion or sensation herself (Gallese & Sinigaglia, 2011, p. 512).

In virtue of the translation of others' bodily movements into something that the observer is able to grasp as being part of a given motor act accomplished with a given motor intention, the observer is immediately tuned in with the witnessed motor behaviour of others. This enables the observer to understand others' motor goals and motor intentions in terms of her/his own motor goals and motor intentions (Gallese et al., 2011, p. 370).

Also emotions, in fact are part of the story. They are strictly connected to the motor behaviour and action understanding mechanisms. The sensorimotor patterns that underlie our engagement with the world are not executed in a mechanical way, but incorporate some kind of *emotional colouring* (Sinigaglia & Sparaci, 2010), provoking a response of fear, disgust, interest, and so on, thus representing an essential feature for consciousness and intersubjectivity. In light of these considerations, it comes as no surprise that this discovery has been incredibly relevant for further investigations in cognitive neuroscience, and open to discussion in other fields, such as psychology and philosophy.

For decades one of the most common assumptions in neurological studies was to describe the motor areas of the cerebral cortex in terms of a tool evolutionarily designed for merely *executive* processes. Thus, movements and actions were considered ascribable to the classic scheme: perception → cognition → movement (Rizzolatti & Sinigaglia, 2008). This simple picture is coherent with the idea that sensations, perceptions and actions, being distinct and hierarchically organized psychological functions, should be located in different cortical areas, where the stream of information would proceed from a *brain that knows* to a *brain that does* (Boria, 2009, p. 32). According to Gallese, indeed, if we follow the traditional cognitivist model for the analysis of perception and action we would find that:

sensations would prevail in the primary sensory areas, and perception would be the product of primarily temporal-parietal, associative areas, while movements would be controlled by motor and premotor areas located in the posterior portion of the frontal lobe, also known as agranular frontal cortex. The analysis of the external world would be configured as a unidirectional stream of information which proceeds from the (associative and sensory) posterior cortical areas to the frontal motor areas where they would integrate with the prefrontal cortex' elaboration product, location of the decisional processes and, more generally, of the more sophisticated aspects of our intelligence. Experimental data acquired during the last twenty years, however, show us a completely different scenario. The motor cortex of the frontal lobe [...] is constituted by a mosaic of distinct anatomo-functional areas, which relate to each other forming distinct cortico-cortical circuits [...]. Each of those parieto-premotor circuits integrates motor and sensorial information related to a particular body area ensuring its control within the distinct systems of spatial and reference coordinates (Gallese, 2010, *my translation*).

The previous quote also explains why the Parma team was so interested in the re-definition of the anatomical description of the premotor cortex. Their data suggest in fact that the agranular frontal cortex and the posterior parietal cortex, however interconnected, present a strong anatomical and functional distinction, creating parallel circuits that integrate sensorimotor information with regards to a specific effector. The same holds true for the prefrontal and the cingulate cortices, which are concerned with planning the appropriate way to act (Rizzolatti & Sinigaglia 2008). Brincker is quite clear on this point:

This project of reinterpreting the organization of the cortical motor system might seem to be simply an issue of anatomical labelling, but it represents a significant split with the overall model of the mind as an input-cognition-output system [...]. Importantly they started to conceptualize the functional divisions of premotor areas not only according to effectors, i.e. around different body parts of an internally represented homunculus, like Penfield and colleagues thought, but rather as depending on more abstract categorizations of various *kinds of actions* and *goals*. Further, based on the systematic differences in sensorimotor connectivity within the large heterogeneous frontal motor area, they argued that it was *functionally* much less uniform than formerly thought and that many areas showed a level of cognitive complexity well beyond the kinetic movement commands, which traditionally had been seen as the only legitimate functional role of motor areas (Brincker, 2010, p. 19).

But a new model of the motor system implies not only a radical separation from the classic cognitivist position but also a redefinition of many assumptions at the basis of physiology and neuroscience: indeed, as reported by Boria (2009) the evidence that sensory and motor information are ascribable to a common format codified by specific parieto-frontal circuits (Gregoriou et al., 2006) suggests that, beyond the organization of our motor behaviours (Rizzolatti et al. 1997) some

other processes commonly considered high-level like space perception (Sakata et al. 1997, Colby 1998), action understanding (Gallese et al., 2002; Rizzolatti & Matelli, 2003), and others' motor intentions predictions (Fogassi et al., 2005; Fogassi & Luppino, 2005) may have their neural substrate in the motor system (Rizzolatti & Sinigaglia, 2008).

A well-known study by Kohler and colleagues (2002), moreover, showed that a subclass of F5 mirror neurons fire not only when observing or performing a given action, but also when *hearing* the *related sound* of the action itself. These neurons become active not only when a monkey witnesses the observer carrying out a sound producing motor behaviour, but also when the monkey only hears the sound without seeing the action¹³. A goal-directed action, therefore, can be understood independently from the format of the sensory information. The motor acts employed in the experimental paradigm were *breaking a peanut* and *ripping apart a sheet of paper*, acts that are relevant for the monkey, being present in their motor knowledge.

The conclusion therefore is that when the monkey was only hearing the sound of a given action, it automatically activated in its brain the motor plan necessary to perform the very same action. Since mirror neurons encode the witnessed or heard action along with its goal, irrespective of the agent who performs the action, it has been assumed they play a crucial role for action understanding through a mechanism of *embodied simulation* (e.g. Gallese, 2005). Namely, an agent can automatically *re-enact*, with her own cognitive system, the witnessed or heard action by recruiting the same neural population, “as if” (Damasio, 2003) she was performing the action herself.

[W]hen we observe goal-related behaviours [...] specific sectors of our pre-motor cortex become active. These cortical sectors are those same sectors that are active when we actually perform the same actions. In other words, when we observe actions performed by other individuals our motor system “resonates” along with that of the observed agent (Gallese, 2001, p. 38).

This “resonance” would give rise to an inner *representation* of the witnessed or heard action (Gallese et al., 1996). However, as Gallese (2009) points out, the term *representation* is employed in a very different sense from the standard cognitivist tradition. Its content, rather than being solipsistically generated by symbol manipulations, stems from the relationship that our ‘situated and interacting brain-body instantiates with the world’ (ibid., p. 524). This move allows Gallese, as well as other proponents of this view, to reject most of the criticism that the standard theory of embodied simulation and action understanding has been facing so far (i.e. Jacob, 2008). As Dan Hutto puts it, many “sceptics” assume in their arguments that

action understanding necessarily involves making mentalistic attributions of contentful attitudes of some kind - i.e. beliefs, desires or their analogues. It is then argued that mirror neuron activity does not - by itself - suffice for *any kind* of action understanding because such activity lacks some or other aspect that is required for attributing

¹³ ‘As a control, the monkeys were also tested for white noise and other sounds unrelated to the actions. The control sounds were used to rule out the possibility that mirror neuron responses to action sounds were simply due to the arousing, nonspecific effect of any sound’ (Iacoboni 2009, p. 35).

contentful mental states. Thus it is concluded that mirror neuron activity - in lacking the relevant features - necessarily falls short of what is required for action understanding proper (Hutto, 2012, p. 4).

However, this criticism is valid only if one accepts that explicit mindreading (like TT, RTT or ST) is the only possibility to achieve intentional understanding (Gallese et al., 2009, p. 109). Mirror neuron theorists, on the contrary, assume that the kind of understanding embodied by mirroring activity is pre-reflexive and therefore not attributable to deliberate mentalizing. But how exactly could this mechanism instantiate a representation of goal directed actions into one's motor repertoire? Is this mechanism *sufficient* for understanding the meaning of the actions? And is it really a form of *simulation*? Before answering these crucial questions, however, it is important to notice that this theoretical framework has been extensively applied to the realm of music (see D'Ausilio, 2009, for a review).

The next sections, therefore, will offer a brief overview of the most relevant studies of this issue, showing that the notion of simulation, when applied in musical contexts, presents some difficulties in fully explaining the mutuality of musical sense-making. Another aim of the following paragraph is also to provide an analysis of the closest position in musical research to Gallese's ES, namely, the SAME model proposed by Katie Overy and Istvan Molnar-Szacaks (2006; 2009). This latter model, as I see it, is a promising starting point to develop a research paradigm that takes seriously the notions of embodiment and mirror neurons, combining theoretical and empirical evidence.

3.6 - Mirrors in music

Many studies in the last few years have focused on the neural aspects of the sensory-motor integration embodied by the mirror mechanism, stressing in particular the cross-modal plasticity of the motor cortex through the development of musical expertise: among others, Bangert and colleagues (2006) showed with fMRI an activation of the left premotor regions during passive listening tasks for musicians, compared to non-musicians, implicitly suggesting that a *musical vocabulary of acts* could develop through musical training, underlying our musical understanding (for a TMS study see D'Ausilio et al., 2006).

Lahav and collaborators (2007) explored the brain areas recruited when musically naïve subjects listened to sounds associated with sequences of actions they learned during a prerecording training period, finding that 'music one knows how to play (even if only recently learned) may be strongly associated with the corresponding elements of the individual's motor repertoire and might activate an audio-motor network in the human brain' (ibid., p. 309). Haslinger and colleagues (2005), moreover, found that the simple observation of meaningful musical acts elicits a stronger activation for musicians in the fronto-parietal-temporal network when compared to non-musicians, reflecting 'the operation of a mirror-matching system' (ibid., p. 289). A stronger activity for the musicians' primary motor cortex has also been reported during a passive music listening task in a study by Haueisen and Knosche (2001), suggesting how the auditory-motor mapping implemented by the mirror-like activities requires an appropriate repertoire of acts to be successfully fulfilled.

From this empirical evidence it thus might be inferred that the development of the ability to play an instrument would build up in the performer's motor system a vocabulary of (musical-directed) motor acts that can be recruited to simulate the actions, emotions and the intentions evoked by a musical piece played by another individual, thus allowing the agent to ascribe a motor-based intentional meaning to the musical object. This is however, a matter of *degree* of understanding. Some layers of the sense's stratifications of an auditory feedback, in fact, can obviously be grasped also by individuals who (i) don't have any musical training or (ii) don't have that particular chain of acts in their motor repertoire (i.e. although being musicians, they don't know how to play that particular piece, musical instrument or phrase). Indeed, the mechanism of Embodied Simulation (ES), as described above shapes the degree of the agent's *pragmatic* knowledge of the piece, but it does not play an *extensive* explanatory role in the way a listener meaningfully interacts with all the features of a given musical event. Music theory, for example provides a listener with a particular knowledge to ascribe meaning to a musical object, yet (i) it is not the only tool that allows sense-making and, above all (ii) it is not the primal. As we previously stated, ES refers to a *basic* and *minimal* form of (action) understanding, which regulates the pre-conceptual responses to the musical stimulus according to the listener's motor expertise, providing her with a different, intrinsically motor, modality to relate to the musical object.

However, in most literature on music cognition, this standpoint remains not sufficiently addressed, leading in turn to problems in drawing a coherent phenomenological description of such processes. The following two examples provide a more detailed idea of how simulation and embodiment are applied in musical contexts and how it can play a key role for musical understanding. On the other hand, they present some intrinsic problems when compared to the version of ES before mentioned, which can be seen as more *parsimonious* in spirit.

1) Tom Cochrane (2010) provided what he dubbed the simulation theory of musical expressivity, in which 'music is seen as hijacking the simulation mechanism of the brain' (ibid., p. 20). Drawing on the work of Damasio (2003), Cochrane suggested that musical empathy is grounded in the recognition of others' emotions by perceiving their expressive behaviour. In particular, Cochrane's argument is based on three main steps. At the beginning of the causal process, he suggests the (i) triggering of a brain's emotion detecting simulation mechanisms, which - and here is the important point for our discussion - is done either by belief or imagination of the agency generating the sound. Next, (ii) the intermodal connection between sound and bodily movements is utilized, leading to (iii) the mirroring of these movements from a first person perspective, which elicits a simulation of emotions in the listener. Cochrane's contribution is certainly fascinating, having the advantage to unify the listener's sense-making abilities, where memory, imagination and other cognitive functions can be integrated in a motor-grounded framework. However, a few aspects of his model might require a different interpretation: firstly, it relies on an autonomous domain to simulate emotions, which according to ES is not necessary, and prone to circularity in the context of embodied approaches to sense-making where imagination is conceived as an instantiation of ES among others (Gallese, 2011). Furthermore, the entire process of musical simulation is described by Cochrane more as a matter of mental states rather than as a motor-grounded phenomenon, as he posits that, since a musical stimulus was "deliberately constructed by a human being", we proceed to 'interpret that work as the product of certain mental states, and derive the nature of those mental states from the characteristics of the work' (Cochrane, 2010, p. 19).

2) Secondly, Rolf Inge Godøy (2003) proposed that sound-producing actions (hitting, stroking or blowing, etc.) largely influence the formation of our images of musical sounds. Godøy posits that we mentally imitate sound-producing actions when we listen attentively to music, or that we may ‘imagine actively tracing or drawing the contours of the music as it unfolds’ (ibid.). Following the trends in cross-modal research (Calvert et al., 1998), he hypothesizes that the motor-mimetic element translates ‘musical sound to visual images by a simulation of sound producing actions [...] forming motor programs that re-code and help store musical sounds in our minds’ (Godøy, 2003, p. 318). We can understand Godøy's framework as an attempt to integrate ES theory into music studies. However, his proposal seems unfit to radicalize the claims of ES in the context of musicality: the action-perception cycle, as portrayed by him, is meant as a ‘feedback loop of an incessant process of top-down hypothesis-generation followed by bottom-up driven comparison with what we assumed in our hypothesis, successively adjusting and refining our top-down generated hypothesis by each period of the perception-action cycle’ (Godøy, 2009, p.212). His account, coherent with the so-called control theory, contrasts with two of the most peculiar features of ES, namely the intra-personal character of the resemblance primarily involved in the simulation process (Gallese & Sinigaglia, 2011, p. 513), and the absence of propositional aspects from the core mechanism of action understanding (ibid.). Despite embodied and simulationist approaches being still relatively new to the field of music research, during the last decade they have both become increasingly influential (see Leman, 2007). The speed of this diffusion may account for some peculiarities that seem to distinguish most of theoretically elaborated ES theories of musical understanding from the general field of embodied and enactive paradigms. While embodiment and, more specifically, embodied simulation, still represent a controversial topic in philosophy of mind (Clark, 1997), cognitive neuroscience (Gallese & Sinigaglia, 2011), as well as other domains, we find just a weak echo of this discussion in music-related studies. As a consequence, the diluted versions of ES as it appears in current debates are, *de facto*, unsuitable for providing the much-needed reconceptualization of musical understanding. Such an achievement would require at least a closer attention towards some basic theoretical issues implicit in the adoption of an ES theory, even before considering the problems in applying this framework to musical understanding. In particular, the positions we briefly analysed seem to reject the basic statements of ES. That is:

- ES does not involve any kind of mental states/cognitive involvement.
- The acts mirrored in ES are goal-directed acts within the motor repertoire of the perceiving subject.
- The resemblance on which ES relies should be characterised as intra-personal, as the perceiving subject doesn't have direct access to the other's mental states or acts

As briefly illustrated, the two contributions presented instead of moving from these preliminary points, seem to take for granted options and interpretations that not only should be discussed, but are often in contrast with classic ES theories, introducing additional structures and processes that are not required by the standard version. As I see it, in fact, when ES is coherently implemented in a musical context, it might present a more parsimonious model than other simulation-like theories for defining the complex mutuality between an agent and the auditory feedback. However, as will

emerge in the next chapter, it carries some problematic assumptions. Moreover it might fail to provide a model for music cognition that extends beyond the boundaries of skull and skin. Anyway, a better understanding of the principles of ES - and of the mirror neurons in general - is shown by the seminal publications of Katie Overy and Istvan Molnar-Szakacs (2006; 2009). The two scholars were the first to apply systematically this new neuro-cognitive paradigm to music research, developing the SAME (Shared Affective Motion Experience) model, suggesting that ‘musical sound is perceived not only in terms of the auditory signal, but also in terms of the intentional, hierarchically organized sequences of expressive motor acts behind the signal’ (Overy & Molnar-Szakacs, 2009, p. 492). From this perspective, the listener (according to his or her motor expertise) is able to extract different levels of motor information:

- The *intention* level, which defines the long-term goal of a given action.
- The *goal* level, which describes the basic goals that lead to the achievement of long-term intentions.
- The *kinematic* level, which deals with the space movements of the body.
- The *muscle* level, which comprehends the pattern of muscular activity required for the actual execution.

This means that the *degree of comprehension* of a musical object is shaped by the *degree of practical musical expertise* of the listener, *before* and *below* any kind of cognitive, or high level, subordination. As Mark Reybrouck, states, in fact, ‘sounds [...] are the outcomes of human actions. Even if they are not self-produced, they can induce a kind of (ideo)motor resonance that prompts the listener to experience the sounds as if they have been involved in their production.’ (Reybrouck, 2005a, p. 3). A naïve musical subject will be able to extract information only with regards to *goals* and *intentions*, without being able to gain an insight into the detailed motor behaviour that allow a particular performance, as she does not have it in her motor repertoire. On the other hand, an expert musician, when listening to a piece that she has been rehearsing for a long time, is able to constitute the musical object in much more detailed terms, as she integrates her practical motor knowledge to *understand* (pre-reflectively) the goals and the intentions of the heard action. The authors put it as follows:

For example, at one extreme, a professional musician listening to music which they know how to perform (e.g., a saxophonist listening to a saxophone piece they know well) is able to access precise information at all levels of the hierarchy, from imagined emotional intentions to specific finger movements and embouchure. At the other extreme, a musical novice listening to unfamiliar music from an unknown sound source (e.g., someone who has no knowledge of the existence of saxophones) is not able to access precise information at any level, but may feel the beat, sub-vocalise, and interpret emotional intention accordingly (e.g., fast, loud, and high in pitch might be considered emotionally charged). Thus, the resonance or simulation mechanism implemented by the human MNS [Mirror Neurons System] matching perceived and

executed actions allows a listener to reconstruct various elements of a piece of music in their own mind (bringing together auditory, motion, and emotion information), and the richness of that reconstruction depends on the individual's musical experience (Overy & Molnar Szacacs, 2009, p. 493).

The authors use the word "mind" in a different (internalistic) sense from what I intended it during this work. Fair enough. What matters here is that this model matches quite strongly with Gallese's ES theory and it allows the proponents to explain the nature of musical understanding from an embodied standpoint. We will consider in the next chapters in which terms the *enactive perspective* can enrich this model but, for now, we can notice that the SAME model represents a huge step forward for the redefinition of music cognition.

According to this model, the goal directedness of the musical acts mirrored by the MNS are necessary for the constitution of a basic musical meaning and essential for the study of the (inter)active engagement that links musical agents (listener or performer) and musical objects (an auditory feedback experienced as musical). This relation, that I previously named *musical intentionality*, in its motor-based characterization, implies that:

- Music is something intentionally constituted by the way an agent intends it, through her dynamic interplay with the musical environment¹⁴
- Music is something that we do: we engage in music by moving, dancing, playing, composing, and listening. Listening is also, in fact, a skilful bodily activity (see Noë, 2004).

As previously shown, the shared intimacy between action and perception lies in the assumption that perceptual experience consists in the implicit knowledge of the sensorimotor contingencies and results embodied by the recruitment of the same neural population that a subject - according to her motor expertise - employs when performing as well as listening to the related sound(s) of the same (chain of) goal-directed action(s) (Kohler et al., 2002). This idea has also implications for to the nature of musical ontogeny, the development of musical expertise, the perception-action coordination in collective music making and the emotional response to music - given that the auditory mirror-like properties seem to be valid for a wide range of functions which can elicit very different behaviours (D'Ausilio, 2007). These points will be examined in greater detail in Chapters 5 and 6. For now, instead, we can conclude this chapter stating that from a neurophenomenological standpoint it is our ability to be directed towards the musical object and to constitute it that makes the experience of music possible, and not *viceversa*. However this seems to be a too radical conclusion. The next chapter will therefore explore another possible solution, still employing the

¹⁴ This point will be further clarified and implemented in the following chapter. As emerged, according to classical mirror neurons theories, it seems that this interplay is defined in terms of *simulation*. Rather I will argue for a different interpretation.

crucial evidence carried by the discovery of mirror mechanism, yet not committed to a too internalistic perspective, and not to be identified with a *simulation mechanism* for sense-making.

Chapter 4

Music beyond mirroring¹⁵

An individual perceptual system no more implicitly represents laws determining, or principles governing, transformations of states than the solar system implicitly represents Kepler's or Newton's laws. In both cases, the laws are real. In neither case are they represented by elements of the subject matter governed by the laws

- Tyler Burge, *The origins of objectivity*

The causal relations between nervous systems and environments are intricate and continuous. There is nothing specially oomphy about causal relations inside the skin, or inside the head, nothing specially capable of pushing or shoving. So there is nothing causally mysterious or inhospitable to materialism or naturalism or realism about relational states of persons. And there is no magical causal boundary around persons. Viewed subpersonally, they are in principle transparent to causality

- Susan Hurley, *Consciousness in Action*

4.1 - A wolf in sheep's clothing

In the last chapter I explored the conceptual geography at the basis of embodiment and mirror neurons, introducing and developing the notion of *musical intentionality* as well as highlighting a few studies laying at the intersection of the (mirror-implemented) embodiment thesis and musical experience.

In this chapter, I will continue the discussion of the relevant literature on embodied music cognition through a critical analysis of the book *Embodied Music Cognition and the Mediation Technology* written by Marc Leman (2007). A second - and more important - aim of this section, however, will be drawing a comparison between classical embodied approaches to music experience and the radical interpretation of embodiment posited by *enaction*. I will, therefore, take the ecological notions of affordances and *musical affordances* as a starting point for this second

¹⁵ This chapter contains material previously published as Menin, D., & Schiavio, A. (2012). Rethinking musical affordance. In *AVANT. The journal of the philosophical-interdisciplinary vanguard*, 3(2), pp. 202-215, and as Matyja, J., & Schiavio, A. (2013). Enactive music cognition. Background and research themes. In *Constructivist Foundations*, 8(3), pp. 351-357. Schiavio, A., & Menin, D. (2013). Embodied music cognition and the mediation technology. A critical review. In *Psychology of Music*, 41(6), pp. 804-814. Moreover, it contains material from the following article in preparation: Schiavio, A., & Gerson, S. (in preparation). An enactive perspective on infants' musicality.

part, pushing the non-representationalist perspective on musical experience further, as well as analysing some problems that may arise by employing the concept of (embodied) *simulation*. Finally I will show how alternative views on mirror neurons can implement and perhaps overcome the simulationist claims - with regard to intersubjectivity, shared musical experience and music perception. This move will allow me to radicalize the *embodied* perspective on music cognition, defining the main themes of the *enactive* approach to human musicality through a confrontation with (i) embodied and (ii) externalist views on sense-making and musical experience.

Generally, this chapter will present a full rejection of the concept of mental representations and a serious consideration of the active forms of dynamic interplay between living systems and environment through music-directed sensorimotor patterns. A *full rejection of mental representation* is in fact a necessity, if we are to develop an approach that takes the notions of embodiment and situatedness seriously. As we will see, however, many *embodied* claims with regards to musical experience not only dangerously flirt with fully blown representative objects, but also define as ‘embodied’ something that is - on the contrary - completely *disembodied*, eventually marrying happily the dualistic tradition as well as cognitivist models to study the musical mind.

To understand the roots of the classical perspective on music cognition - where the musical stimulus is conceived in terms of an abstract and unidirectional stream of information encoded and processed by the brain - we can go back as far as the seminal work by Hermann von Helmholtz (1863). He provided one of the first accounts aimed at investigating music perception in terms of mere physiological processing. His work gave a neurophysiological explanation of some of the key aspects of Western musicality, such as perception of consonance, dissonance, harmony and tonality, as well as providing ‘the physiological grounding for gestalt psychology in the first half of the twentieth century, and for the cognitive sciences approach for the second half of the twentieth century’ (Leman, 2007, p. 29).

In a well-known paper, Marc Leman and Albrecht Schneider (1997) described the origins of cognitive and systematic musicology (see also Parncutt 2007), analysing the revival of gestalt and cognitive perspectives. Considering the cognitive perspective, they referred to the study by Allen Newell (1982) as representative of the *symbol system approach*, where ‘propositional representations of music were believed to be a proper starting point for the study of musical cognition’ (Leman & Schneider, 1997, pp. 18-19). The rise of cognitivist musicology was also characterised by a strong link with the research in linguistics by Noam Chomsky (1965) and the computational approach in the field of artificial intelligence (Laske, 1977). For instance, Leman and Schneider examined musical semiotics - the study of the musical signs and their meaning (see Monelle 1992) - from a purely linguistic point of view (Ruwet, 1975) as well as from Charles Sanders Peirce’s theory of signs. These approaches, in league with the classic *modular* account of the mind previously discussed could be summarized in the massive trend in current music psychology that ultimately aims to draw an accurate map of specialized brain areas involved in music processing in order to provide a universal description of musical abilities. The following quote clarifies once again the idea behind this area of research:

As with language, specific areas of the brain seem to be devoted to the processing of music information. If we could grasp universal principles of musical intelligence, we would get an idea of how our music understanding gets refined and adapted to a particular musical style as a result of a developmental process triggered by stimuli of

that musical culture (Purwins et al., 2008, p. 152).

In other words, the localisation of musical brain modules (which would represent “universal principles of musical intelligence”) might lead to a satisfactory explanatory model of musical understanding despite developmental and cultural differences.

The representationalist account is in fact still widespread, even if sometimes it manifests itself in disguise. In a more recent work, for instance, Charles Nussbaum (2007) claimed to have developed an *embodied* approach to musical meaning. This approach assumes that a listener would develop a *mental map* of the bodily actions elicited by the music. The role of the body is therefore merely *passive* because, for Nussbaum, what allows the understanding of music is the listener’s ability to infer non-musical contents and build a representation of them in her/his brain.

Music learning, too, is understood as ‘a process by which mental representations (genuine musical conceptions) are developed and gradually altered, differentiated, extended, and refined’ (Gruhn, 2006, p. 17). This shows that the cognitivist paradigm is still used as an important tool for musical understanding and education, despite the rise of the promising antirepresentationalist frameworks in cognitive science that I have previously described. According to Wayne Bowman, moreover, another (yet strictly related) problem is that

[w]e find ourselves advocating music study for reasons that fit with prevailing ideological assumptions about the nature of knowledge and the aims of schooling, but on which we are ill-equipped to deliver, and that neglect what may be most distinctive about music: its roots in experience and agency, the bodily and the social. Our most revered justifications of music education are built upon deeply flawed notions about mind, cognition, and intelligence (Bowman 2004, p. 33).

However, something has recently changed. During the past few decades, in fact, when trying to explain how we understand, or make sense of the musical material, scholars in musical research have tried to give an increased importance to the role of the body for musical experience. Probably the most famous and quoted work within this context is *Embodied music cognition and mediation technology*, written by Marc Leman (2007). The aim of his book is to clarify the sense-giving process at the basis of musical comprehension, focusing on the *cognitive* relationship - in the broadest sense - that ties musical subjects and objects. His approach assumes that the (musical) mind results from an embodied interaction with music:

The human body can be seen as a biologically designed mediator that transfers physical energy up to a level of action-oriented meanings, to a mental level in which experiences, values, and intentions form the basic components of music signification. The reverse process is also possible: that the human body transfers an idea, or mental representation, into a material or energetic form (Leman, 2007, p. xiii).

This promising and fascinating perspective has been developed according to the previous literature in different disciplines such as psychology, neuroscience, musicology and cognitive science. In the next few paragraphs I will discuss the particular way Leman tackles the problem of musical understanding, as well as its internal and external coherence. Specifically, I will explore some implications and sub-problems that arise from the adoption of an embodied approach to the study of musical experience, focusing on the following four points

- Overcoming of the subjective-objective *dualism*
- The *simulation* issue: what kind of processes could enable musical understanding?
- Musical *intentionality*: how can we define a conceptual topography of musical contexts?
- *Mediation* in musical practice and experience: the role of musical instruments and of the body

4.2 - Mind the gap!

From the outset of the book, Leman identifies one of the crucial problems of music research as explaining how to *fill the gap* between music seen as *encoded physical energy* (the way modern digital media considers music) and music as a matter of *beliefs, intentions, interpretations, evaluations and significations* (the human way of dealing with music) (2007, p. xiii). In his words, the book is based on a hypothesis about musical communication, which is supposed to be rooted in a particular relationship between musical experience (mind) and sound energy (matter). Despite some noble intentions, the way Leman tackles this issue are not beyond question. By postulating an open dichotomy between mind and matter, the author clearly misinterprets the embodied stance, which on the contrary assumes that the mind is already a category that cannot be divided from the body of an agent. Mysteriously, the author further assures us that '[n]euroscience has provided compelling arguments that the Cartesian division between mind and matter can no longer be maintained and that a disembodied mind as such does not exist' (ibid., p. 13).

We can notice a strong contradiction between this point and the problem the book wants to solve. Indeed, how is the body supposed to bridge the gap between mind and matter if they cannot be considered anymore two different categories? In assuming this ontological separation, the author actually follows Descartes' stance, as previously discussed in Chapter 1. Among the main theoretical models proposed to overcome the difficulties raised by the Cartesian dichotomy (how are mental events and physical events linked? How can an immaterial mind interact with a material body?), the classic cognitive perspective, as we saw, has emerged as one of the most discussed and controversial approaches during the past century. It has been developed, however, leaving the dualistic presuppositions unaltered, as in the *bottom up/top down* model for information processing (see Dennett, 1994). Leman's account for 'an action-oriented approach, based on the notion of corporeality' (2007, p. 26) is, in fact, supposed to overcome the problem of dualism, but its aim to provide an 'epistemological foundation for bridging the gap between musical mind and matter' (2007, p. 26) falls inevitably short due to the intrinsic contradiction in its assumptions.

This *dualistic* theoretical background, moreover, will affect most of the topics discussed in the book. After discussing the problems of a subjectivist approach (ibid., pp. 12-13), we are told: '[a]

way to proceed is based on the idea that action may play a key role in mediation processes between the mental and the physical world' (ibid., p. 14). In order to focus on *action*, Leman advocates an intermediate level between subjective and objective methodologies. As Keller and Janata pointed out in their book review, Leman's analyses focus on the levels of descriptions used in musical research:

[p]revious approaches to understanding music (e.g., traditional musicology) have fixated upon these two levels of description without giving adequate treatment to the 'rules' that govern the mapping between objective representations and subjective interpretations. Such rules are needed to achieve his scientific goal of developing a complete theory of music, as well as his practical goal of developing a successful mediation technology. The key to Leman's solution is the proposal that an understanding of musical intentions requires third-person and first-person descriptions to be linked via second-person descriptions, which are corporeal in nature. At this intermediate level, expressive bodily gestures from an individual's repertoire of actions are used to describe moving sonic forms in a manner that the individual can interpret based on his or her personal experience of interacting with others in the world (Keller & Janata, 2009, pp. 289-290).

Leman's approach aims to replace an objectivistic third-person description and a standpoint linked to subjective intentions' interpretations (first-person description) with a second-person description based on the notion of *corporeal intentionality*. The next section will discuss this crucial notion after highlighting Leman's understanding of (embodied) simulation and its role for musical understanding.

4.3 - Simulating and understanding - part I

In Leman's work the concept of *embodiment* has a pivotal role in the meaning-giving processes that tie musical subjects and objects. This notion, however, is strictly connected with a disembodied simulationist claim that has very little to do with a truly embodied kind of sense-making. The author, in fact, basically states that music understanding consists in a *mental simulation* of acts appropriate to make sense of a given musical surface. As briefly outlined in the last chapter, *simulation* is a very extensive issue. Therefore, in order to gain an insight into this claim we would need to explore the conceptual topography of Leman's approach starting from its basic formulation. The first question to be answered is about the characterisation of the *subject* of musical experience, and the identification of other agents that she perceives and understands *via* simulation.

The above findings agree with the notion of self as a mental model having access to a proprioceptive and kinesthetic representation of one's own body in combination with a representation of intended relationships with the other (Gallagher, 2000; Metzinger,

2003). Both aspects of the self, its inward and outward directedness, can be considered aspects of the action-oriented ontology. In the present context, the other is of course the music. The above findings are consistent with the idea that empathy involves regulatory mechanisms by which the subject keeps track of the self in relation to music (Leman, 2007, p. 126).

According to Leman's account, therefore, the subject of musical experience should be identified with a mental model, whereas "the other", the agent, whose behaviour is understood in intentional terms, should be the musical stimulus itself. Of course Leman does not mean musical surface as some sort of subject to whom we can ascribe beliefs, values and meanings. His claim should be rather understood in these terms: what matters for musical understanding to take place is the subject's intentionality; the attribution of intentionality, moreover, relies on some kind of mirroring, that is, 'on the basis of a simulation of the perceived action in the subject's own action' (Leman, 2007, p. 92). Here we have to notice the radical incompatibility of such a position with the ES theory, according to which motor understanding takes place without and before any kind of inferential mediation.

Another crucial issue in defining a simulationist approach concerns the way we mean *simulation* itself. I have previously identified ST (in contrast to TT and RTT, but also to ES) as a claim that holds the following perspective: an agent can understand the behaviour, the intentions, the emotions and the actions of another individual by simulating them internally, within her cognitive system. As Rockwell (2008) pointed out, however, a naïve approach to this topic is no longer sustainable, in light of the long-term controversy between *Theory Theory* and *Simulation Theory*. According to Gordon (1986) and Rockwell (2008), there are two main ways *simulation* has been intended in literature:

- As a kind of pictorial representation. In these terms "simulating" is equated to "creating an image in the mind".
- As a form of pre-tense or hypothetical "acting out". According to this view, any aspect of our mental life can be turned into a simulation by taking it off-line.

Both these approaches have been widely questioned regarding their theoretical presuppositions and empirical evidences, yet, what it seems problematic is the demand to embrace both of them at the same time. According to the first view, indeed, there is no need to postulate any kind of conceptual understanding *preceding* simulation, whereas the second view is known to collapse into a hybrid TT/ST theory. The fact that Leman does not address explicitly the issue of this distinction raises several coherence problems for his account, questioning whether he always means "simulation" in the same terms or not. The following excerpt (in which Leman quotes his colleague Godøy) highlights a view of "simulation" as pictorial representation, where the subjects would create an image in the mind from which to extract information for action planning.

When a trajectory in the inner space is simulated - for example, a walk from the front door of the university building to my office - I can associate this motor image with a trajectory in the outer space. I can explore the visual-audio-tactile features of the objects I encounter along my imaginary walk. In that sense, the sensorimotor couplings allow the transition from imagined movement to predicted sensory qualities. If the moving sonic forms can engage humans in body movements, then it is straightforward to assume that this movement will engender sensory qualities which can be attributed to music as well. If corporeal imitation of movement in sound is possible, then the association with sensory qualities is straightforward (Godøy, 2003, [quoted in Leman, 2007, p. 88]).

A second quotation, however, will show a different - and, *de facto* incompatible - way he means “simulation”, here understood in terms of *imitation*.

Mimesis theory assumes a transitivity relationship: (1) music imitates something, (2) the subject imitates the music, and hence (3) the subject imitates that same something [...]. A related version of this transitivity relationship is based on the notion of expressiveness: (1) music expresses something, (2) the subject captures that expression, and (3) the subject captures the source of that expression (Leman, 2007, p. 128).

This view, exemplified in several excerpts, outlines a circular scenario in which simulation (or mimesis) cannot account by itself for an understanding of actions, concepts or musical objects, requiring therefore, as the ST/TT dispute showed, a conceptual knowledge prior to simulation itself, as we can argue from this well-known argument of Dennett:

If I make believe I am a suspension bridge and wonder what I will do when the wind blows, what “comes to me” in my make-believe state depends on how sophisticated my knowledge is of the physics and engineering of suspension bridges. Why should my making believe I have your beliefs be any different? In both cases, knowledge of the imitated object is needed to drive the make-believe “simulation”, and knowledge must be organized into something rather like a theory (Dennett, 1987, pp. 100-101).

Furthermore, it seems that the emphasis on *attribution of intentionality* reflects a position similar to Dennett’s, where *understanding* is a matter of *concepts*, not *percepts*. As Leman puts it, in fact:

This attribution of intentionality can also be extended to material things that move, such as cars. In traffic, another car is not just a moving object. It is an object with particular intentions, which I can understand by using the experience of my being a driver. Responsible drivers aim at understanding the intended movements of other cars, and on

that basis predict their future behaviour. It is likely that for a dog, a car is just another moving object. It is not an intentional object because the dog is not involved with driving a car. The moving car is not something that the dog can relate to its own actions. Hence, it may not have an understanding of it in terms of a mental simulation. Like cars, music can be understood as an intentional object (Leman, 2007, p. 78).

This view seems to compromise Leman's understanding of the mirror neurons issue. The author, in fact, assumes that intentionality can be mentally attributed to agents, or even to *material things that move*, via simulation. This view, however, highlights not only remarkable internal incoherencies, but also points to a substantially disembodied approach, in which the subject mentally attributes intentional states to musical surfaces. On the contrary, as Sinigaglia (2008a) noticed, mirror-like activities might suggest that the basic aspects of intentional understanding can only be fully appreciated through a *motor approach* to intentionality, without any form of mentalizing. It is probably to this kind of intentionality that Leman is referring when he uses the expression "corporeal intentionality" (2007, p. 84). It might be assumed that the author's mistaken interpretation of embodiment could be forgiven in light of a coherent intentional theory that would explain the primarily motor-based relation that ties musical objects and subjects. But, again, the author inexplicably refers to another kind of intentional understanding, a 'cerebral intentionality, which - in music - explores the speculative pursuit of potential interpretations. The essence of cerebral intentionality is interpreting *the source of intentions* attributed to music. The essence of corporeal intentionality is the articulation of moving sonic forms' (ibid., p. 84 - *emphasis added*). This quote clearly shows that Leman attributes the understanding of the intentions to purely high-level cognitive abilities, rather than to a truly embodied kind of sense-making.

Despite its revolutionary ambitions, Leman's view of simulation shows a clear lack of theoretical grounding, nevertheless portraying the intrinsic problems that stem from the employment of terms such as "simulation" and "understanding". After a closing remark on Leman's work in regards with the notion of *mediation*, I will explore from a different perspective the concept of *embodied simulation*, highlighting different interpretations of this framework, heading towards a more enactive perspective. But let's first conclude our discussion on Leman (ibid.).

4.4 - Mediation, Body, Instruments

According to Keller and Janata (2009), *mediation* refers to the mappings between the intentions and desires on the part of active musical participants and the technology that renders the music. This first meaning, related to the dualistic presuppositions previously examined, is not, however, the only one adopted by Leman: he also refers to mediation with respect to the role of technology itself as mediating between the performer and actual sonic outcomes. I have already discussed the first kind of mediation, concerning the link between the *allegedly* separate domains of mind and matter. Now it is the turn of the second one, wondering what kind of mediation Leman poses between a musician and the music he plays. Simply put, we could also look at the problem from this side: what is the role of musical instruments, whether they are classical acoustic ones or electronic devices, in musical practice and experience? We have at first to notice that Leman is facing two

different issues:

- How mediation technology actually works in musical practice.
- How *mediation technology* should evolve (with particular regard to electronic devices for music making) and how *reproducing activities* should work, in order to be fully *transparent*, giving the illusion of non-mediation.

Though Leman seems not to distinguish between these two sub-problems, he made clear his persuasion that an understanding of the more theoretical one would be useful in order to implement more effective music technologies (2007, p. 2). In order to delineate Leman's approach to this crucial issue, we should look back to his position about the body-mind problem. Once recognized the dualistic bias, which informs his whole theoretical paradigm, we can fully understand the operating concept of "mediation" adopted, though not explicitly discussed. As emerges, his view about musical technologies should be understood as based upon the previously examined dualistic account: musical instruments work as mediation technologies in a similar way to how the body functions 'as a biologically designed mediator that transfers physical energy up to a level of action-oriented meanings, to a mental level in which experiences, values, and intentions form the basic components of music signification' (Leman, 2007, p. xiii).

This is, apparently, a classical embodied claim: instruments, as *tools*, can be seen as extensions of the human body or, more precisely, of the *peripersonal* space defined by the subject's action possibilities (see Costantini et al., 2011). Nevertheless, a closer analysis will show that Leman reduces embodied cognition to a kind of mediation compatible with classical, disembodied, paradigms about tool use, once more misunderstanding the deeper meaning of embodiment. The reason for this misunderstanding is to be found precisely in the dualistic presuppositions previously examined: if mind and body are conceived as distinct substances the main theoretical problem becomes to *justify the mediation between these separate dimensions* (see also Maravita & Iriki, 2004).

This brief analysis aimed to show that Leman's attempt to develop a new approach defining musical experience is not sufficient in providing a coherent theoretical paradigm, for most of his intentions fall short according to the implicit contradictory background underlining his arguments. The explicit use of an ontological distinction between mind and matter is indeed incompatible not only with an embodied approach to musical understanding, but also with a plausible notion of musical intentionality based on subject's motor repertoire. At the basis of Leman's perspective there is, as previously emerged, a misunderstanding of the role of the sensorimotor integration provided by the mirror neuron system: while the mirror mechanism 'maps the sensory representation of the action, emotion or sensation of another onto the perceiver's own motor, visceromotor or somatosensory representation of that action, emotion or sensation' (Gallese & Sinigaglia, 2011) without any cognitive subordination, the author, conceiving the subject as a disembodied mind, prefers to investigate other (high-level) forms of sense-giving abilities, shaped upon a rough notion of *simulation*.

On the contrary, in order to provide a persuasive theory of musical understanding based on embodied cognition, the author should have considered the *goal-directedness* of the motor acts

grasped and mirrored by the subjects (for example by studying their involuntary resonance during passive listening tasks, or measuring their pre-attentive neural response) focusing on the *basic* level of the intentional relationship that links the subject and the object in any musical experience.

4.5 - Simulating and understanding - part II

It might be assumed that embracing the form of embodied and internalistic processing with *bodily formatted* (i.e. sensorimotor) representations posited by ES for explaining the way humans *understand* actions, emotions and intentions of others, is a price well worthy paying in order to overcome the traditional *high-level* forms of intentional understanding (as portrayed by TT, RTT or ST). However, it still remains unclear how the mechanism instantiating such an immediate and automatic comprehension (ES), would work in detail. Furthermore, there are several questions concerning the passage from the actual empirical findings of Mirror Neurons to the development of the theoretical framework based on ES (Gallagher, 2012; Hutto, 2013) that need to be answered. The traditional notion of Embodied Simulation, as described several times by Gallese, Goldman, Sinigaglia and other spokesmen of the Mirroring stance, does not require any sophisticated form of explicit *mentalizing*, providing a parsimonious and incorporated model for understanding the others' actions, emotions and intentions.

As noticed by Gallagher (2012), Gallese uses the following quote from Lipps to clarify the equivalence between perceiving and understanding: 'when I am watching an acrobat walking on a suspended wire, Lipps (1903) notes, *I feel myself inside of him*' (Gallese, 2001, p. 43). As Goldman (2006) recognizes, this low-level, *enacted* form of simulation stands in contrast with the high-level, *pretense*, simulation theory (Gordon, 1986; Currie & Ravenscroft, 2002). The relevance of this paradigm for the enactive approach should rely on its automaticity, which avoids the recruitment of fully blown mental representations. However, this automaticity - as embodied by simulation-like processes - has been recently questioned. As Shaun Gallagher puts it, in fact:

if simulation were as automatic as mirror neurons firing, then it would seem that we would not be able to attribute a state different from our own to someone else. But we often do this in cases where we see someone acting in a way that actually motivates the opposite reaction in us, for example, if I see someone enjoying acting in a way that for me is disgusting (Gallagher, 2012, p. 168).

Let's have a closer look into the main nuances of this implicit form of simulation by firstly focusing once more on the explicit Simulation Theory, in order to define some shared assumptions that might fail to develop a coherent explanatory paradigm. ES, in fact, *is* a form of *simulation*. Standard simulation theory (ST) implies (i) pretence and (ii) some kind of instrumental characterisation (Gallagher, 2012). Indeed, following the analysis provided by Gallagher (ib), we notice that Goldman asserts that ST employs 'pretend states [...] deliberately adopted for the sake of attribution's task [...]. In simulating practical reasoning, the attributor feeds pretend desires and beliefs into her own practical reasoning system' (Goldman, 2002, p. 7). Moreover, classical ST

accounts are ‘characterised in terms of a mechanism or model that we manipulate or control in order to understand something to which we do not have direct access’ (Gallagher, 2012, p. 177). But, as Gallagher concludes, these two conditions are not met by Mirror Neurons, raising several problems in attributing a simulationist account for their activation. Also in the embodied model of simulation, it seems that pretence should play an important role, considering that ‘I put myself “as if” in other person’s shoes’ (ibid., p. 178). My brain does not have any *instrumental control* on mirror neurons’ activation - it cannot decide to make them fire. Rather, being an automatic process, they are elicited by the other person’s actions. ‘It is not us (or our brain) *initiating* simulation; it is the other who does this to us. This is a case of perceptual elicitation rather than executive control’ (ibid.).

It seems therefore that using the term “simulation” for these kinds of sub-personal processes might create enormous theoretical difficulties in the analysis of first- or third-person’s specification, as mirror neurons are neutral with respect to the subject. ‘In that case, it is not possible for them to register my intentions as pretending to be your intentions; there is no “as if” of the sort required by ST because there is no “I” or “you” represented’ (ibid., p. 179). Gallagher and Hutto summarize this interpretation as follows:

Nothing (or no one) is using a model [...] and neuronal processes cannot pretend. As vehicles neurons cannot pretend - they either fire or they don’t. More importantly, in terms of relevant content, if they are neutral with respect to first- and third-person, pretence in just these terms (I pretend to be you) is not possible. In effect, simulation, as defined by ST, is a personal-level concept that cannot be legitimately applied to sub-personal processes (Gallagher & Hutto, 2008, pp. 18-19).

There are also other problems. What happens when, for example, I am witnessing at the same time different actions? Let’s imagine the following scenario. Since I am a guitarist, it seems that while listening to a piece for *guitar orchestra*, I should somehow automatically simulate all the meaningful acts that would constitute the musical stimulus that I am listening to. But how could I possibly do that? The intentions, emotions, actions of the guitarists involved can be extremely different, though directed towards the same musical goal - *the execution of a chord, a motif, a passage or - in general - of the piece*. Of course one could posit (as I have argued myself when discussing music and modularism) that since a musical piece is defined not by the different parts, but rather by their unity, I would experience the *totality* of the different emotions, intentions and acts, therefore simulating the sum of them. Moreover it is also true that the mirror mechanism codes the goal of a given action rather the actual movement. But, again, how could I put myself in the other person’s shoes “as if” I am performing the piece by myself, in this paradigmatic case? There is no “other person” here, but a totality of different goal-directed motor behaviours that are performed by different actors. A possibility is that what I am actually doing in this context is *resonating* with my brain-body system in a way that does not involve any kind of *direct matching* between the witnessed and neurally represented given action.

Rather, instead of simulating, the brain processes embodied by mirror-like activities could be seen as underlying a form of *intersubjective enactive perception* (Gallagher, 2012, p. 181). This move allows getting rid of the notion of neural representation and of neural simulation, without

nevertheless denying that neural processes play a decisive role in musical sense-making. To the enactive eye, in fact, the problem is the supposed sufficiency of such sub-personal processes¹⁶. But how would it be possible to characterise mirror neurons' activation in non-simulationist terms? As Gallagher puts it:

The articulated neuronal processes that include activations of mirror neurons or shared representations may underpin a non-articulated immediate perception of the other person's intentional actions, rather than a distinct process of simulating their intentions. On this view, perception is a temporally dynamic and enactive process (Gallagher, 2012, p. 181).

This model, in other words, maintains that perceiving the actions of another individual *is already* to understand their affective, sensorimotor (i.e. intentional) meaning. Moreover, it implies that this is not a process encapsulated in the brain. In fact:

[t]he explanatory unit of social interaction is not the brain, or even two (or more) brains, but a dynamic relation between *organisms*, which include brains, but also their own structural features that enable specific perception-action loops involving social and physical environments, which in turn effect statistical regularities that shape the structure of the nervous system (Gallagher, 2005). The question is, what do brains do in the complex and dynamic mix of interactions that involve full-out moving bodies, with eyes and faces and hands and voices; bodies that are gendered and raced, and dressed to attract, or to work or play; bodies that incorporate artefacts, tools and technologies, that are situated in various physical environments, and defined by diverse social roles and institutional practices? The answer is that brains are part of a system, along with eyes and face and hands and voice, and so on, that enactively anticipates and responds to its environment. How an agent responds will depend to some degree on the overall dynamical state of the brain and the various, specific and relevant neuronal processes that have been attuned by evolutionary pressures, but also by personal experiences (the historicity [...]) of the agent [...]. How an agent responds depends on the worldly and intentional circumstances of the agent, the bodily skills and habits she has formed, her physical condition, a variety of so-called extraneous factors [...], with whom she is interacting, and what the other person may expect in terms of normative standards stemming from communal and institutional practices (Gallagher et al., 2013, p. 422).

This idea is also valid for musical experience. The *enactive* approach to musical experience sees music cognition as a process that involves a living system considered as a whole, in its sensorimotor coupling with the environment. It is thus not reducible to structures inside the head. Given these considerations, it seems that the mirror neuron activation might not be sufficient *per se* but rather,

¹⁶ I am grateful to Ezequiel di Paolo, who shared with me his thoughts about this point through personal correspondence.

acts as an enabling prerequisite for sense-making. The enactive story is in fact compatible with inner processing as long as it is characterised not in the usual representational terms. What enactivism denies, instead, is the idea of explaining thinking by assuming that the brain is “playing around” with terms like “belief” or “pretend” (De Jaegher, 2010).

A more concrete example about the enactive approach to musicality can be discussed with regards to the ability of the musical agent to *incorporate* tools in order to make music. Music making, indeed, often employs musical instruments extending *de facto* the body through a new *organ* - a musical instrument could be seen an actual extension of the body - for coupling with the musical environment in a meaningful way. I will return to this example in the last section of this chapter.

What we can say here is that also the way I perceive music is already sense-making, it is already a process that is direct, interactive and not solely caused by (or determined in) the brain. As O'Regan and Noë put it, ‘the knowledge of the ways movements affect sensory stimulation is necessary for experience’ (O'Regan and Noë, 2001a p. 1055). The experience of the world, therefore, can be such and such mainly on the basis of the creature’s sensorimotor abilities, without any sort of (explicit or neural) simulation.

The way we engage with the world is never a matter of (explicit or neural) representations and inferential mediations but it is rather radically entwined with the ability to actively and meaningfully interact with the environment. In other terms, with regards to a musical context, the motor expertise of an agent (her vocabulary of musical-directed actions) is constituent and indispensable for musical sense-making as it allows the possibility to interact meaningfully with the musical environment. But what does this mean? How could one *interact* with the musical environment, and, furthermore, what is a *musical environment*? I do not intend to imply that the musical environment is a sort of *sonic world* where pre-given entities (musical sounds) meet the listeners’ ears, but rather, I consider it as an *in-constitution* category, to be intended in terms of the modalities in which an agent intentionally interacts with and constitutes the musical world, ascribing meanings and values. As Johnson puts it:

Meaning includes patterns of embodied experience and pre-conceptual structures of our sensibility (i.e., our mode of perception, or orienting ourselves and of interacting with other objects events, or persons). These embodied patterns do not remain private or peculiar to the person who experiences them. Our community helps us interpret and codify many of our felt patterns. They become shared cultural modes of experience and help to determine the nature of our meaningful, coherent understanding of our “world”. (Johnson, 1987, p. 14)

This is exemplified by collective musical contexts, from rituals to orchestral rehearsals, where the very same “world” - though *enacted* in different modalities - is constituted by each agent who participates in the event through embodied sensorimotor patterns. The directedness towards musical objects constitutes them continuously through every intentional act *bringing forth* the agents’ own domain of meanings.

This intentional directedness, therefore, needs to be radically intermixed with the object itself, in order to ascribe its configuration, resulting in a co-determined process. To interact meaningfully

with music means to enact a domain of meaning, to *give sense*. A pianist does that in every performance, in every chord, in every gesture intentionally directed towards the musical execution. But also a listener is doing that; a listener is in fact motorically entwined with music, and the way her cognitive system resonates is already sense-making. Joel Krueger describes this position quite clearly:

This sensorimotor knowledge consists in the practical understanding that modulations of bodily movement and attentional focusing affect sensory change. For instance, when we perceive a visual scene, movements of the head or body change the way that occluded objects (e.g., part of a bush obscured by a tree standing in front of it) gradually reveal themselves as I move closer to or around them. We possess similar knowledge of how bodily movements and attentional modulations shape the character and content of musical experience. Rudimentary sensorimotor knowledge is thus the implicit, practical understanding that, as an embodied agent, I possess the sensorimotor skills needed to secure experiential access to different features of my world by using my body in different ways. Being sensitive to the sensorimotor contingencies governing my relation to perceptual objects is what it means to be a “skilled” perceiver (Krueger, 2011, p. 12).

I shall call the motor action constituting the sensorimotor knowledge necessary to be engaged to a musical feedback *teleomusical acts* (in Greek τέλος means more or less “goal” or “result”). These actions with a musical goal-directedness, therefore, not only represent the basis for the musical repertoire of acts that the musician develops throughout her musical life, but are also the necessary pre-requisite of a fully constituted musical intentionality. This last point will be clarified in the next chapters, in particular when focusing on exploratory behaviours and infants’ musicality. For now, we have to acknowledge that this characterisation of sensorimotor skills as fundamental properties for musical intentionality forces us to admit that musical experience consists in situated, skilful, coping (Noë, 2004). As Krueger notices, in fact, ‘we enact music perception via sensorimotor manipulation of sonic structures’ (2009, p. 104). This “manipulation” as he concludes, is to be understood through the notion of *affordance* and *musical affordance*. The next paragraph offers an analysis of this.

4.6 - Visual-motor affordances and the power of action

The notion of affordance has been introduced by James Gibson (1977; 1979) as the feature of an object or the environment that allows the observer to perform an action, a set of ‘environmental supports for an organism’s intentional activities’ (Reybrouck, 2005b). The aim of the Gibsonian perspective was to challenge the idea that perception is based on sensation (Gibson & Gibson, 1955). His starting point was the assumption that an array of light in the environment has simply too much information to be fully represented in the brain as an analogous entity.

If perception has no mediating categories such as representation, then it can be defined as *direct*. According to Rockwell (2005), Gibson’s claim that visual experience exists in the light, should be

understood not in terms of ecological sufficiency (light, in fact, cannot be experienced if there is no one to perceive it) but rather, through the *interaction* between organism and the environment. The status of affordances is indeed unclear in absence of the animal (Chemero, 2009, p. 150).

Studied under very different perspectives, this concept has become a crucial issue not only for ecological psychology, but also for cognitive sciences, artificial intelligence studies, and philosophy of mind. This variety of approaches has widened the already ambiguous definition originally provided by Gibson himself, contributing to the development of different standpoints in open contrast with each other (see Zipoli Caiani, 2011).

During the last two decades several researchers tried to extend the notion also to musical experience, with the aim to draw a coherent theory of musical affordances (Clarke, 2005; Reybrouck, 2005b; Nussbaum, 2007; Krueger, 2011a; 2011b). In this section I will argue for a particular concept of musical affordances, that is, as I see it, narrower and less ambiguous in scope and more closely related to its original. In the context of motor action research, there is a common agreement regarding a basic understanding of the notion of *affordance*, usually intended in terms of a set of possible motor actions evoked by the intrinsic properties of an object or the environment. According to Gibson (1979), indeed, the visual perception of an object leads to an *automatic selection* of those of its intrinsic properties that support the individual's physical interactions with it. These properties, however, are not only abstract, physical or geometrical features, but 'incarnate the practical opportunities that the object offers to the organism which perceives it' (Rizzolatti & Sinigaglia, 2008, p. 34). When, for example, an object like a cup is located in the subject's *peripersonal space* (Costantini et al., 2011), it can represent the goal of the individual's grasping act and this subject/object interaction is codified through an *affordance*.

A key role for this current understanding is played by the previously discussed neuroscientific evidence of *canonical neurons*, a set of neurons which discharges when an individual simply observes an object without performing any movement, as well as when he/she grasps that object. The discovery of these visuo-motor neurons shows how an object can *afford*, according to the subject's motor expertise, a set of possible actions that can be performed thereon, relying on a *sub-cognitive* form of understanding, not linked to mental representation or higher mental faculties. Vittorio Gallese, in describing canonical neurons, states that

[t]he most interesting aspect [...] is the fact that in a considerable percentage of neurons, a congruence is observed between the response during the execution of a specific type of grip, and the visual response to objects that, although differing in shape, nevertheless all 'afford' the same type of grip that excites the neuron when executed [...]. The intrinsic relational functional architecture of primates' motor system likely scaffold the development of more abstract and detached forms of intentionality, as those characterising thought in our species (Gallese, 2009, pp. 489-490).

The basis of the intentional relationship between an organism and the environment, therefore, can be reconsidered in terms of how the *motor possibilities* (Poincaré, 1908; 1913) of the subject's body can interact with the surrounding objects, advocating for a *motor approach* to intentionality (Sinigaglia, 2008a). This form of intentionality doesn't require any high-level, metacognitive ability as 'the intentional character [...] could be deeply rooted in the intrinsic relational character of body

action' (Gallese, 2009, p. 489). In other terms, cognition arises from the bodily interactions with the environment, depending on 'the kinds of experiences that come from having a body with particular perceptual and motor capacities that are inseparably linked and that together form the matrix within which memory, emotion, language, and all other aspects of life are meshed' (Thelen et al., 2001, p. 1).

However, it seems that the current ecological account for *affordances* inspired by Gibson is not totally committed to this characterisation. Rather, it displays *epistemological vagueness*. One of the assumptions at the basis of this perspective is that every set of behaviours ascribable in terms of a unitary action has the right to be described as an act potentially evoked by its related affordance. In order to clarify this statement we can think about the possible analogy between these two different conditions:

- A cup affords the act of grasping.
- The sight of a movie-trailer affords the act of going to the cinema.

From an ecological standpoint, these actions (grasping and going to the movies) are both homogeneous and describable as unitary. But the *inferential* feature of the second one makes it unsuitable: otherwise, every *a-posteriori* correlation between events and actions should be considered *affordative*. To get an idea of the variety of phenomena included under the concept of affordance we can indeed have a look at this excerpt:

Air, the medium we live in, affords breathing. It affords walking or driving through, and seeing through, at least in communities that are free of smog (E. J. Gibson, 1982, p. 55).

This example presents some incongruences with the necessary setting of a scientific research on the notion of affordance. Even if sometimes the air is breathable, this does not imply that breathing is a *goal-directed act*, intentionally linked with the portion of air considered as an *intentional object* (Menin, 2011, p. 12). We cannot indeed integrate every possible interaction between the two poles of action (subject and object) into a genuine reflection on the issue of affordances without applying a rough *objectivation*, hence depriving the notion from any phenomenological characterisation. The relationship between a subject and the air that the agent is breathing, simply put, cannot be described as constituent of any intentional relationship. It can only be described in terms of physical-chemical events¹⁷.

But if some events do have affordances, we should investigate the *object* of these, and also the

¹⁷ With regard to this point, it should be clear that, if we are to develop an experimentally expendable notion of "affordance", the original Gibsonian ambiguity needs to be explicitly overcome. If it is true that one of the most fascinating features of affordances resides in its being direct and somehow 'automatic', thus not requiring any cognitive or attentional mediation, we cannot go, on the basis of this not-needed mediation, to the extent of saying that no intentional sensorimotor relationship is required for an affordance to take place. The notion of 'basic motor act' introduced in the recent literature (i.e. Rizzolatti, Sinigaglia 2008) clearly highlights goal-directedness as the central feature of sensorimotor interactions with the environment.

role of the relevant *subject*. Gibson (1979) stated that a niche is the set of affordances for the agent and that different organisms with different sensorimotor skills ‘may have physical collocated but nonetheless non-overlapping niches. For example, a human and a bacterium may share a physical location [...], but their niches will not overlap. [...] Gibson suggests that this is the way to make sense of the mutuality of animals and environments. An animal’s abilities imply an ecological niche. Conversely, an ecological niche implies the animal’ (Chemero, 2009, p. 147).

The *abilities* in play, here, are the sensorimotor skills that would allow the living creature to make-sense of the world through intentional goal-directed patterns of embodied action. We see through our eyes, which are movable, placed in a head that can turn around, linked to a body that can explore the environment (Gallagher & Zahavi, 2008). The objects of our perceptual experience are shaped by what we do or by what we are able to do. It must be clear, however, that although the discovery of canonical neurons provided the discussion of affordances with an evidence-based framework that focuses on embodied actions, perceptual experience is something “wider” than neural states. Perception, to put it shortly, is a form of action. And any form of action implies sense-making. A given musical pattern can be experienced as *far* or *close* to me, as spatially- expressive- and movement-relevant. It will never be an abstract object of symbolic manipulation, in a disembodied mind. It might resonate *in me* through my cognitive system; yet it could be constituted only on the basis of the dynamics of interaction *outside of me*, as action possibilities that would allow me to perform the same phenomenologically relevant goal-directed actions in the world.

In the realm of music, many authors recognized the relevance of the concept of affordance, which has been investigated during the last few years in order to clarify the subject/object relationship characterising musical experience, with a particular emphasis on the surrounding, sound-made environment. What does music *afford*? How can a subject *interact* with a musical environment? Which modalities are involved in this sense-giving process? The following paragraph presents some of the answers to these questions provided in the current debate, proposing incentive for further discussions as well.

4.7 - What is a musical affordance?

This section will offer a closer look at the notion of musical affordances, focusing on the works by Clarke (2005), Nussbaum (2007), Krueger (2011a; 2011b), and Reybrouk (2005), probably the most influential and well-known studies on this specific and problematic issue.

1) *Ways of Listening* by Eric Clarke (2005) aims to face the problem of musical understanding through a Gibsonian interpretation in order to overcome the interrelated dichotomies of subjects/objects, passive/active listening and autonomy/heteronomy of musical experience. The author firstly suggests that musical structure is not a construction within the mind, but, rather, something inherent in the environment, with which our auditory system would *resonate*. Secondly, Clarke describes the dynamics of these resonances, introducing the key concept of *affordance* (ibid., p. 36). Two of the most quoted excerpts of Clarke’s book help to highlight what kind of *phenomenic topography* is included in his notion of musical affordance.

Music affords dancing, worship, co-ordinated working, persuasion, emotional catharsis, marching, foot-tapping, and a myriad other activities of a perfectly tangible kind (ibid., p. 38).

In the specific contexts of musical hermeneutics, musical material can be conceived as affording certain kinds of interpretation and not others [...]. Interpretation is also action - the speaking, writing, gesturing and grimacing in which interpretation is manifest [...]. The recapitulation of the first movement of Beethoven's Ninth Symphony affords writing (or speaking) about in terms of murderous sexual rage, or the heavens on fire. Interpretative writing and speaking are forms of action (ibid., p. 204).

Looking at the examples provided in the first list, allegedly belonging to the class of *perfectly tangible acts* that can be afforded by music, we can notice a partial overlap with E.J. Gibson's holistic stance. However, it appears problematic to assume that music is in some way the intentional object of all this variety of behaviours. What kind of phenomenological characterisation can be ascribed to musical experience if, *dancing, persuasion or speaking about music* are all under the umbrella term of *affordance*? Since, in fact, there is no substantial difference between the way music affords foot-tapping and writing pieces of musical criticism - in Clarke's view - the mutualism between agents and environment does not result grounded into the sensorimotor patterns that eventually underlie music and musical experience. As Clarke admits:

A reader might object that this discussion [...] is unacceptably speculative, full of interpretive license, and basically out of step with an ecological approach: it seems to depend heavily on verbal and dramatic information to interpret the perceptual information [...] rather than relying on specification by stimulus invariants. But this overlooks the fact that all of the elements mentioned (the drama, the characters, the sounds) are part of the available information for a viewer/listener (ibid., pp. 87-88).

But how exactly could the direct agency of the cognizer pick up from the (musical) environment conceptual information related to a speculative interpretation of the stimulus? Writing pieces of musical criticism implies, in fact, an *inferential* level of mediation. In this sense, the meaning of *to afford* emerging from Clarke's contribution does not exceed the concept of *to evoke*, or *to elicit*, showing that musical experience is, broadly speaking, evocative. This, however, does not mean that Clarke is wrong. His strategy to unify percepts and concepts (Nonken, 2008), cannot be valid for a basic level of sense-making, but there is no doubt that high-level forms of musical engagement could take place in a more inferential level. For example, it seems reasonable that I can be moved by Purcell's *Dido and Aeneas* because I have read the tragic story of the two lovers' fate two hours before the concert, allowing me to make more than legitimate inferences during the actual performance. It is quite clear, in fact, that there might be a distinction between my experience of the music, and the experience of another individual who does not know anything about the story.

Another important aspect of Clarke's standpoint is the idea that a living system can adapt to a particular sonic environment, maintaining the direct characterisation that 'specifies properties of the

object itself to an organism equipped with an appropriate perceptual system' (2005, p. 15). The resonance that embodies this adaptation is *active*, considering that 'the ecological approach presents perception as a mutual relationship between organism and environment, so that every description of perception is therefore specific to an individual's capacities and perspectives' (ibid., p. 156). This description of circular specification, as well as the example about *Dido and Aeneas* previously provided, seems to contradict the purely ecological idea of a fully given musical environment. If meanings are *already there*, then how could mutual relationships be possible? How can *different* musical subjects develop *different* ways of listening? And what would happen if an individual does not grasp the actual meaning of a given musical structure? Mutual specification, rather, implies that categories like "meaning", "environment", or "listening" would be *in-constitution* - and not fully given in the world, or inside someone's perceptual system. Despite acknowledging the rare pertinence of Clarke's book to themes strictly related to the development of a post-Cartesian cognitive science of music, it seems that the broad notion of musical affordance employed, as well as his coherent ecological focus on properties of the musical environment, reflects a rather direct application of Gibson's view - although the concept of *mutual relationship* used by Clarke could have been developed into a more *enactive* - and sensorimotorically grounded - perspective.

2) More ambitious is the theoretical proposal of Charles Nussbaum (2007) who actually aims to unravel the *riddle of musical experience* (ibid., p. XI) in light of its *representational* nature. To understand the problem he is dealing with, we can first have a look at this citation:

It takes only a small amount of perspicacity to realize that music is remarkable, indeed an astounding, phenomenon. The emergence of human musical experience from the audition of organized tones remains deeply puzzling, truly 'a riddle wrapped in a mystery inside an enigma,' a riddle, moreover, of very long standing (ibid., p. XI).

But this statement is highly misleading, especially with regards to the set of problems involved in musical understanding. The apparently naïve implication that musical experience builds up from the auditory perception of analytically isolable basic elements is not only unwarranted, but it also stands in open contradiction with the most fruitful studies of action-related aspects of perception (Rizzolatti & Sinigaglia, 2008, pp. 50-52).

If we are to unravel the riddle of musical experience, we need a thread on which to tug. Construing music as representational, as a symbolic system that carries extra-musical content, I hope to persuade you, exposes such a thread (Nussbaum, 2007, p. 1).

As this quotation clearly shows, the enigmatic position of the musical problem and the representative option proposed in the book are closely interrelated. This epistemological situation does present analogies with the *post-Cartesian* dilemma concerning the emergence of a thinking substance from an extended body: in both cases, the hypostatisation of the starting dichotomy necessitates an *ad-hoc* solution in order to mediate between the two substantialised realms. It is

worth noticing, with regard to this topic, that the notion of *affordance* is usually connected with a strong *anti-dichotomist* position, as Gibson (1976, p. 129) first pointed out. In Nussbaum's work, however, an affordance is conversely meant as a *mediation tool*, functional to the perpetuation of a radically dualistic stance. Indeed, according to this scenario, musical affordances are considered to be conceptual bridges between a *low-level* dimension of musical experience, conceived in terms of a meaningless isomorphic transcription on a pitch-time diagram of the stimulus, and an idealised *high-level* dimension that includes every aspect of musical experience, broadly meant. Even without discussing the proposed theoretical framework, we can rule out Nussbaum's contribution, as he considers affordances to be a *cognitive form of understanding*, linked to mental representation or higher mental faculties, thus in contrast with any position developed from an ecological standpoint.

3) Another approach aimed at developing a sustainable notion of musical affordance is provided by the recent works of Joel W. Krueger (2011a; 2011b). The author argues that 'an affordance is a relational property of the animal's environment perceived by that animal as having a functional significance for that animal' (Krueger, 2011a, p. 4). He also states that music is perceived from birth as an *affordance-laden* structure that *affords* a sonic world (ibid. p. 1) that further *affords* possibilities for, among other things, (i) emotional regulation and (ii) social coordination. From an epistemological standpoint, however, this proliferation of *affordative levels* seems - at least - suspicious, and gives rise to three questions:

- Does music afford a sonic world in the *same sense* as this sonic world affords emotional regulation?
- What would be the relevance of such a claim in the study of the *intentional relationship* between a subject and a musical object?
- How should we describe the *animal-environment* relationship if music affords a sonic world that further affords acts of any type?

In trying to answer these questions, we find out that the *mediation* offered by the notion of "sonic world", besides being unnecessary, if coherently implemented would substantially compromise the *direct character* of the concept of affordance, explicitly acknowledged by Krueger (2011a, p. 7). Besides that, however, the crucial point of Krueger's argument is the characterisation of musical space, developed in league with the tradition of *spatiality-for-action* (i.e. Poincarè 1908) mainly discussed in the *visuo-motor* domain (Rizzolatti et al., 1997; Sakata et al., 1997). In his paper *Doing things with music* (2011a), Krueger faces the problem of *musical space* from a purely ecological standpoint, defining musical environments (or sonic worlds) as *comfortable* or *stressful*, whereas in his other work, *Enacting musical content* (2011b), he does contrast inner (or structural) and outer musical space. *Outer musical space* is here identified with the localisation of the occasional sound source, whereas *inner musical space* is described as 'the piece's inner syntactical structure established by the way that constituent components (e.g. tones, rhythmic progressions, etc.) go together, lending the musical piece its sonic coherence as a composed object' (2011b).

Since outer space is meant as *non-musical* (related to the localisation of musical stimulus), we would assume that *inner space* and *sonic world* are interchangeable notions, defining from different standpoints the same musical *space for action*. But this assimilation is hardly accomplishable. The notion of *sonic world* arises indeed from a standard ecological standpoint, while the concept of *inner space* descends from an approach in which the musical surface is identified with a Cartesian diagram with time and pitch as axes, embracing *de facto* the “pharmaceutical model” (Sloboda, 2005, p. 319) of musical understanding which Krueger explicitly refuses (Krueger, 2011a, p. 3). As a result of this irreducible duality of approaches, the notion of *affordance* connected to the concept of sonic world seems extremely relational, as it is associated with *every kind* of activity that music could possibly elicit, whereas the one connected to the concept of inner space is conceived from a completely *objectivistic* point of view. What both of these concepts are missing is the *intentional character* needed to make musical affordances a *phenomenologically* relevant notion, which maintain the co-determinative character between subjects and objects.

4) A better awareness of the range of issues implicit in the *enactive* approach to human musicality is shown by Mark Reybrouck (2005b), who addresses in his work an embodied characterisation of musical experience. The author, using syncretic integration of different perspectives (from classic pragmatism to cognitive economy), aims to overcome the prevailing *objectivism* in the realm of musical understanding, by applying the key notion of *sensorimotor coupling* (which defines the perception-action loop) to the analysis of this topic. His strategy consists of defining the two domains of (i) musical *experience* and (ii) *motor cognition*, showing how they can be connected in such a coupling. ‘Musical experience’, he claims, ‘is not basically different from an auditory experience at large. It is continuous with the natural experience or experience proper [...] with a difference in degree rather than in quality¹⁸’. This equivalence between *experience proper* and *aesthetically connoted perception* justifies the application of a general concept as sensorimotor coupling to the peculiar realm of musical understanding, considering the importance of action in acoustic perception (Kohler et al., 2002). On the other hand, the processes of motor cognition are introduced through the discussion of the pivotal notion of “image schemata”, defined as ‘recurring, dynamic pattern[s] of our perceptual interactions and motor programs that give[s] coherence and structure to our experience’. The two classes of image schemata presented as the most relevant to the study of musical experience are the “container schema” and the “source-path-goal schema”. While the *container schema* is ‘a pervasive mode of understanding everyday experiences in terms of “in” and “out”’, the *source-path-goal schema* represents the feature of being oriented towards a goal in a continuous, temporally extended path. The author then illustrates the musical analogies of these image schemata, introducing the concept of “musical affordance”:

There are, as yet, many possibilities that stress the ‘action aspect’ of dealing with music. I mention five of them: (i) the sound producing actions proper, (ii) the effects of these actions, (iii) the possibility of imagining the sonorous unfolding as a kind of movement through time, (iv) the mental simulation of this movement in terms of bodily based image schemata and (v) the movements which can be possibly induced by the sounds

¹⁸ Being an online publication, it is impossible to specify the page of this and of the following quotes

(ibid.).

There are, however, some doubts about the robustness of the conceptual framework developed here. In particular, the choice of defining *separately* the two realms of musical experience and motility seems weak, considering that Reybrouck himself acknowledges *non-objectivism* as one out of the two main claims defining an embodied approach (2005a, p. 10). Moreover, the two definitions proposed appear to be questionable: indeed, as Clarke (2005, p. 1) pointed out, one of the dichotomies that a notion of musical affordance could help to overcome is that which opposes *autonomy* and *heteronomy* of musical experience, but the *full integration* of musical experience into mere acoustic experience proposed by Reybrouck seems to neglect this aspect. On the other hand, the notion of *image schema* and the *actual schemata* proposed could not be accepted by many proponents of embodied cognition as these concepts seem to individuate a class of mental schemata, rather than a truly embodied kind of sense-making. This problem emerges clearly from the analysis of the notion of “goal” provided by the author: indeed, the kind of *goal-directedness* described by Reybrouck cannot be assimilated into the classic *motor-grounded* one (see, for instance, Murata et al., 1997; 2000, for a focus on acts such as *prehension*). However, it does present similarities with the *teleological stance theory* advocated by authors such as Csibra and Gergely (2003), where the understanding of the others’ intentions and goals can only be possible from an *external*, ascribing, perspective (see also Dennett, 1987). On the other hand, Reybrouck’s contribution is certainly fascinating, as he explicitly appeals to enactive and embodied approaches. In a more recent paper, for example, we are told:

Musical sense-making, [...] can be addressed in terms of interactions with the sounds, both at the level of perception, action and mental processing. It is a position that broadens the scope of music research, encompassing all kinds of music and sounds, and going beyond any kind of cultural and historical constraints. Music, in this broadened view, is to be defined as a collection of sound/time phenomena which have the potential of being structured, with the process of structuring being as important as the structure of the music. As such, it is possible to transcend a merely structural description of the music in favour of a process-like description of the ongoing process of maintaining epistemic contact with the music as a sounding environment (Reybrouck, 2012, p. 392).

The following quote, moreover, provides a concrete example of this process of *epistemic contact*, portraying the situation of a violin player dealing with the intonation of a pitch:

in order to produce a beautiful sound he relies on sensory-motor integration, “shaping” the sound through the perception of the sounding result which must match the internalized representation of this sound. Sound production, on this view, entails the reciprocity of “doing” and “undergoing”[...]. The reciprocity of doing and undergoing is typical of sensory-motor integration. This is obvious in “playing” music, but it applies also to the process of “listening”, with the listener imagining or

simulating mentally the manifest movements of the players [...]. There is, in fact, empirical evidence that motor imagery and motor execution involve activities of very similar cerebral motor structures at all stages of motor control (Crammond, 1997). It allows us to stress the continuity between “sensory-motor integration” and “ideomotor simulation”, the former dealing with movements that are actually executed in real-time, the latter with movements that are simulated at an ideational level of motor imagery (Reybrouck, 2005b, p. 231).

However, as we can notice, the notions of mental abilities are featuring strongly, raising several problems when compared to a more *direct* form of understanding and interaction with the musical feedback. In the next paragraph I will provide a definition of musical affordances in light of a more enactive account to perception and sense-making, without any involvement of higher mental faculties.

4.8 - Teleomusical acts for musical sense-making

Another way to deal with musical affordances is to consider them as *properties of the intentional relationship between musical subjects and objects*. Indeed, the scenario described above advocates a new phenomenological characterisation of musical perception, where the (musical) object is constituted with respect to the motor repertoire of the listener - through embedded and embodied forms of interaction. This pivotal role of the body and its motor knowledge has been addressed in the phenomenological equivalence between perceiving and giving sense to the percept (Merleau-Ponty, 1945), where

my body is geared to the world when my perception offers me a spectacle as varied and as clearly articulated as possible, and when my motor intentions, as they unfold, receive from the world the responses they anticipate. This maximum distinctness in perception and action defines a perceptual ground, a basis of my life, a general milieu for the coexistence of my body and the world (Merleau-Ponty 1945, [1962], p. 250).

Rather than postulating high-level cognitive abilities or simulation-like mechanisms to account for basic forms of musical understanding, the phenomenological explanation provided requires only my body, conceived as ‘the meaningful core which behaves like a general function’ (ibid. p. 46) in its *perceptual, continuous mutual and active* interaction with the world. A skilled guitarist might be unable to say where to put her/his finger to perform a solo, but s/he can use the *motor knowledge* of the fingers to reconstruct the actual set of notes played, by just putting the hand on the strings. I believe that this kind of sensorimotor process not only represents the basis of musical understanding, but it can also shed light on the notion of musical affordance, relying on a sub-cognitive, pre-linguistic, intrinsically motor, form of intentionality. In particular, the studies on the *ontogenetic basis of musicality* can provide further evidences of a non-cognitive characterisation of

the subject-object relationship of musical experience, as we will see in Chapter 5.

This account for musical intentionality, hence, leads the discussion to the analysis of *teleomusical acts*, chains of actions with a musical goal-directedness constituting the musicians' motor knowledge. Indeed, a correct characterisation of these acts cannot be limited to the executive side of motility, because what allows the possibility of understanding a musical object in terms of its actions (Overy & Molnar-Szakacs, 2009, p. 492) is the *goal* rather than the actual performed movement (Ticini et al., 2011).

I have argued against the fitness of the presented theoretical frameworks to consistently make sense of the number of problems implicit in the notion of musical affordance by proposing an approach that radically diverges from the standard accounts, considering musical objects as *entities constituted within the intentional motor-based relation that defines a musical context*. From this standpoint, the notion of *musical affordance*, correlatively with the key concept of music-directed (or *teleomusical*) acts, becomes then crucial for understanding the ontogenetically originary elements of music experience, and the processes that lead to their development to a fully constituted musical - embodied and enactive - intentionality. In light of these considerations we could consider an approximate distinction between *original* and *constituted* teleomusical acts.

1) Original Teleomusical Acts (OTAs) are autonomous and relevant to the foundations of human musicality, considering their basic kinematics and spontaneous emergence. Nevertheless these acts represent a *shared substrate* to human musical behaviours as they are executed early in infancy, yet clearly directed towards the constitution of a musical object. The infant brings forth her domain of meaning by interacting with the environment in order to constitute and manipulate sounds objects. In this sense, OTAs enable the intentional relationship between music and the perceiver, as they constitute the basic sensorimotor skills necessary to make sense of the musical object.

2) Constituted Teleomusical Acts (CTAs) are not authentic chains of acts in the sense that they are built through the unification of a set of OTAs. Yet, the fluidity of the kinematics allows the performer to execute them as *unitary goal-directed actions*. Playing a chord on the piano requires temporal coordination, sensibility of the fingers, wrists, arms and back, expressivity and so on. However, a skilled pianist does not simply integrate these categories one by one. Rather, she would achieve the goal (playing *this* chord in *this* way) through an intentional, fluid, active, rapid execution.

These two different, yet interrelated, levels of actions, might represent an important turning point for music cognition, as they can provide a biologically plausible theory of non-musicians and musicians' abilities to enjoy, make sense, and perform, music. One of the obvious counterarguments against the centrality of action for musical sense-making is that also musical non-experts display some kind of motor resonance or entrainment to a given musical stimulus. Therefore, it can be argued, the development of musical directed chains of acts does not play a meaningful explanatory role, since people who lack this motor expertise could be engaged in a musical experience as well. A first answer is that non-musicians relate to musical objects through a theoretical kind of intentionality, as the one that Leman named "cerebral intentionality". This solution, however, seems weak and stands in open contrast with the phenomenological notion of motor intentionality previously defined.

Rather, if we are to take seriously the challenge of OTAs and CTAs, noticing that OTAs are *ontogenetically original*, it could be argued that they would represent the roots of the sensorimotor skills required for musical sense-making also for non-musicians. These acts are basic (simple to perform), plastic (they can be easily improved, and performed through different strategies and motivations), spontaneous (they emerge in infancy without any imposition or help from the caregiver), ecologically relevant (they make sense only in light of a given subject-object relation) and goal-directed (they are directed towards sounds' properties and *protomusical* structures). I will get back to infants' musicality in the next chapter and explore in greater detail the dynamics from which a minimal form of musical intentionality emerges in infancy. For now, however, it goes without saying that non-musicians possess the basic skills to perform acts of hitting, plucking or scratching, and it is obvious as well that they would intend these actions non-inferentially in the very same way that listening to someone knocking at the door provides the perceiver with a clear understanding of what is going on (i.e. someone is outside the door). It can be argued that the way non-musicians make sense of music is an intentional relationship based on motor skills qualitatively different from the ones *enacted* by an expert performer. Both OTAs and CTAs, moreover, can modify the expectations of the objects' affordances (E.J. Gibson, 1988) showing again the level of mutual determination of the two inseparable categories of subjects and objects. A guitarist listening to a piece for guitar transcribed for piano, in fact, will display a different intentional relationship from the kind of sense-making that would emerge when hearing the same piece for her own instrument (guitar). For the categorization of the motor skills employed in sense-making would be hierarchically different. As Chemero admits, indeed, 'over developmental time, an animal's sensorimotor abilities select its niche - the animal will become selectively sensitive to information relevant to the things it is able to do. Also, over developmental time, the niche will strongly influence the development of the animal's ability to perceive and act' (2009, p. 151).

Matching the study of autonomous living systems in their sensorimotor and affective coupling with the environment with the notion of musical affordance provided here, makes the enactive approach to musical experience a dynamic science of the brain-body-environment nexus. The challenge, in this sense, would be maintaining a sharp equilibrium between the two poles, trying not to collapse into an objectified view of musical phenomena on both sides of the skin taken autonomously, without considering their mutual interactions. Since I have already described the difficulties in embracing a complete *internalistic* perspective to music cognition, the next paragraph will focus on the so-called Extended Mind Hypothesis (i.e. Clark & Chalmers, 1998) exploring its assumptions and perspectives. In doing this, I will mainly focus on the *radically enactivist* critique offered by Hutto and Myin (2013) providing also examples in current musical literature (i.e. Krueger, 2014), nevertheless showing that considering the (musical) mind as merely *extended* could represent a step backwards for a truly post-cartesian cognitive science of music.

4.9 - Transparency and sensorimotor primacy

We could draw an approximate distinction between two trends in philosophy of mind, when it comes to understanding cognitive processes. On one side we have *internalists*, who maintain a position where cognition is defined in terms of the animal's internal factors, usually relative to the boundaries of the skull (i.e. Adams & Aizawa, 2001). As we have seen, this idea insists on

considering the environment and any exogenous influence, as a mere causal contribution that may play a supportive role in shaping cognition (Hutto & Myin, 2013, p. 136). The core body of standard music cognition offers the paradigmatic example of this standpoint, seeing music as an external stimulus that *causally* affects the brain of the subject. On the other hand we have proponents of an *apparently* opposite claim. *Externalists*, in fact, deny that cognition can be reducible to what happens inside the skull or the brain-body system of the living creature. Although the enactive approach embraces this openness of cognitive process (to be radically imbricated into the brain-body-environment nexus), there are some differences in the assumptions the externalists and enactivists appeal to, in order to defend this position.

The Extended Mind Hypothesis (EMH) basically maintains that ‘the local material vehicles of some aspects of human cognition may, at times, be spread across brain, body and world’ (Clark, 2009, p. 966). In particular, this can occur when ‘the local operations that realize some human cognizings include (possibly quite complex) tangles of feedback, feedforward and feedaround loops that promiscuously criss-cross the boundaries of brain, body and world’ (ibid.). This claim, is justified on the basis of the so-called *parity principle*, which is usually formulated as follows

If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process (Clark & Chalmers, 1998, p. 8).

The principle, in other terms, holds that sometimes, when given conditions are met, ‘features of the environment can co-constitute the mind’ (Hutto & Myin, 2013, p. 139). Clark and Chalmers’ (1998) famous example to mount the defence of their claim involves two main characters, Otto and Inga. Otto suffers from a slightly amnesic condition, which forces him to write down on a notebook any relevant information that might be useful later but could easily be forgotten. When Otto wishes to visit the museum he needs to use his notebook in order to find the right address. On the contrary, Inga, uses her biological memory to find out where the museum is, thus ‘activating her unconscious, lingering belief concerning the museum’s location. If this suffices for regarding Inga’s unconscious belief to be classed as mental, then, by parity of reasoning, we have every reason to adopt the same stance with respect to what is written in Otto’s notebook’ (Hutto & Myin, 2013, p. 139). In fact,

[The] organism is linked with an external entity in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behaviour in the same sort of way that cognition usually does. If we remove the external component the system’s behavioural competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head (Clark & Chalmers, 1998, p. 7).

The problem with these statements, however, is the definition of “cognitive processes”: only after a clarification of what is intended for cognitive process we can justify the analysis of its location (skull-bound, bodily, extended, neuronal) (Di Paolo, 2009, p. 10). But, as Hutto and Myin admit, ‘the appeal to parity is utterly silent about the nature of mentality’ (2013, p. 140). Moreover, it seems that there are some differences between Otto’s notebook and Inga’s biological memory. As pointed out by Adams and Aizawa (2001), for example, while the notebook is not simply automatically updated by new information, Inga’s memory is non-inferentially modified every time she receives new information (for example, when she learns that an earthquake changed the whole topography of the city where the museum is located) (Weiskopf, 2008). Of course, die-hard proponents of externalism claim that technology would offer some arguments that would make these differences disappear. For example, ‘it is easy enough to imagine an electronic notebook being updated through wireless connections to the brain’ (Hutto & Myin, 2013, p. 141). Similarly, in a musical context, a performer improvising on her instrument would *offload* her musical expertise into the instrument, in processes that are functionally coupled with the constituting musical environment. This holds true also for joint musical events: ‘the players in a jazz trio, when improvising, are immersed in just such a web of causal complexity. Each member’s playing is continually responsive to the others’ and at the same time exerts its own modulatory force’ (Clark, 1997, p.165).

Moreover, Joel Krueger, in a very recent paper, insists that also a musical *listener* would offload certain regulative structures into the music, reducing her cognitive burden and perceived exertion (Krueger, 2014). Although Krueger’s proposal is certainly fascinating and well argued, I think that a concrete assimilation of the enactive and affordance-based notion of musicality to a - Clark’s style - extended mind model, is counterproductive. A first objection is that one can accept the notion of “offloading processes” in these terms, only admitting that these processes do exist inside the head while listening to or performing music. Di Paolo explains this concern as follows:

Just imagine how an EM [extended mind] researcher would go about determining whether extra-bodily process X is constitutive of a cognitive process. He is at this point already convinced that the skull boundary no longer makes sense for determining whether something is cognitive. And yet, on consulting his EM handbook he finds this schizophrenic piece of advice: “Even though you know that you should not rely on the skull boundary in order to call process X cognitive, you must still use it to check whether it would have made sense to those confused people who did not take EM seriously and consider how they would have judged process X were it to happen in the head”. Isn’t this the absurd equivalent of “Even though in the past several cases of mental illness were erroneously diagnosed as cases of demonic possession, in modern days, in order to determine whether you have a case of mental illness you must always consult an exorcist first?” (Di Paolo, 2009, p. 11).

What is occurring inside the head, for example in terms of neural discharges, is not a cognitive process. It is only part of the story. Consider the previous analysis of motor resonance. The mirroring processes have explanatory value only in a relational domain between the agent and the (social) environment, constituting the enabling condition for sense-making. But ‘to call internal

processes as such cognitive is to confuse levels of discourse or to make a category mistake (neurons do not think and feel; people and animals do)' (Thompson & Stapleton, 2009, p. 26). Sense-making needs to be understood in terms of dynamic sensorimotor patterns of embodied action immersed into the reciprocity with the environment. In these terms cognition and musical cognition have no location. If we admit a notion of mind as "skilful capacity", then there is no need to posit internal processes that would extend into the environment.

The enactive mind focuses on the communal resources that expand beyond the skull, without any decoupled activity by means of external structures. Playing an instrument is a form of musical sense-making as the performer explores and actively engages with the environment by means of sensorimotor patterns of meaningful coupling and not because her computational or problem-solving analysis of the musical structure and rules is facilitated by her interaction with the instrument. Music playing (or music listening) is not problem solving. If this statement is too strong, it could at least be claimed that a subjective musical experience is not *always* problem solving. This would be a too limited account for music cognition, resulting in problematic positions as the one formulated by Huron, where the agents' body is essentially decoupled from the active dynamics of sense-making. This comparison with Huron shows in fact that functionalist and extended approaches share a similar assumption about the bodily involvement in sense-making.

As reported by Thompson and Stapleton, proponents of EMH see the body as just an element 'in a kind of equal-partners dance between brain, body, and world, with the nature of the mind fixed by the overall balance achieved' (Clark, 2008, pp. 56-57). 'For the enactive approach, the body (including the brain) leads in this dance because it is what realizes the autonomous organization necessary for individual agency and sense-making' (Thompson & Stapleton, 2009, p. 28). While playing a musical instrument, the agent *incorporates* it into her cognitive dynamical system rather than using it *instrumentally*. This *sense of body ownership* that can be applied to the musical instrument is what distinguishes the enactive account to the extended one (see also De Preester & Tsakiris, 2009). Does Otto experience the notebook in the same way Glenn Gould experiences the piano?

Let's think about Merleau-Ponty's famous example of the blind man that uses a cane. According to the French philosopher, as the cane becomes a transparent tool, incorporated into the agent's body image like an 'extension of the bodily synthesis' (1945 [1962], p. 153) the blind man becomes able to *see through his cane* by interpreting the data of the world without any inferential mediation (i.e. without measuring the cane in order to understand the distance of an object or feeling the pressure on his hand when the stick hits an object). Similarly, musical instruments are *transparent tools* during performances.

Merleau-Ponty's analysis of *the organist* is, in this sense, illuminating. An organist settles in a new organ like he would settle in a house, without analysing the instrument in a disembodied way. He does not prepare a *plan*, or cognitive *representations* of the registers and the pedals (ibid.). In fact, during the rehearsal or during a concert, the keys, the registers and the pedals are not simply located in an objective space. Rather, they become a horizon of musical-directed, motor, possibilities, therefore intermixed with the musician's physiology and with the (in-constitution) musical environment. This position could be clarified by appealing to the "transparency constraint" proposed by Thompson and Stapleton:

For anything external to the body's boundary to count as part of the cognitive system it

must function transparently in the body's sense-making interactions with the environment (2009, p. 29).

If this analysis turns out to be correct, the distinction described by Fishkin (2004) between *full embodiment* and *distant embodiment*, based on the locations of input and output, would be fully compromised (see Menin, 2011). According to this perspective, while violin playing can be seen as a case of full embodiment, given the high degree of integration between sound and action, organ playing would represent a situation of distant embodiment, considering the nature of the instrument itself (to produce sounds the wind must move within pipes after pressing a key or a pedal). But this is simply a *reification* of the relationship between musical subjects and their instruments, as Merleau-Ponty clearly observed. Moreover, as this entire chapter tried to suggest, a serious consideration of the role of the brain-body-environment systems, in defining musical contexts, challenges any non-intentional interpretation of human musicality, and reorientates the discussion of musical affordance to a motor-based level. These challenges, however, are not to be simply interpreted as a set of theoretical assumptions at the basis of a new model of the musical mind. A post-cartesian, enactive, cognitive science of music should explicitly address themes such as musical development and musical learning.

The hypothesis of teleomusical acts, as previously presented, will therefore play a key role in the next two chapters. Firstly I will discuss some minimal forms of musical sense-making considering also the sphere of affectivity. Secondly, in the final chapter, I will present two behavioural studies that my colleagues and I have run at the Music Mind Machine research center of the University of Sheffield and at the Donders Institute for Brain Cognition and Behaviour of the Radboud University of Nijmegen. Both these experiment will corroborate my hypothesis on the sensorimotor bases of music cognition.

Chapter 5

Exploring Musicality¹⁹

A movement is learned when the body has understood it, that is, when it has incorporated it into its “world”, and to move one’s body is to aim at things through it; it is to allow oneself to respond to their call, which is made upon it independently of any representation

- Maurice Merleau-Ponty, *Phenomenology of Perception*

Perceiving how things are is a mode of exploring how things appear. How they appear is, however, an aspect of how they are. To explore appearance is thus to explore the environment, the world. To discover how things are, from how they appear, is to discover an order or pattern in their appearance. The process of perceiving, of finding out how things are, is a process of meeting the world; it is an activity of skillful exploration.

- Alva Noë, *Action in Perception*

5.1 - At the basis of musical sense-making

The previous chapter aimed to reconsider different assumptions at the basis of human musicality, with particular regards to the notions of embodiment and musical affordance. As it emerged, moreover, the centrality of action for musical sense-making cannot be coherently addressed in terms of simulation-like mechanisms or fully external affordative properties that would elicit a particular action. Rather, without any sort of *reification* of musical subjects and objects, the enactive approach to music cognition allows a more interactive stance, based on the mutual engagement between living system and their world. The characterisation of musical affordance provided - defined as a property of the intentional relationship between cognizers and musical

¹⁹ This chapter contains material previously published as Menin, D., & **Schiavio**, A. (2012). Rethinking musical affordance. In *AVANT. The journal of the philosophical-interdisciplinary vanguard*, 3(2), pp. 202-215, and as **Schiavio**, A. (2013). Review: Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness, and Language. In *Philosophical Psychology*. Moreover, it contains material from the following articles in preparation: **Schiavio**, A., & Gerson, S. (in preparation). An enactive perspective on infants’ musicality, and **Schiavio**, A. (in preparation). Music, Evolution, Mirror Neurons.

environment - aimed to emphasise the direct (and intrinsically motor) fashion of the mutuality between musical subjects and objects.

An ideal area to provide a more detailed analysis of this interactive engagement is represented by musical learning and infants' musical behaviours, mainly because of their *minimal* and *basic* forms of intentionality. In this chapter, therefore, I will discuss the notion of *teleomusical acts* with regards to the development of musical expertise from a phylogenetic and ontogenetic point of view, mainly focusing on infants' affective and (inter)active, sense-making. If we reject the idea that human musicality is a pre-wired biological module, underlying specific (music related) mental states - as remarked in Chapter 2 - how could we possibly explain the emergence of human musicality in infancy? How can a child familiarise with rhythmic and melodic features? The pedagogical work provided by Delalande (2009), might reveal some fascinating features with regards to the *birth of musicality*, without positing neither an innatistic constraint nor a distinction between internal states and external fully defined categories.

The French psychologist invites us to admit that the child's ability to familiarise with the musical codes of her culture depends on her *practical engagement* with musical structures rather than mere exposition (or any kind of direct imposition). In order to facilitate this process, the work of the music teacher should be based on the relevant *motivations* that would allow the infant to employ different strategies that would direct her to achieve a *musical goal*. In other terms, a teacher should motivate the child to actively engage with musical structures through the development of OTAs. Bondioli (1996) defines three ways in which the instructor can do this; in particular he refers to (i) *mirroring* - where the adult reproduces the spontaneous sound-based plays of the child - (ii) *modelling* - in which the teacher helps the infant to reach a musical-directed goal - and (iii) *scaffolding* - based on the active interaction between the two in order to develop musical ideas. The relevance of these methodological strategies for enaction relies on their focus on the sensorimotor patterns of mutual engagement, abolishing the boundaries between mind and behaviour and between the self and the other (see Reddy, 2008).

From birth, in fact, humans are social animals that actively engage with the (social) environment. In contrast to previous assumptions that infants are passive observers of the "blooming, buzzing confusion" (James, 1890) surrounding them, recent evidence suggests that they are actively involved in interactions (that are not reducible to stimulus-response behavior) within the first several months of life (Bloom, 2001). The classic finding by Meltzoff and Moore (1977) that neonates will imitate facial gestures such as mouth openings and tongue protrusions was some of the first evidence that infants are engaged in social interactions from birth. Although the interpretation of these findings is still a source of controversy (see, for example, Anisfield et al., 2001), at the very least, they suggest that neonatal infants are *interactors* with other social beings from the onset.

More recently, Reddy and colleagues (2013) hypothesized that 2- and 3-month-old infants are more actively engaged in simple coordinated activities such as being picked up than previously assumed. They found that, when being approached by an adult, infants "help" adults by performing anticipatory body adjustments prior to the lifting act. This is just one of many examples in which infants actively engage with others in the first year. Similar instances occur when infants engage in turn-taking behavior during dyadic interactions or stiffen their leg when the caregiver is dressing them in the morning. Rather than focusing on whether infants are able to infer the intentions and plans of the caregiver at this stage or simply learn patterns of activity, the enactive account would focus on the infants' movements as the 'tools of her cognition' (De Jaegher & Di Paolo, 2007 p.

489). In particular, this mutual interaction between infant and caregiver portrays a legitimate form of sense-making, through the creation and appreciation of the basic meaning attribution process (De Jaegher, 2009). These goal-directed actions can hardly be considered mere reflexes or unintentional movements; rather, they describe the cognitive interaction through a process of mutual coordination. The autonomous agents involved in this interactive process present a regulated coupling, in which they both *bring forth* their own domain of meaning. As interpersonal engagement is a fundamental property of neonates' behaviors, it comes at no surprise that minimal forms of musical intentionality can be observed in young infancy, involving social and physical environment. Delalande (2009), hence, individuates three different musical conducts²⁰, each of them associable to the *phases of the game* postulated by Piaget (1964):

1) The *explorative conduct* is based on the discovery of sounds and noises, and corresponds with the *sensorimotor game*, which, according to Piaget (1964), dominates the first two years of life. After a few months from birth, the infant explores the *auditory possibilities* of the surrounding objects, using different sensorimotor dynamics and producing sounds that could potentially capture her attention (Imberty, 1983).

2) The *expressive conduct* corresponds to Piaget's *symbolic game*, characterising the years of primary school. Delalande assumes that in this period the child attributes to the sound some extra-musical values such as situations, roles, expectations and so forth, somehow enriching the primary form of sensorimotor intentionality with a broader domain of meaning attribution. The adult musicians would be immersed in the same dynamics through the emphatic gestural expressions used to convey further expressivity in concrete musical practice (Castellano et al., 2006; Leman et al., 2009).

3) The *organizational conduct* emerges when the child discovers the enjoyment of applying rules to her own musical games (phase of *game of rules*). In a musical context, this idea is best understood when considering joint musical practices and turn taking behaviours. These rules would also play a crucial role in practices such as musical analysis and composition, where the agent is employing a particular strategy - governed by a given rule. In this sense, even using *chance* as a compositional methodology is still a choice, a self-imposed rule²¹.

Any pedagogical work aimed to develop musical expertise, Delalande claims, should focus on these three interactive and sensorimotor conducts. Moreover, as the author interestingly points out, all of the most important stages of Western musical history could reflect these developmental phases, showing their deep continuity in ontogenesis and phylogenesis. Although this last idea is certainly fascinating, it does not seem to be relevant for our aims or falsifiable in any sense. Therefore, instead of embarking on a discussion about historical genres and musical forms, it would be better to have a closer look at the most basic motor behaviour described, in order to gain an insight into the minimal kind of musical sense making.

²⁰ A *conduct* is defined by a set of coordinated actions, relevant for a specific goal.

²¹ The reference here is for John Cage and his well-known use of *I Ching* for composition, during the Fifties.

5.2 - Explorations and teleomusicality

Consider the following scenario: a 10-month infant stands in front of an unknown object - say, a new doll. What is she going to do? Most probably, if the object would capture her attention, she will try to grasp it, act upon it or interact with it. In one word she will *explore* it. In this way, she will also discover some properties of the sounds provoked by her actions on the object. Having realized that sounds are dependent by the motor behaviours employed to produce them, the child will not simply use the same chain of acts every time she explores the environment. Sometimes she will squeeze an object; sometimes she will hit it (and so forth). Eric Clarke perfectly described these embodied dynamics when discussing the first encounter of a child with a xylophone: ‘the child’s more-or-less unregulated experiments with hands or sticks will result in all kinds of accidental sounds. With unsupervised investigation, the child may discover that different kind of actions [...] give rise to differentiated results [...], and even that these distinctions can themselves be used to achieve other goals’ (2005, p. 23). Delalande (2009) observes that when the attention of the infant is captured by the produced *sounds*, rather than by her own action, or by other (tactile, visual) properties of the objects, she would start playing with the sounds in a meaningful way. For example she can apply some basic rhythmic or even melodic variations.

This process emerges properly between 6 and 10 months of age. Before this period, when young infants see events in which some objects are targeted as goals for actions, they selectively encode the action and ignore object appearance (Bahrick & Lickliter, 2002). But by ten months, infants can encode both object appearance and the actions performed thereon (Perone & Oakes, 2006). Between six and ten months of age, in fact, both sensitivity to object appearance in the context of actions increases, and the ability to integrate information about object appearance and action emerges. Objects are viewed as rich in information that potentially specifies what they afford for action (Perone et al., 2008). As already hypothesised by Ruff (1982), infants’ exploratory behaviours are a necessary tool to understand the actual objects’ configuration, providing the child with a form of comprehension that does not rely only on visual perception.

Infants continually shape their movement spatially to the intentional urgings that prompt them to move. In a very real sense, they play with movement, discovering kinetic awarenesses and possibilities in the course of moving. Over time, they hone their movement to better effect - changing their orientation, for example, or the range of their movement. Their focus of attention is not on themselves as objects in motion, but on the spatiality of movement itself, what it affords and does not afford with respect to touching things that are near, grasping them, pulling them toward themselves, crawling toward those that are distant, pointing toward them, and so on. This experiential space, or better this *tactile-kinesthetic spatiality*, has nothing to do with measured or measurable distances but is an experiential dimension of movement itself. Space in this experiential sense is precisely *not* a container in which movement takes place but a dynamic tactile-kinesthetically charged created space (Sheets-Johnstone, 2010, pp. 171-172).

But how exactly is this manual exploration characterised? Which strategies are employed by the infants to explore the environment playing with sounds? Michel Imberty's research on children's musical improvisation (1983) showed that their musical production (both freely improvised or based on imitation) has two permanent features, named *pivot* and *colmatage*. While "pivot" refers to a stable and defined musical element, "colmatage" specifies the unstable and variable nuances of a musical production. These two sets of musical behaviours would require a particular attention to the sound properties rather than to the sensorimotor behaviours employed in order to obtain the sound. Therefore, *pivot* and *colmatage* emerge after (or during) the six-to-ten months' attentive shift. This means that this shift would allow the infant to create musical material, being *intentionally* directed towards the sounds without being focused on the actions performed. In other terms, it permits the constitution of a first musical context where the infant's goal is *intentionally musical*. We could name this basic form of intentionality as *teleomusicality*, considering the fundamental role of the goals of the action for its constitution. If the sound related to the object is now the goal of an intentional action, then the strategies employed by the infant are always directed towards new and meaningful horizons of sensorimotor possibilities.

The *explorative conduct*, hence, allows the infant to understand the relationship between the physical properties of the explored object and its practical opportunities for action (see Kemler-Nelson, 1999). As Delalande (2009) states, in fact, a truly musical behaviour would emerge only when the child's attention would shift from the sensorimotor patterns of actions to the sound itself. In other terms, according to Delalande, what matters for a minimal constitution of musical contexts is the goal (sound) rather than mere kinematics (movements). In this characterisation, the sound produced by these basic musical-directed actions (OTAs) would become the *noematic pole* of the intentional relationship with the musical object, constituting a basic form of musical intentionality.

Another cardinal idea of Delalande regards the continuous employment of the same behaviours for all the musical life. In this sense, the adult musicians' improvisation practices could be seen as a continuation of the ontogenetic processes of exploration, where the object intentionally explored is the relevant musical instrument (including the voice). As previously stated, this behaviour is essentially based on sensorimotor patterns of goal directed actions employed to produce a (series of) sound(s). It is put into effect through OTAs such as plucking, or hitting, creating a shared horizon of motor and musical goals which fosters concrete strategies to modify OTAs in order to further explore the musical environment. The child, therefore, actualizes OTAs, only after the six-to-ten months' attentive shift, being immersed in the musical dynamics without any mediation or inference that would spring from her attentiveness towards the action. Rather she will learn by means of her own situatedness in the world.

Development, and therefore learning, is essentially an endogenously self-generating process; it is therefore unnecessary - and impossible - to "instruct" it from the outside. This runs directly counter to the widespread notion that "learning" is a process of "instruction", by which is meant a process of information transfer from teacher to pupil (Stewart et al., 2010, p. 9).

Before the 6-to-10-month attentive shift, an infant has already familiarised with musical objects, as it emerges from the analysis of her *protoconversations* with the caregiver (Stern 1974; 1985).

However, the evidence for grades of conscious agency that depend on innate emotions of sympathy for other persons' rhythmic, "musical", patterns of intention does not prove an innate musical intentionality (see Trevarthen, 1999). We can only define those behaviours as musical-inherent *a-posteriori*, as they lack any intentional relationship with music itself. With regards to this point, Imberty (1983) invites us to notice that the infants' auditory system develops to a greater extent than the other perceptive mechanisms. The reason is probably that the voice of the mother should be recognized as soon as possible, providing the infant with the recourses necessary to further develop the ability to actively *interact* with the sounds' environment. This *affective tuning* (Stern 2004) relies on the cognizers' ability to meaningfully engage in mutual activities, involving parameters clearly employed in musical dynamics such as intensity, durations, form and so forth. Trevarthen (1997) states that the infant and the mother are united by a single rhythm, that is a turn-taking in a slow Adagio (a beat every 0.9 seconds), responding to each other in light of the prosodic characters of the meanings co-enacted. Dissanayake (2000) essentially confirms these ideas when suggesting that

the enjoyment and capacity of producing musical notes are faculties of indispensable usage in the daily habits of life of countless women, specifically mothers, and their infants, and that it is in the evolution of affiliate interactions between mothers and infants [...] that we can discover the origins of the competencies and sensitivities that gave rise to music (ibid., p. 389).

The author does not refer to actual infants' songs but rather to the trans-modal affective interactions between infant and caregiver. From a phylogenetic perspective, these active emotional patterns would have helped the neonate to improve her sensorimotor abilities. This is very important considering that very often in our evolutionary history, birth was premature (Leakey, 1994; Morgan, 1995). The acquired competencies, therefore, would develop in the next ontogenetic processes through the mutual interaction of the infant with the parents or with other infants (Dissanayake, 2000, p. 399), disclosing new horizons of sensorimotor creative and interactive activities. In other terms, although the mother-infant interaction described is a goal-directed and intersubjective motor behaviour involving many musical components, the realisation of musical sense-making should reflect a clearer intentional activity directed towards a musical goal. In this sense, OTAs developed after the 6-to-10-month attentional shift would allow the constitution of a new kind of intentionality (musical intentionality) that links the musical subject and the object in a co-constitutive relationship.

Simply put, the foundation of a *Music-directed set of acts* provides a child with the possibility to develop this new kind of intentionality, which is no more related to a *physical* object only or to the caregiver, as in the affective tuning. Motor acts like hitting or scratching (Delalande, 2009) make the child able to constitute a *musical context*. To sum it up, we would have not only OTAs and CTAs, but also a motor behaviour that only *a-posteriori* can be seen as *musical*. We can define this early (before six months of age) behaviour as *protomusical*, being directed only *inferentially* to a musical goal, therefore not yet constituting an intentional kind of musical sense-making. According to this perspective, after the attentive shift, the development of musicality stems from the unification of OTAs into CTAs in a way that would allow the agent to experience these acts in a

non-associative way. The ability to make sense of music, from early infancy, is thus defined by the agent's motor expertise rather than by the recruitment of fully blown mental representations. As Dreyfus puts it in fact:

A phenomenology of skill acquisition confirms that, as one acquires expertise, the acquired know-how is experienced as fine and finer discriminations of situations paired with the appropriate response to each. [...] Thus, successful learning and action do not require propositional mental representations. They do not require semantically interpretable brain representations either (Dreyfus, 2002, p. 367).

My point is that the outlined scenario portrays the origin of the musical vocabulary of acts of the agent who, by storing sets of OTAs, allows the constitution of CTAs leading to a more meaningful intentionality. This means that the abundance of these kinds of actions in the agent's own repertoire is the basis of the musician expertise, while OTAs that do not develop into CTAs would provide the agent with a superficial musicality without leading to a better constituted musical intentionality. The idea, however, can be elusive. So an example may help. When a child stands in front of an unknown object she would explore it, eventually producing sounds. However, if her attention would be captured by tactile or visual nuances of the object, then the exploration will be directed towards a different - non-musical - goal. The already acquired OTA would be still comprehensible in light of its basic categorization but the horizon of goals disclosed by the employment of the explorative behaviour would go beyond a musical context. It thus might be assumed that an original minimal musicality, which is actualisable in motor OTAs, would lead to a fully developed musicality, crystallizing into new CTAs. Musical ontogenesis, according to this perspective, is characterised by a progressive cross-modal association (sensorimotor but also perceptual-imaginative and visual-acoustic), corresponding to a better clarity in intending CTAs. This model offers a phenomenologically relevant interpretation of the development of musical expertise, stating that the constitution of a first musical context is immersed in the dynamics of action. In particular, this process reflects the idea that what matters for musicality is the goal of the motor behaviours, rather than abstract movements. The development of Constituted Teleomusical Acts, moreover, depends on the attentive, strategic and motivational behaviours of the infant, whose relationship with the musical objects can be modulated by meaningful social interactions with the caregiver or with the music teacher. That being said, it must be clear that this speculation does not exclude other modalities of meaningful engagement (i.e. exposure, dance, and so forth). Rather, it promotes the view that the *primary* way to constitute a musical intentionality is based on the spontaneous interaction between musical subjects and objects (Bannan & Woodward, 2009; Custodero, 2009). Furthermore, these circular dynamics of interaction are best understood without any reference to the standard input-output model, being constituted by co-determinative processes. An infant constitutes the musical context on the basis of her motor expertise, her motivations and goals, through her embodied interaction with sounds, whose *discovery* would eventually lead to other (more fluid and precise) forms of musical engagement. The following representation sums up the theoretical framework here developed:

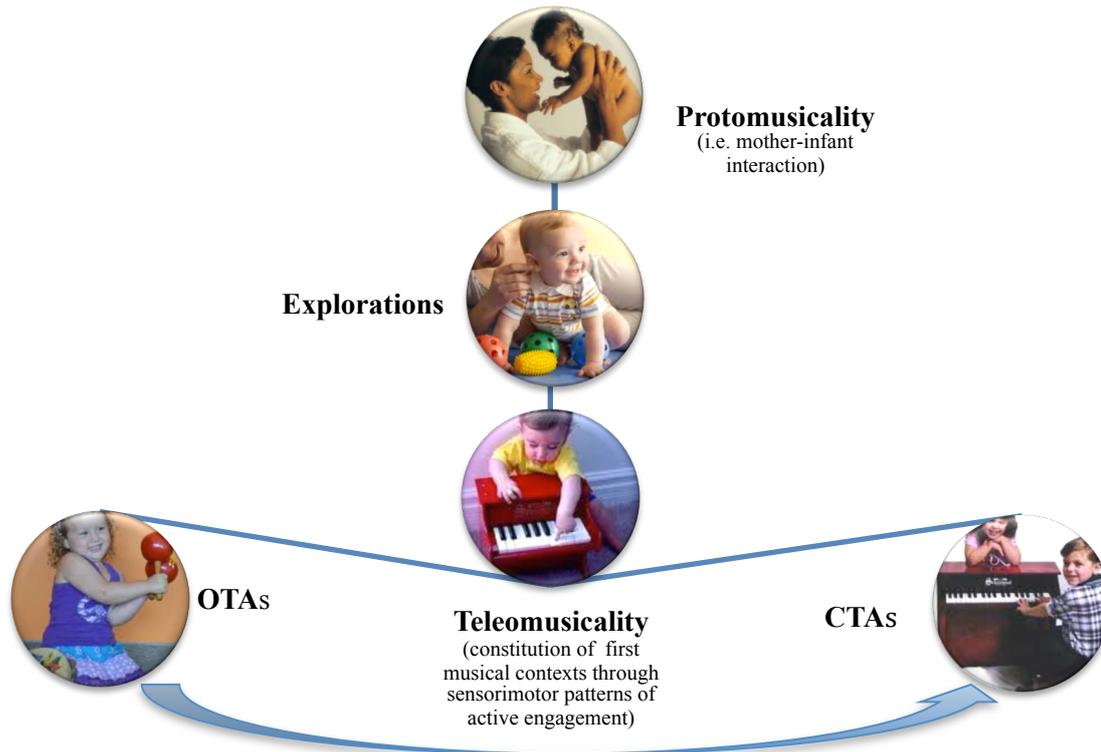


Fig. 5

This figure encapsulates my enactive interpretation of musical development. While affective tuning provides the infant with a superficial and non-directed musical expertise, explorative behaviours allow the constitution of teleomusical acts through the active engagement with the environment, modulating motivations and strategies employed by the infant to engaging in musical practices. Note that OTAs could develop into CTAs thanks to dynamical interactions with caregivers (or music teachers) or through self-organized (spontaneous) modalities of meaningful explorations.

This process of constitution, however, will never solve in a fully determined musical intentionality for the simple reason that musical intentionality is an *in-constitution category*. There will always be for the cognizer new meaningful ways to interact with, and constitute, the musical object. The horizon of sensorimotor patterns of mutuality is never fully defined. Musical practice does not simply end with the mastery of CTAs but, rather, it keeps on developing according to the strategies and motivations (see Sloboda et al., 1994) that the adult musician employs to engage with the musical environment.

5.3 - Interactive and affective sense-making in infancy

The role of early social interaction in infants' sense-making has been described by researchers and philosophers spanning an array of theoretical perspectives. For example, Meltzoff's (2005) "like me" hypothesis focuses on the links between the infant and co-actor's "body schema". That is, he suggests that there is an inherent connection between the bodily mappings of self and conspecifics that infants can act on from birth. In contrast to the hypothesis that there is a direct

mapping between self and other, Tomasello (1999) has put forth the idea that our motivation to engage with others and attempt to understand them plays a critical role in our understanding of others' intentional states.

This view more closely follows the cognitivist approaches, implying that infants come to understand others through analogy with their own internal states. Barresi and Moore (1996) emphasize the important role that interactions, and particularly triadic interactions (in which two individuals share attention on a common object), play in making sense of others' actions. They propose that the physical alignment between intentional relations (e.g., the relation between an infant and the object of his or her gaze or the relation between a mother and the object of her gaze) facilitates recognition of others' intentions. In this way, they take a step closer to enactivism by emphasizing the role of the interaction in the sense-making process.

The key here is to realize that because there is an individuality that finds itself produced by itself it is *ipso facto* a locus of sensation and agency, a living impulse always already in relation with its world. There cannot be an individuality which is isolated and folded into itself. There can only be an individuality that copes, relates and couples with the surroundings, and inescapably provides its own world of sense (Weber & Varela, 2002, p. 117).

Or, as Susan Oyama puts it, 'biological persons are constructed, not only in the sense that they are actively construed by themselves and others, but also in the sense that they are, at every moment, products of, and participants in, ongoing developmental processes' (2000, p. 180). Having established that social interactions are common, salient, and thought to be sources of cognitive and social development early in life, we now focus on empirical evidence of the action-perception link in infancy.

Growing evidence from both correlational and intervention studies suggest strong links between action production and action perception in the first two years of life. Infants' motoric capacities increase greatly in the first two years of life and beyond. For example, at around six months of age, infants become proficient at grasping objects within their reach. Near the end of the first year, they begin to produce object-directed points, engage in object-directed gaze, produce containment actions, and use tools as means to retrieve objects (see Clearfield & Thelen, 2001 for review of motor developments). In conjunction with these fine motor developments, infants become able to move themselves in self-propelled ways by crawling, cruising, and eventually walking. The natural progression of these skills allows researchers to assess relations between variations in these motoric abilities and infants' perception of these actions when produced by others.

Sommerville and Woodward (2005), for example, found that individual differences in ten-month-old infants' ability to produce a cloth-pulling action (in which they needed to pull on a cloth to retrieve a toy at the end of the cloth) was related to their perception of the goal of this action (the toy) in a habituation paradigm. Similarly, Woodward (2003) found that production of object-directed pointing was related to infants' perception of the relation between an agent and the object of his or her point. Interestingly, this relation was specific to the particular action in that pointing production was unrelated to perception of object-directed gaze (whereas engagement in shared

attention was related to gaze understanding). Similar relations between action experience and action perception have been found for gross motor actions on a neural level.

For example, van Elk and colleagues (2008) measured motor activity in 14-month-old infants (using electroencephalography) when they viewed videos of infants crawling and walking. At this age, as a group, infants had more experience crawling than walking and showed greater motor activity when viewing crawling actions than walking actions. Further, variability in walking experience (as measured in months since walking onset) related to the amount of motor activity when infants viewed videos of walking infants. Importantly, this effect of experience remained after accounting for differences in age, indicating that motor experience, rather than general development, was underlying these changes.

These findings are of particular interest because they suggest that motoric experience (with crawling), rather than visual experience (with walking), drove motor activity during the observation of gross motor movements. That is, infants likely had a great deal more experience observing walking than crawling, but motor activity changed as a function of motor, not visual, experience. Together, these findings support the notion that action production and perception are not separated, autonomous, categories as the standard cognitivist view posits. A quote by Merleau-Ponty can clarify this point:

The organism cannot properly be compared to a keyboard on which the external stimuli would play [...]. Since all the movements of the organism are always conditioned by external influences, one can, if one wishes, readily treat behaviour as an effect of the milieu. But in the same way, since all the stimulations which the organism receives have in turn been possible only by its preceding movements which have culminated in exposing the receptor organ to external influences, one could also say that behaviour is the first cause of all stimulations. Thus the form of the excitant is created by the organism itself. (1945, [1962], p. 13).

Correlational studies, however, do not allow causal claims to be tested. To determine causal directions, intervention studies are needed. A series of recent studies have done just this: given infants experience learning an action when they are at the brink of being able to perform that action and then tested the infants' perception of that action when produced by others. For example, at three months, infants are not yet proficient at effectively reaching for and grasping objects and they do not yet recognize the goal-directed nature of this action when performed by others. Sommerville, Woodward, and Needham (2005) fitted infants of this age with Velcro mittens that allowed the infants to pick up and move around toys (also covered in Velcro). After gaining this object-directed action experience, infants' recognition of the goal of a mittened grasping action was assessed in a habituation paradigm. Infants who had gained experience producing object-directed reaching actions using mittens, but not those who did not get this experience, then recognized the goal of this action when performed by an experimenter, as assessed in a habituation paradigm.

Research by Gerson and Woodward (in press) contrasted active mittens experience with visual experience observing mittened actions and found that active experience was uniquely beneficial for perception of a grasping action at three months. Similar findings demonstrate the effect of active, above and beyond observational, experience in two-step means-end actions at eight and ten months

of age (e.g., Sommerville et al., 2008). The unique effects of active experience at the origins of action perception are congruent with an embodied perspective of action in infancy. The above-reviewed findings, however, focus on motor experience gained as an individual, independent of social context. The enactive account, in contrast, focuses on the embedment of living bodies in the social world. Infants are engaged in interactions, not only with objects, but also with other people. Recently, Gerson and colleagues (Woodward & Gerson, in press) have put forth an account in which the knowledge gained through one's own self-produced actions (as described above) can then be used as the base for building upon this knowledge through social interactions. Similar to the effect of experience in triadic interactions described by Barresi and Moore (1996) above, Woodward and Gerson suggest that comparisons between self and other during social interactions can facilitate generalization of action understanding. In one study, they examined how comparison facilitated understanding and imitation of a novel action in seven-month-old infants (Gerson & Woodward, 2012). At seven months, infants can produce grasping actions and understand and imitate grasping actions when they view them performed by others. In contrast, they do not yet produce or recognize the goal of tool-use actions, in which an individual uses a tool as a means to retrieve an object. In this study, infants played a simple game in which the experimenter passed the infant a series of toys using a tool. During this interaction, infants could compare their grasping action with the experimenter's tool-use action and recognize that the two actions shared the same goal structure; that is, they both aimed to reach the same object. Following this interaction, infants recognized and imitated the toy-choice of the experimenter when she used a tool to reach one of two toys. Infants who received other kinds of experience with the tool and toys that did not allow them to compare the grasping and tool-use actions did not imitate the experimenter's toy choice in the test trials.

Thus, participation in the interaction facilitated comparison between actions and understanding and imitation of the novel action. Gerson (in press) suggests that comparison processes during social interactions can provide a rich source of information about intentional actions. Consistent with an enactivist perspective, sense-making takes place in the *meaningful* interaction due to the sensorimotor coupling that links embodied agents with their world. For this reason, the sphere of affection must be considered as a constitutive and fundamental part of (musical) sense-making and not treated as an autonomous category. Rather, affectivity is radically intermixed with perception and inter(en)action. There are two reasons for assuming this:

First, there is a large amount of anatomical overlap between the neural systems mediating cognition and emotion processes, and these systems interact with each other in reciprocal and circular fashion, up and down the neuraxis. Second, the emergent global states to which these interactions give rise are "appraisal-emotion amalgams", in which appraisal elements and emotion elements modify each other continuously. Such modification happens at each time-scale, through reciprocal interactions between local appraisal and emotion elements, and circular causal influences between local elements and their global organizational form (Thompson, 2007, p. 371).

In the realm of infant research, many scholars in the field of attachment theory have long discussed the value of early, everyday interactions between infants and their caregivers. As in the

enactivist account, they suggest that these social interactions provide the basis of subsequent affect, cognition, and behaviour. The affective component of these interactions is critical to accounts of attachment in that the feeling of security with a caregiver when faced with fear forms the basis of the attachment (Bowlby 1988). According to Bretherton (1990), infants create models of their world that act as “affective-cognitive filters”, and are created through representations of the self, attachment figures, and relationships. These models then influence a child’s actions and view of self in future relationships and interactions with social partners. Attachment theorists label these models “internal working models”, and they are often thought of as representations of the world, thus taking on a cognitivist perspective. Others, however, have proposed that the information contained in these models need not be cognitive and explicit. As Suomi puts it:

Cognitive constructions per se may not be necessary for long-term developmental or cross-generational continuities in attachment [...]. That is, such continuities are essentially “programmed” to occur in the absence of major environmental disruption and are in fact the product of strictly biological processes that reflect the natural evolutionary history of advanced primate species, human and nonhuman alike (Suomi, 2008, p. 186).

The proposal that affective and cognitive information is integrated through these interactions and influences future behaviour and interactions has been studied in adults and preschoolers using a variety of techniques, including interviews and storytelling games. In general, expressive emotional communication is a theme that should be tightly intermixed not only to an embodied and enactive perspective on early musicality, but also to human cognition in general. However, even an author like Piaget “largely ignored” (Trevarthen, 2012, p. 472) this aspect, as he mainly focused on the sensorimotor advances that characterise infants’ development. An enactive approach, nevertheless, considering the whole creature in its situated and autonomous self-organization, cannot avoid a discussion of the basic forms of sense-making. Giovanna Colombetti, for example, considered bodily arousal not as a mere ‘response to the subject’s evaluation of the situation in which she is embedded. It is rather the whole situated organism that subsumes the subject’s capacity to make sense of her world’ (2010, p. 15).

In general, the main idea is that the sphere of affectivity constantly emerges in the openness to the world that characterises a cognizer (Ratcliffe, 2009). The sphere of emotion cannot be simply bracketed out. Rather, it does act as a ‘transparent background that constrains and informs the features of the environment which show up for a perceiver’ (Ward & Stapleton 2012). This is true not only for the *cognizer-as-perceiver* but also for the *cognizer-as-embodied agent*, which ultimately are exchangeable notions. As Trevarthen (2012) states, in fact, ‘enactivism itself must allow that one must care about what one is doing, must have an emotional investment in the moving and its consciousness’ (p. 468). Let’s consider, for example, the semantic analysis of the terms “motion” and “emotion” provided by Bloem (2012). The author claims that Descartes’s philosophy might have influenced the processes related to the current use and comprehension of the word “emotion” (2012, pp. 416-417). Starting the analysis from the French verbs “é mouvoir” and “mouvoir”, Bloem posits that while “mouvoir” has been treated as a term to indicate physical movement, “é mouvoir” has a psychological connotation, thus separating de facto the mind from

the body, in league with the Cartesian dualistic stance. Before that, the two terms were used almost equally to express either physical movement or stimulation. In the following section, hence, I will discuss the enactive analysis of emotions and musical emotions, in order to provide further insight into the categories employed for musical sense-making.

5.4 - Enacting Emotion

The word “emotion” originally comes from the Latin verb *emovere*, which literally means an outward movement (Thompson, 2007, p. 363). This categorization is very similar to the very first meaning of *intentionality*, which we encountered in Chapter 3. Thompson (ibid.) provides a useful quote from Walter Freeman, in order to clarify the relation between these two terms:

A way of making sense of emotion is to identify it with the intention to act in the near future, and then to note increasing levels of complexity of contextualization. Most basically, emotion is outward movement. It is the ‘stretching forth’ of intentionality, which is seen in primitive animals preparing to attack in order to gain food, territory or resources to reproduce [...]. The key characteristic is that action wells up from within the organism. It is not a reflex. It is directed toward some future state, which is being determined by the organism in conjunction with its perception of its evolving conditions and its history (Freeman, 2000 [quoted in Thompson, 2007, p. 364]).

Similarly, a pianist’s motor behaviour is intentionally directed towards a musical goal, whose sense’s stratifications include affective and sensorimotor properties by its own definition. A pianist does not play mechanically, like a puppet; she does not spend hours and hours of practice to perform, simply by moving the fingers and the hands on the keyboard. Rather she *constitutes* the musical object by imposing a meaning, an intentional meaning that is emotionally driven and thus co-determined by the perceiver(s) in various aspects. As we previously saw, in fact, musical sense-making is a process that allows the specification of a musical environment by *enacting* or *bringing forth* the participants’ meaningful perspective.

This holds true for performers as well as for listeners. In a performance both of them *co-enact* a world (see McGee, 2005) in a completely different manner from how the cognitivist viewpoint would tend to portray this process. There is no need to postulate cognitive mechanisms for coping with perceptual structures resulting from individuals’ activity in an objective musical environment. The direct, yet neurally enabled, process of sense-making reflects the assumption that music is not a unidirectional entity that meets a passive listener. Meaning making involves the cognizer as an *embodied* and *affective* system, ‘constantly coping with the demands of being-in-the-world’ (Geeves et al., 2009, p. 107).

Perception and emotion therefore should be considered as *dependent* nuances of intentional motor behaviours. With regards to this point, Ellis and Newton (2012) recently claimed that an agent is an organism that meaningfully acts in a world with an affective-driven behaviour, mainly elicited by the affordances of the environment. As Trevarthen admits, indeed, ‘evidences from

neuroscience and neuropsychology abundantly demonstrates how skilful action, every kind of knowledge, and all the “mirroring” of the intentions and feelings of others, depend on core processes of “action imagery” and *specific emotional charge*’ (2012, p. 468, emphasis added). From this perspective, an emotion would be *distributed* over brain, body and environment, manifesting itself in the subjective evaluations of bodily sense-making.

This account of emotions as embodied in the circular dynamics of the sensorimotor loops (see Colombetti, 2014) goes against two typical models of affectivity: the cognitivist perspective (i.e. Solomon, 1976) and the bodily-feeling theories (i.e. Damasio, 1994). According to the classic intellectualistic view endorsed by proponents of cognitivism, an emotion depends on the belief about a certain state. In this sense, I can feel ashamed only if I would acknowledge that my behaviour was wrong. Realizing my mistake would therefore enable my feeling of shame. ‘If you don’t believe you did anything wrong, you will not feel ashamed. Shame can be caused by beliefs and cured by beliefs’ (Hutto 2012, p. 2). But could this model account for a complete story of emotional states?

Pure cognitivist accounts of the emotions are woefully incomplete in that they overlook the importance of feelings. [...]. For it is easy enough to imagine the relevant cognitions taking place in disembodied, entirely “cold”, “detached” and, wholly, “unemotional” ways. This observation is especially pertinent even if we consider quite ordinary cases. Familiarly, even when our emotions are stirred up by having certain beliefs or judgements, the associated feelings relating to such emotions can outlast the changes in those beliefs and judgements. Thus my wife’s seething anger and feelings of outrage at what she takes to be my transgression might not immediately subside upon discovering me innocent of that of which I am accused. While, as a consequence of this discovery, she may no longer direct anger at me (and certainly not with justification) it does not follow that her feelings of anger entirely dissipate, and in such circumstances, it is the having of such feelings that seems sufficient for her being in a continued emotional state of anger (ibid.)

On the contrary, a theory of emotions as body-feelings, like the one proposed by Damasio (1994; 2003) assumes that emotional consciousness is not a matter of beliefs but, rather, depends on receiving interoceptive information (a signal from the *inside* of the animal) aimed at maintaining the agent’s homeostasis (the somatic marker hypothesis). Many authors, however, reject the idea of an efferent signal that ‘travels towards the central nervous system from more remote areas’ (Ellis & Newton, 2012, p. 62). Damasio’s view ‘provides no account of how conscious information processing is different from a robot’s passive receiving of inputs and then transforming them into behavioural outputs’ (ibid.). In contrast, the focus on participatory sense-making (De Jaegher & Di Paolo, 2007) reveals a more ‘complex interplay of degrees of connectedness’ (Colombetti & Torrance, 2009, p. 17) that is never a matter of solipsistic inner events but, rather, it emerges in intersubjective practices. As Hutto puts it:

Such theories [bodily feelings theories] are problematic in that they “have little to say

about the processes by which external stimuli are evaluated for ecological and social significance” (Hill, 2009, p. 199). Basically, in reducing emotions to bodily feelings or perceptual states exclusively targeting such, “somatic theories have trouble explaining what it is for an emotion to have an intentional object or target” (Hill, 2009, p. 200). They lack appropriate reach (Hutto, 2012, p. 2).

The idea that sensorimotor patterns of motor loop between organism and environment would modulate the emotional meaning enacted by each organism implies that any *isolationist* proposal would face difficulties in integrating such mutual processes for understanding (musical) emotions. While proponents of the cognitivist framework would in fact describe the emotional sphere as a top down mechanism, the enactivist will consider “top down” and “bottom up” as ‘heuristic terms for what in reality is a large-scale network that integrates incoming and endogenous activities on the basis of its internally established reference points’ (Thompson, 2007, p. 366).

In musical contexts, the subject of emotion is probably one of the most important, having received much interest in behavioural, electrophysiological, and theoretical studies (i.e. Juslin & Sloboda, 2001). Among others, Peter Kivy (2001; 2007) argues, for example, that our emotional response to music depends on our ability to identify certain sounds’ qualities that would elicit a particular feeling. According to this traditional view, therefore, a listener would be affectively moved by a particular piece, thanks to the way in which music represents, or portrays, a particular emotion. This idea, which eventually recalls Huron’s model of expectations, heavily relies on the agent’s past listening experiences - since they could facilitate the tendency of the musical properties to evoke an emotion in the listener. It is quite clear, however, that this perspective is based on the input-output dichotomy, as well as on the attribution of given qualities to the musical object, thus objectifying the two poles of musical intentionality. The input-output commitment is a die-hard assumption not only in music cognition, but also in the broader field of neuroscience. As Thompson puts it, in fact, ‘traditional neuroscience has tried to map brain organization onto a hierarchical, input-output processing model in which the sensory end is taken as the starting point. Perception is described as proceeding through a series of feedforward or bottom up processing stages, and top-down influences are equated with back-projections or feedback from higher or lower areas’ (2007, p. 366). Endorsing such a view would make us lose any chance to consider the brain as an autonomous system (Freeman, 1999).

Self-determining or autonomous systems [...] are defined by their organizational and operational closure, and thus do not have inputs and outputs in the usual sense. For these systems, the linear input/output distinction must be replaced by the nonlinear perturbation/response distinction. From an enactive perspective, brain processes are understood in relation to the circular causality of action-perception cycles and sensorimotor processes. Hence emotion is not a function in the input-output sense but rather a feature of the action-perception cycle - namely, the endogenous initiation and direction of behaviour outward into the world. Emotion is embodied in the enclosed dynamics of the sensorimotor loop, orchestrated endogenously by processes up and down the neuraxis, especially the limbic system (Thompson, 2007, p. 365).

A musical emotion therefore is not an autonomous category, a property of auditory feedback that would be actualised in the mind. Rather, it is a constitutive part of the holistic process of musical sense-making, being distributed over brain, body and musical environment, thus playing a fundamental role in the sensorimotor loop characterising musical intentionality's mutual specification. Once again, the distinctions between internal and external and between subjective and objective must be reconsidered:

Cognition is a form of embodied action [...] the enactive approach implies that we need to move beyond the head/body and subjective/objective dichotomies that characterise much of emotion theory. Appraisal is not a cognitive process of subjective evaluation "in the head" and arousal and behaviour are not objective bodily concomitants of emotion. Rather, bodily events are constitutive of appraisal, both structurally and phenomenologically (Colombetti & Thompson, 2008, pp. 56-58).

A post-Cartesian cognitive science of music, therefore, should take into serious consideration the challenge posed by the primacy of corporeality in cognitive and affective economy. Empirical frameworks aimed to shed light on the nature of musical emotion should hence consider embodied agents and their environment as inseparable systems, without any reification of subjective and objective categories.

If a living system is best understood 'by identifying the significant variables that constitute its behaviour, both inside and outside the head' (Rockwell, 2005, p. 197), then the study of musical emotion should not be limited to the understanding of the (inner or external) mechanisms that would provoke a feeling. Rather, the direct coupling between performers (or listeners) and their musical world implies a primacy of action, which is already an *affectively enacted* category. As Varela and colleagues remind us:

[t]he enactive approach consists of two points: (1) perception consists in perceptually guided action and (2) cognitive structures emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided. The overall concern [...] is not to determine how some perceiver-independent world is to be recovered; it is, rather, to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually guided in a perceiver-dependent world (Varela et al., 1991, p. 173).

In this sense, the proposal to reconceptualise emotions here described is only a sketch and simply argues for a more phenomenologically relevant working methodology for the analysis of musical emotions, for example in league with dynamic systems theory's techniques (see Garvey & Fogel, 2008).

5.5 - Phylogenesis and ontogenesis

The study of infants' musicality and musical emotions, being directed towards the basic forms of mutual engagement between musical subjects and objects, implicitly asks us to investigate other possible *antecedents* of musical behaviours. A biologically plausible interpretation of human musical behaviours, in fact, should take into consideration also the mutual constraints between ontogenesis and phylogenesis, without any strong distinction between the two (Tomasello 1999). Cross (2003) suggested that musical behaviours in infancy could be seen as a cognitive and motor play leading to a decreasing of social tensions inside a group of individuals, confirming the idea of Robin Dunbar (1996), who claimed that a complex vocal exchange - without any clear communicative significance - could have been used among our ancestors as a "social lubricant". Phylogenetically, it is useful to note that group activities are fundamental also for chimpanzees (Boesch, 2005). For example, during hunting activities, although there is no proper "role division", every animal tries to maximize its chance to capture the prey and the group responds to each animal's spatial position (Tomasello, 2008). According to the classic perspective, chimpanzees do not co-operate as humans, as they do not understand the goals of their partners (Povinelli & O'Neil, 2000). The discovery of mirror neurons, however, contributed in changing this assumption as they show how monkeys intentionally respond to goal-directed motor behaviours (Rizzolatti & Sinigaglia, 2008). Moreover, chimpanzees are also aware of the others' own perceptions, considering that during food competitions they take into consideration whether the rival sees the food or not (Hare et al., 2000), sometimes trying to get closer to hidden food (Melis et al., 2006).

Chimpanzees, therefore, are able to understand the others in terms of their goals and perception (Tomasello, 2008). According to Kirschner and Tomasello (2009a; 2009b) these nuances seem to overlap with human musical behaviours, leading the two scholars to integrate the studies on primates with research on humans. For example, they demonstrated that 4-year-old humans who shared a musical experience with some partners, tends to behave in a more social and cooperative way. If the hypothesis is that some kind of protomusicality could improve social bonds inside a group, it remains unclear which kind of concrete musical experiences are involved. Merker (2000) assumes that synchronisation plays the cardinal role, given that protomusical behaviours such as affective tuning are usually characterised by regular beats. In 1974 Fraisse noted that the easiest rhythmic impulse to synchronise with seems to correspond to the spontaneous timing that is the subjective velocity employed when performing natural rhythmical movements - such as leg-swinging - providing an argument in favour of Merker's hypothesis. According to the latter, in fact, these kinds of basic synchronisation would have led groups of primates to perform synchronised vocalisations with functions related to territory's defence and courtship (Tomasello, 2008). Most of the monkey's species display a basic shared repertoire of vocalisations, suggesting an actual predisposition for shared vocal activities.

On the ontogenetic side, Delalande (2009) observes that the explorative behaviour previously described develops from basic sounds' discoveries to the mastery of more meaningful affective and sensorimotor conducts underlining musical sense-making. With regards to this point, it is important to note that the ontogenetic differences in the adoption of sensorimotor patterns - as in the elaboration of short phrases - led researchers to postulate the existence of a *personal style* since the very first months after birth (Imberty, 1983). Repetitions or variations of given OTAs or CTAs, in this sense, would represent a basic morphology of musical sense-making, together with the

development of body schema (Freschi, 2006). As observed by Overy and Molnar-Szakacs (2009), moreover, the capacity of musical subjects to synchronise and interact might improve their ability to predict events. In their words:

At a basic level, this is evident in the simple pulse that underlies most musical behaviours, which is highly predictable and allows for spontaneous, enjoyable synchronisation (e.g., group clapping, dancing). At a higher level, it is evident in our strong emotional responses to, and preferences for, extremely familiar music (e.g., Peretz et al. 1998), a phenomenon that seems to appear in early childhood and continue throughout adulthood. It might appear that this proposal contradicts the classic theory of musical expectancy, which suggests that emotional responses occur to unexpected features in music (Juslin, 2001; Steinbeis et al., 2006; Meyer, 1956). Rather, we suggest that the capacity for music to create such a strong environment for minimized prediction error (and resultant affect) provides the very basis for a strong emotional response to an unpredicted event. While familiar, predictable music can be enjoyed to its fullest, the violation of expectancies can be more emotionally dramatic, as evidenced, for example, in the Romantic era of classical music (ibid., p. 494).

The distinctive aspect of synchronisation seems to be its empathic nature. While a study by Himberg (2006) shows that human adults tend to better synchronise with a human partner rather than an audio beat coming from speakers, Kirschner and Tomasello (2009a) showed that young infants are facilitated in synchronisation tasks (like clapping the hands to a beat) in intersubjective events. It comes at no surprise, then, that Walter Freeman provides a definition of music as ‘biotechnology of group formation’ (2000, p. 419). Although this definition appears too exclusive, it stresses the importance of intersubjective practices and their role in phylogenesis, creating and modelling social bonds:

Here in its purest form [music] is a human technology for crossing the solipsistic gulf. It is wordless, illogical, deeply emotional, and selfless in its actualization of transient and than lasting harmony between individuals [...] and perhaps even among higher apes [...]. It constructs the sense of trust and predictability in each member of the community, on which social interaction are based (ibid., p. 420).

Many theories have tried to shed the light on the origins of music from a phylogenetic standpoint. According to Antonio Serravezza (1996) the first evolutionary theories of music have been provided by Herbert Spencer (1857) and Charles Darwin (1871). For the British philosopher the birth of music was a necessary evolutionary adaptation aimed to trigger stronger emotions than language. In particular, first forms of affective human vocalisations would have induced emotions, being characterised by the use of cadences and vocal inflections (Serravezza, 1996, p. 232). In this scenario, music developed as a tool to communicate emotions, independently from, and more

strongly than language. Moreover, Spencer argued that musical sounds²² have a natural relationship with the organic conditions from which they are produced and consequently they are representative of a defined mood, physiologically linked to us through a sort of associative memory (ibid., pp. 238-239). Hence, the phylogenetic value of music is highly important, as it would have facilitated social cohesion through its feature of being a *medium* of physiological moods. While the theory provided by Spencer is based on a physiological standpoint, the perspective adopted by Darwin is genuinely biological.

Although Darwin himself considered his hypothesis on music as “exactly opposite” (Darwin, 1871, [quoted in Serravezza, 1996, p. 239]) from Spencer’s approach, in his work *Expression of the emotions in man and animals* (1872) his analysis appears to some extents similar to the one provided by his intellectual rival. Darwin, considered the *voice* as the actual root of the developing process of music, because - he claimed - every musical feature is reducible to chant. In fact, ‘primeval man, or rather some early progenitor of man, probably first used his voice in producing true musical cadences, that is in singing, as do some of the gibbon-apes at the present day’ (1871, p. 133). Despite this affinity, however, the two theories effectively present some differences: while Spencer claims that music is an emotionally intensified way of communication, Darwin considers this argument as too reductive as it can be applied only to very simple contexts (e.g. a scream for help must be very high and intense) and it might be inappropriate considering that, often, the emotions provoked by music are more connected to the laws²³ of sounds’ associations than to the features of single vocal modification (Darwin, 1872). Their disagreement, according to Kivy (1959), reveals their different philosophical approaches as Darwin criticizes the *deductive* argument developed by Spencer from a very general premise. In Darwin’s view, in fact:

The impassioned orator, bard, or musician, when with his varied tones and cadences he excites the strongest emotions in his hearers, little suspects that he uses the same means by which his half-human ancestors long ago, aroused each other's ardent passions, during their court-ship and rivalry (Darwin, 1871, pp. 942-943).

From his perspective, music can be considered as *seductive* tool for reproduction (see also Miller 2000), indeed ‘[b]esides this natural means of selection, by which those individuals are preserved, which are best adapted to the place they fill in nature, there is a second agency at work in most bisexual²⁴ animals tending to produce the same effect, namely the struggle of the males for the females. These struggles are generally decided by the law of battle; but in the case of birds, apparently, by the charms of their songs, by their beauty or their power of courtship’ (Darwin F., 1909). The proof of this argument, for Darwin, lies in the *economy* of nature itself. If female birds were unable to appreciate their partners’ voices, then all the effort to produce sounds would be simply useless (Darwin, 1871, p. 599; see also Fisher, 1915).

²² He mainly refers to chant and singing sounds.

²³ No reference to tonal harmony has been provided by Darwin.

²⁴ Peter Kivy (1959) writes about this term that ‘a certain ambiguity here with regard to the term *bisexual*. Darwin is clearly using the term here to refer to species in which the male element is present in one individual and the female element in another. However, modern scientific dictionaries give *hermaphroditic* as the synonym for *bisexual*. In this sense, a bisexual animal is one with both male and female sex elements. Darwin must have seen this problem for, when this passage appeared in print in 1858, the term *bisexual* was replaced by the term *unisexual*’.

The Darwinian hypothesis of music, however, should be integrated with some clarifications. First, it cannot pretend to be *exclusive*: the fact that some birds use vocal sounds for courtship does not necessarily imply that the complexity of the musical phenomenon has all its substrates in it. And, second, even if it may appear granted, we consider birdsongs as “songs” only *a posteriori* for they seem similar to what we think music is. In fact, Fitch (2006) noticed that there are at least two conditions that made the birdsongs relevant to an evolutionary enquiry on music: they are ‘*complex and learned* vocalisations’ (ibid., p. 182). In literature, many studies have investigated the learning abilities of birds (Marler, 1970b), while the study of the vocalisations’ complexity - apart from aesthetical purposes (Messiaen, 1944) - is considered as a very difficult argument (Weng et al., 1999) considering the ambiguity of the used concepts. Though Scholes (1938) considers birds as ‘the only masters in nature for men’ (ibid. p. 107), the discovery of complex vocalisations by marine mammals has at least put some doubts on this common conviction (Payne & McVay, 1971). Even if a comparative methodology has some advantages in phylogeny, by detecting common traits among different species and formulating inferences on our ancestors’ adaptive functions (Fitch, 2006, p.181), in the case of music it appears insufficient for the development of a coherent hypothesis. Hence, the similarity between the bird syrinx and the human larynx is not enough to draw a definitive conclusion about the birth of musicality and Darwin did already notice that the birds’ sound production can be considered as *analogous* and not *homologous*²⁵ of the human songs assuming that our common ancestor, a reptile from the Palaeozoic era, was not able to sing. Furthermore, according to the definition provided by Fitch (*complex and learned* vocalisations) the only primate capable of singing should be man. In fact, despite the complex vocalisations of similar animals, like gibbons (Geissmann, 2000), there is no evidence about their actual learning (Janik & Slater, 1997).

Enamoured by these argumentations, and trying to defend the uniqueness and the specificity of music as human traits, Hauser and McDermott published a famous study with the title *The evolution of the music faculty. A comparative perspective* (2003). Their argument, considered also by Fitch (2006, pp. 183-184), is based on three different points.

1) The first assumption consists of defining the animal repertoire as very limited in comparison to the human musical production, because it is used only for courtship or territorial defence. Unfortunately, those are not the only contexts in which these behaviours emerge. Many birds do sing alone and not only during the reproduction season. Furthermore, following the example provided by Fitch (ibid.), a different role for their sound production cannot invalidate the *analogy*, otherwise we should not consider as analogous the wings that a specific species uses for preying from the ones of another species used for flying.

2) The second argument provided by McDermott and Hauser is based on the consideration that human music is founded on pure delight, while animals’ proto-musicalities have *communication* as their primary function, considering the prevalence of intersubjective contexts. This analysis, however, turns out to be problematic as it assumes a predefined and a-problematic

²⁵ In biology this distinction is aimed to separate all the common traits of a species by virtue of a common ancestor, from the similar features independently developed across evolution.

notion of music. For example, Merriam (1964) noticed that aesthetic enjoyment is only one of the functions music fulfils²⁶ (Giannattasio, 1998, p. 209).

3) The last point used by Hauser and McDermott to refute the analogy between human and animal music assumes that in most of the animal species, the *singer* is mostly a male while, for humankind, there is no sex preference for sound's production activities. Actually, according to Langmore (1998; 2000) and Riebel (2003) many female birds can sing and, from a socio-cultural perspective, we also notice how in some human cultures only men are allowed to sing (Titon et al., 1984).

One aspect that emerges from this discussion is that all these approaches seem to disregard the role of ontogeny, focusing on an *exclusivist* way to investigate music. As Tomasello (2008) states, indeed, in phylogeny nature selects the ontogenetic paths, which lead to the mature phenotype and the Darwinian approach is a dynamic position where the different factors are not located in a static and a-temporal present. The genetic variation and the natural selection operate across time, with different processes in different modalities (Tomasello, 1999) and the primates' (humans or not) ontogenetic ways would have not developed without external materials and information. If mankind has the ability to learn maths, language or music, it does not mean that this ability is pre-determined in our genes. As Maturana and Varela put it:

Note well that innate behaviour and learned behaviour are, as behaviours, indistinguishable in their nature and in their embodiment. The distinction lies in the history of the structures that make them possible. Therefore, our classifying them as one or the other depends on whether or not we have access to the pertinent structural history. We cannot make that distinction by observing the operation of the nervous system in the present (Maturana & Varela, 1992, pp. 171-172).

For example, in the case of music, this point emerges from the variety of the ways of constituting musicality across cultures; several and with various shapes, those different paths show how musicality is ontogenetically constituted, rather than genetically rooted. Enhancing the positive role of behaviour in evolution would mean amending the notion of adaptation from the feature of passivity often bestowed from Neo-Darwinian adaptationism (Ferretti, 2007, p. 74). In other terms 'every ontogeny occurs within an environment [...] it will become clear to us that the interactions (as long as they are recurrent) between [organism] and environment will consist of reciprocal perturbations. [...] The results will be a history of mutual congruent structural changes as long as the organism and its containing environment do not disintegrate: there will be a structural coupling' (Maturana & Varela, 1992, p. 75).

Since *ontogeny* is defined as 'the history of maintenance of [the system's] identity through continuous autopoiesis in the physical space' (Varela, 1979, p. 32) then it seems that the phylogenetic roots of musicality reflect the ontogenetic processes at the basis of infants' sound discoveries, in their *bringing forth* a world. The musical environment in which we are embedded is

²⁶ For example, we can think about social or religious ceremonies, or about music therapy (Rouget 1968).

at the same time an environment *in constitution*, never fully given, and a category co-determined by the mutual influences that occur phylogenetically (in terms of different musical cultures and styles developed through time) and ontogenetically (in terms of basic sensorimotor modalities of actively engaging with it). Simply put, as musical subjects, we are (in) the musical world that we enact.

Chapter 6

From the cradle to the stage²⁷

It is precisely this emphasis on mutual specification that enables us to negotiate a middle path between the Scylla of cognition as the recovery of a pregiven outer world (realism) and the Charybdis of cognition as the projection of a pregiven inner world (idealism)[...]. Our intention is to bypass entirely this logical geography of inner versus outer by studying cognition not as projection or recovery but as embodied action

- Francisco Varela, Evan Thompson & Eleanor Rosch, *The Embodied Mind*

Living systems are units of interactions; they exist in an ambience. From a purely biological point of view they cannot be understood independently of that part of the ambience with which they interact: the niche; nor can the niche be defined independently of the living system that specifies it

- Humberto Maturana, *Biology of Cognition and Epistemology*

6.1 - Crossing the boundaries

In this final chapter I will draw from the conclusions which emerged in the last section, trying to further elaborate the enactive approach to musical experience through empirical corroboration. In particular, I will present two experiments that my colleagues and I have recently run. In the first one, we focused on the sensorimotor integration underlying an infant's engagement with a musical instrument, and the resulting facilitation in detecting action-effect contingency in musical events. In the second one, we investigated cross-modal processes of active engagement between adult subjects and musical objects, through relevant learning tasks involving atonal music. Taken together, these two empirical studies corroborate some of the main ideas developed until now in this work. In particular they address themes related to the development of musical expertise from early infancy, focusing on the sensorimotor processes underlying the engagement between musical subjects and objects. Musical experience, as I have argued, does not possess an *objectivistic* subject-object

²⁷ This chapter contains material from the following articles in preparation: **Schiavio, A.**, & Gerson, S. (in preparation). An enactive perspective on infants' musicality, and **Schiavio*, A.**, Gerson*, S., Timmers, R., Bekkering, H., & Hunnius, S. (in preparation). Sensorimotor integration and action-effect contingency in infants' engagement with a musical instrument. (*both first author), and **Schiavio, A.**, & Timmers, R. (in preparation). Cross-modal facilitation in learning atonal melodies.

structure but it is rather constituted by a subject-object inter(en)action, immersed in skilful activity. In this sense, musical cognition is a distributed category, emerging from the circular dynamics of brains, bodies, and musical environment.

However, enamoured by the methodological progress of the cognitive neuroscience of music, a large number of researchers prefer to focus only on the *brain mechanisms* involved in musical experience. Recent findings in these areas concern cortical and subcortical neural pathways underlying musical emotions (Juslin et al. 2008), brain plasticity induced by musical training (Schlaug et al., 1995; Bangert & Altenmüller, 2003), premotor response to musical stimuli (Brown & Martinez 2007), and the activation of motor areas during imaginative musical tasks (Kristeva et al., 2003; Lotze et al., 2003). Although most of these studies assume that musical listening, learning, or emotions are actual achievements of an isolated brain, some of the data reported admit a different interpretation, closer in scope to the enactive perspective. In particular, the study of developmental processes clearly highlights the importance of sensorimotor interactions and active explorative behaviours for cognition (Majewska & Sur, 2006). Cortical plasticity in somatosensory and motor systems, for example, seems to reflect the adult musician's mastery of relevant sensorimotor skills, showing the interdependency of functional architectures and active musical behaviours (Münste et al., 2002). In the context of visual research, moreover, it is generally accepted that an agent's forms of interaction with the environment crucially determine the development of neural circuits constituting her visual system (Held 1965).

With regards to this point, we could linger on the work of the American neuroscientist Paul Bach-y-Rita. As reported by Alva Noë (2009), Bach-y-Rita's research on neuroplasticity led to the empirical development of the famous concept of *sensory substitution*. In particular, the scientist developed a device that would allow blind people to "see" again. How? His view holds that the eyes simply are conduits that lead perceptual information to the brain (Bach-y-Rita 1967; 1972). Therefore, changing the pathway would not substantially change the content of the information that would reach the visual system. Indeed, his work was not focused on the "eyes" of an individual but, rather, on the possible alternative ways to convey relevant sensory information to the brain. In order to do this, he developed the famous *sensory substitution* tool named TVSS (tactile-vision substitution system), in which the visual input is captured by a camera and converted into a 'tactile stimulation on a twenty-by-twenty matrix of vibrators that a blind person, for example, could wear on her abdomen' (Myin & O'Regan, 2009, p. 193) or other parts of the body.

According to Evan Thompson, 'TVSS brings about quasi-visual experience of being able to perceive the shape and the movements of objects at a distance only if the subject is able to exercise active control of the camera and thereby integrate it into his or her sensorimotor repertoire' (2007, p. 255). In terms of brain mechanisms, it is possible to postulate that this sort of specific cortical areas' re-mapping would allow people to compensate for the lost perceptual modality. Brain imaging techniques show that congenital blind patients present a cross-modal activation of the occipital cortex during perceptual tasks concerning other senses, including sound localization and recognition (Renier & De Volder, 2005; Amedi et al. 2007). Over the last three decades visual-to-auditory sensory substitution devices like vOICe (Meijer 1992), or open source software like VIBE (Auvray et al., 2005) aim to provide the patients with a visual experience by using sounds. In these cases, research on brain areas indicates not only neural activity in the lateral occipital cortex, but also functional improvement thanks to adequate training sessions (Proulx et al., 2008). It comes as no surprise, therefore, that vision is not the only domain in which sensory substitution is applicable. Recently, in fact, Schurmann and co-workers (2006) examined the consequences of applying the

same paradigm to auditory conditions, showing that tactile sense could generate activation of the auditory cortex. It could be argued, therefore, that for each perceptual domain there is a corresponding pattern of ‘sensorimotor interdependency that is constitutive of that modality’ (Thompson 2007, p. 257).

Contrary to behaviourism, perceptual experience, being in part constituted by endogenous knowledge (skilful mastery) mediates between sensory stimulation and motor behaviour. Contrary to cognitivism, however, experience does not intervene in a linear causal fashion between sensory “input” and motor “output”. Sensory stimulation does not *cause* experience in us, which in turn cause our behaviour, because “skilful activity (consisting in behaviour and sensory stimulation) *is* the experience” (O’Regan & Noë, 2001b, p. 1015). In other words, as a skilful activity of the whole animal [...], perceptual experience emerges from the continuous and reciprocal (non-linear) interaction of sensory, motor, and cognitive processes, and is thereby constituted by motor behaviour, sensory stimulation and practical knowledge (ibid., p. 256).

Although Bach-y-Rita himself believed that his device would allow sensory substitution by means of the transmission of a new kind of information to the brain, which would eventually transform it into a representation, the surprising results obtained by his TVSS allow a different interpretation. In particular, as acknowledged by Noë (2009), Bach-y-Rita’s research shows that it is possible to constitute a relation between an individual and an object where there was no relation before. It suggests that what matters for our experience is not the neural activity inside the boundaries of the skull. Rather, it is the dynamical interaction with the objects of the environment. We “see” with our skin, with our ears or with our bodies because perceptual activity is an active, explorative activity. It is not a matter of symbolic computation.

Again with regards to the visual domain, a classic argument in favour of the enactive approach comes from the studies by Held and Hein (1963). In particular, proponents of enactivism consider their work as a firm proof that vision would not develop normally without motor activity. As reported by Jesse Prinz (2006b) Held and Hein raised two kittens in total darkness, without allowing them to see the light. However, only for a few hours per day the animals were conducted into a room with adequate illumination. Inside the room, one of the kittens was allowed to move and explore the environment while the other one was suspended over the ground in a cage, hence without any possibility to actively explore the room. After this habituation phase, the two animals were freed and the researchers noted that while the kitten that was allowed to move had normal visual abilities, the other one ended up being impaired. In particular, the latter was not able to ‘use vision to direct its paw reaching behaviours and it was indifferent to visual cliffs. Noë [like other spokesmen of the enactive approach] suggests that this experiment supports the view that visual perception necessarily involves understanding of motor responses’ (ibid., p. 9).

In the context of musical research, this discussion influenced Alicia Peñalba Acitores (2011), who indeed considers her recent work as an application of Kevin O’Regan and Alva Noë’s *sensorimotor contingency theory*, where ‘the knowledge of the ways movements affect sensory stimulation is necessary for experience’ (O’Regan & Noë 2001a, p. 1055). Her perspective assumes that bodily movements (such as the rotation of the head or movements of the whole body closer to

the sound source in order to increase the amplitude) and the knowledge of sensorimotor contingencies guiding them are crucial for auditory perception. As Joel Krueger puts it:

Bodily gestures are a form of attentional focusing and the vehicle of perceptual construction. The animate body becomes a vehicle for voluntarily drawing out features of a musical piece (e.g. expressive aspects, rhythmic beats, or melodic progression) and foregrounding them in our attentional field. This drawing out is an enactive gesture in response to felt affordances within the music. The listener perceives the inner space of the piece as a space that can be entered into, experientially, and by doing just this shapes how the experiential content of the piece-as-given becomes phenomenally manifest. In short, “we hear what the body feels” (Philips-Silver & Trainor, 2007 p. 544). What the body feels are sensorimotor contingencies-possibilities for interaction, movement, and coordination that determine the character and content of musical experience. Sensitive music listening is thus a kind of skilled coping with a sonic world (2013, p. 187).

Mark Reybrouck (2005b) captures this idea, defining music as a “tool for adaptation to the sonic world” or as something that is “heard and enacted upon” rather than a merely represented category. The *musical brain* is a brain in action, inconceivable without a body and a world. It is only in this condition that the brain enables embodied agents to coordinate their interactions with the environment. The field of musical learning, for example, represents an excellent domain (Palmer & Meyer 2000) to shed further light on the *motor basis* of this interaction, in league with the idea that musical experience is an embodied and embedded enaction. In particular, learning tasks can be appropriate for investigating the *strategies* used by the musical subjects to actively engage with given musical objects.

For example, in a study by Truitt and colleagues (1997) based on piano sight-reading, the researchers investigated the role of the musicians’ expertise measuring the eye-hand span (temporal distance between the seeing of the note on the score and the playing on the piano) and the perceptual span (the spatial region from which the agent would extract information, close to the note on the score) in expert and non-expert pianists. The results were surprisingly similar: there was no significant difference in the measures of expert and non-expert musicians. According to Jänke (2006) this can be explained by considering the sensorimotor optimization underlying the performance as belonging to the *motor* side of the process, rather than on the *visual* one, regardless of the different participants’ expertise. Musical intentionality is primarily a form of active, sensorimotor interaction, which constitutes the horizon of further motor possibilities aimed at meaningfully exploring and understanding the musical environment. In the next section I will provide additional empirical evidences to this point, once again focusing on the basic forms of musical intentionality characterising infants’ musical behaviours.

6.2 - Sensorimotor coupling and action-effect contingency in infants’ engagement with a musical instrument

Categories like “action”, “perception”, “cognition” and “affectivity” are radically intermixed. In the context of musical research, recent studies show a direct correlation between motor knowledge and auditory experience, in particular after long-term motor training, as for adult professional musicians (i.e. Koeneke et al., 2004).

Gaining adequate motor knowledge is an important prerequisite in order to develop a subject’s musical expertise, as the mastery of specific goal-directed actions is indeed beneficial for recognizing the goal of the relevant actions witnessed or heard, without any cognitive subordination. As von Hofsten argues, indeed, ‘by closing the perception–action loop the infant can begin to explore the relationship between commands and movements, between vision and proprioception, and discover the possibilities and constraints of their actions. It is important to note that the core abilities rarely appear as ready-made skills but rather as something that facilitates the development of skills’ (2007, p. 55). The enactive approach highlights the role of the *active body* as the essential tool for cognition. Without a situated body, in fact, there would be no cognition at all. As Di Paolo and colleagues put it:

the body is the ultimate source of significance; embodiment means that mind is inherent in the precarious, active, normative, and worldly process of animation, that the body is not a puppet controlled by the brain but a whole animate system with many autonomous layers of self-constitution, self-coordination, and self-organization and varying degrees of openness to the world that create its sense-making activity (Di Paolo et al., 2010, p. 42).

As previously acknowledged, however, *how* and *when* the active experience of creating music influences perception is largely unexplored in terms of ontogeny. Very briefly, it is worth reminding that “actions” are different from “movements” in light of the former intentional character. In other words, an action is *goal-directed*. Recent studies suggest that infants start at about 18 weeks of age to perform simple goal-directed actions (Daum et al., 2008). The thesis that I have offered holds that musical-directed motor behaviours can emerge during the 6-to-10-month attentive shift characterising the manual explorative behaviours in infancy, in light of the dynamic engagement between embodied agents and their *sonic environment*. In the domain of action understanding, active experience is initially more beneficial than observational experience for learning about others’ goal-directed actions. The ability to detect the goal-structure of actions plays a fundamental role in learning activities of different domains, such as word learning (Baldwin & Moses, 2001) or the understanding of cultural instruments and artifacts (Bloom & Markson, 1998). In particular, as previously remarked, infants do not solely focus on the kinematics properties of the actions or the physical structures of the objects. Rather, their attention is captured by the outcome of the actions employed to achieve a particular goal (i.e. a sound). As acknowledged by Sommerville and co-workers, ‘for instance, after watching an actor reach for and grasp a toy, infants show a stronger novelty response to a change in the actor’s goal than a change in the spatial location or trajectory of her reach (Woodward, 1998). This basic ability to construe action with respect to external goals may form the cornerstone for an understanding of goals as abstract entities that guide human action and govern event sequences’ (Sommerville et al. 2005, p.2).

In the current study, Sabine Hunnius, Harold Bekkering, Renee Timmers, Sarah Gerson and I²⁸ have connected the work on music cognition and infants' action understanding to assess the effects of active versus observational training versus no training, in a musical context. Our hypothesis was that active training would play a key role in improving infants' sensitivity to action-effect contingencies (in terms of synchronous *versus* asynchronous stimuli) in a musical context. Moreover, the population of interest (6-month-old infants) reflects the above-mentioned ideas about the first constitution of musical intentionality.

- Participants

Sixty six-month-old healthy and typically developing infants (mean age = 6.5 months) participated in one of three conditions. Infants were tested only if awake and in alert state, after their caregiver had provided informed consent. Participants were recruited through a database maintained by the Baby Lab of the Donders Institute for Brain Cognition and Behaviour, Radboud University of Nijmegen, The Netherlands. A phone call was made to the parents in the area whose names appeared in the database. Parents or caregivers who expressed interest in participating in the study were contacted to schedule an appointment. Subjects received 10 Euros or a small book for their participation. No information was asked about their ethnicity.

- Materials

We recorded and manipulated a 30 seconds video in which an experimenter plays a toy snare drum with two drumsticks (full-hand grip - alternate drumming). The video only displays the drumsticks, the toy instrument and the hands of the actor performing the music. The music played uses the rhythmical structure of the excerpt in fig. 6. The video was manipulated in order to create an asynchronous effect of 300 and 600 ms (see Lewkowicz, 1996), in which the sound was delayed or anticipated when compared to the original. Stimuli - two parallel videos per trial - were represented on a high-definition video in front of the infant, randomizing all the present material. In order to analyse the looking time preference of the participants, a camera recorded their behaviours for the entire duration of the experiment (training and videos).

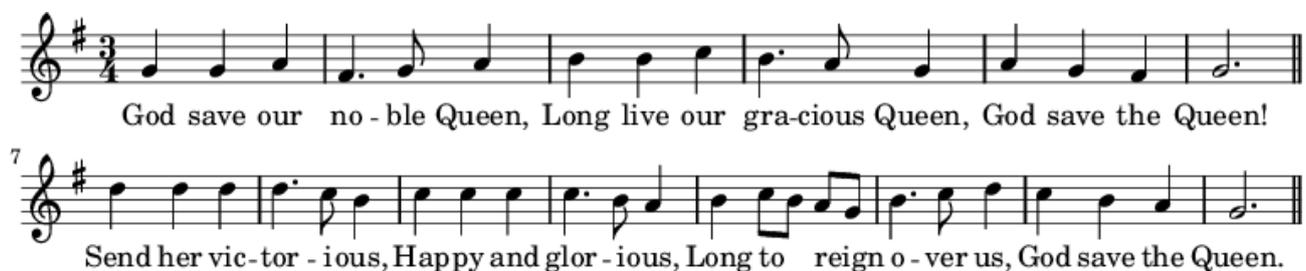


Fig. 6
Rhythmical structure of the experimental stimulus

²⁸ Schiavio*, A., Gerson*, S., Timmers, R., Bekkering, H., & Hunnius, S. (in preparation). Sensorimotor integration and action-effect contingency in infants' engagement with a musical instrument. (*both first author).

- Procedure

The infants witnessed drumming videos in which their recognition of synchronous versus asynchronous drumming actions were assessed before or after the engagement with the drumming task. In all conditions, the infant had the chance to gain live experience with drumming (*active training* [n = 20] or *observational training* [n = 20]) before the videos or active experience after the videos (*control* [n = 20]). For the entire duration of the experiment, all the participants were sitting on their caregivers' laps. In the active training condition, infants had the opportunity to play with a drum from five to ten minutes, according to their degree of awareness.

The drumstick was placed in the infants' hand and, if they did not act, the experimenter helped them by moving their hand toward the drum and demonstrating how it worked. In the observational condition, infants saw an experimenter produce drumming actions for a comparable amount of time. In the videos, infants always saw two videos presented simultaneously (see Figure 8). The drumming sound was synchronous with one of the two video displays and was off by 300 or 600 ms in the other display. The side on which the synchronous and asynchronous videos were displayed was counterbalanced across eight trials. In other words, we recorded a normal (synchronous) video, and four different versions (delayed and anticipated by 300 ms, delayed and anticipated by 600 ms). The normal video was always shown in every trial, along with one of the other versions in a randomized order. This was repeated twice, for a total of eight trials.

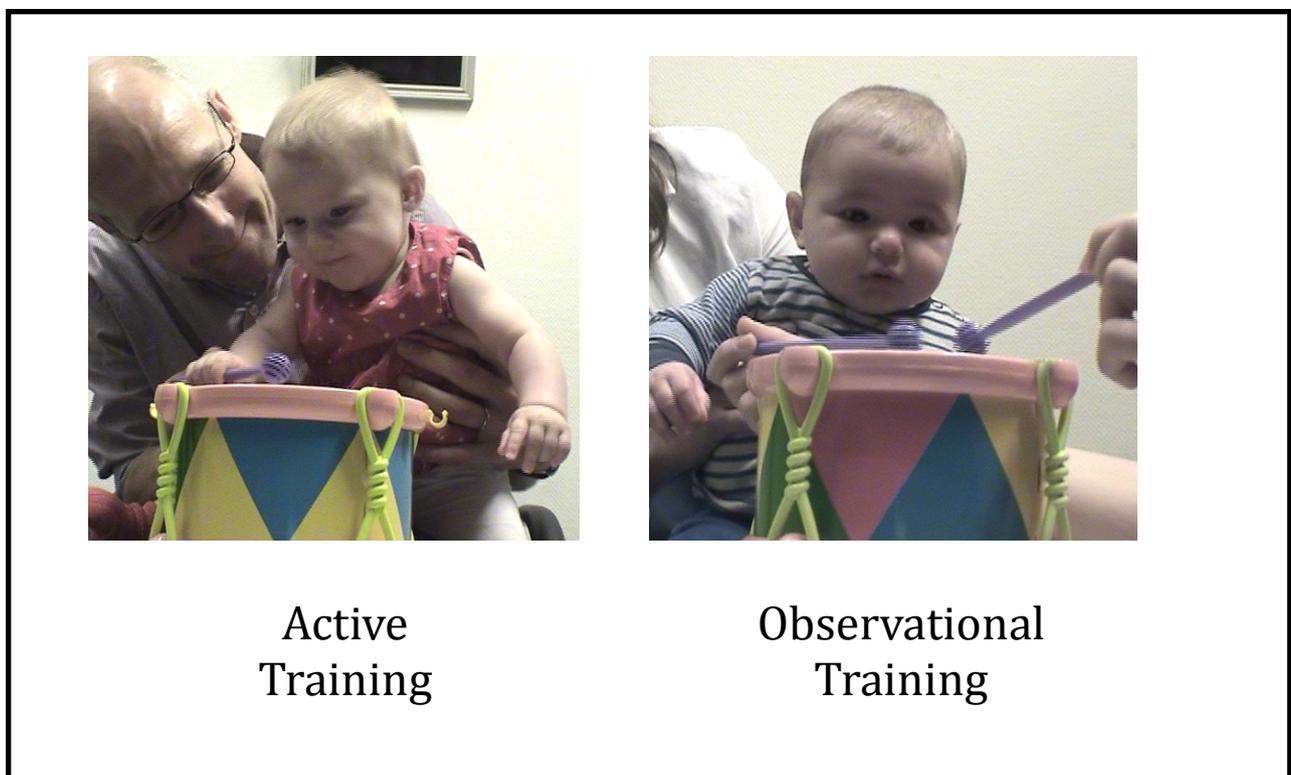


Fig. 7
Two participants during active training and observational training

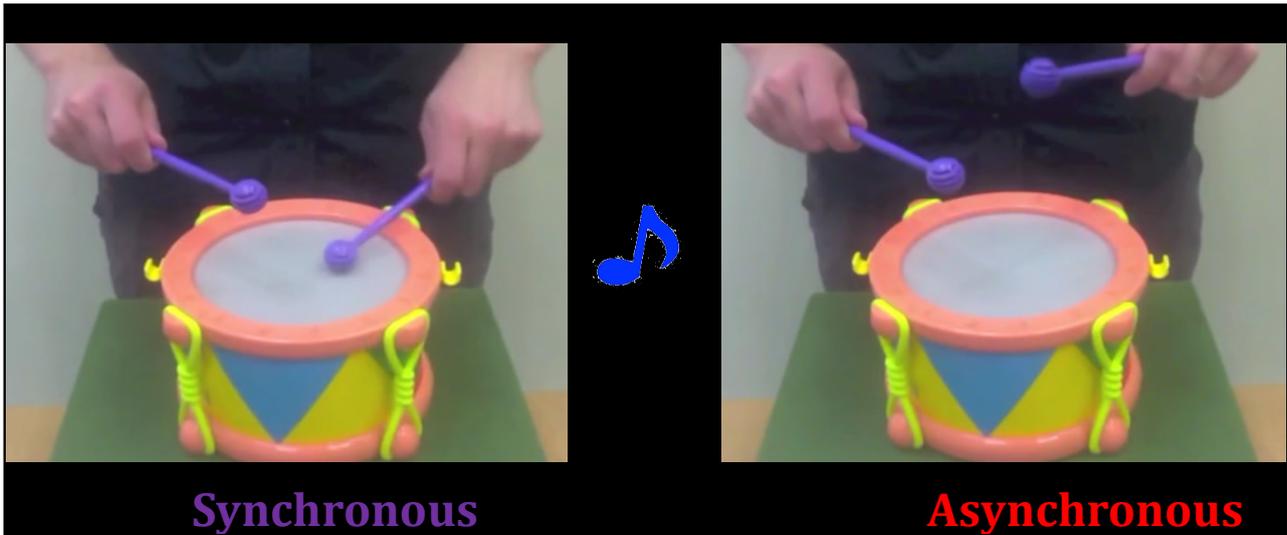


Fig. 8

Synchronous vs Asynchronous video. In every trial the infant was presented with the synchronous video on one side of the screen along with one of the asynchronous versions on the other half of it,, for a total of eight randomized trials.

6.3 - Results and discussion

We examined whether infants looked longer at the synchronous versus the asynchronous videos across trials. Infants in the active training condition looked significantly more often at the synchronous trials (see Figure 9), whereas infants in the observational and control conditions showed no preference.

In particular we found that, in the *control group*, 11 infants looked more at the asynchronous stimulus while 7 displayed more interest in the synchronous video. Two infants looked for the same amount of time at the two different conditions, so we did not use their data. Participants of the active training group, in contrast, showed a different trend: 14 of them looked more at the synchronous video while only 4 preferred the asynchronous.

Also in this case, we decided not to use the data of two infants, as they looked for the same amount of time at both the videos. As for the observational group, we found that 10 participants looked more at the asynchronous video, while 9 infants showed more interest for the synchronous condition. In this group, we found that only 1 subject looked at both the videos for the same duration. Although we are planning to examine the potential role of individual differences during training (which was between 5 and 10 mins, depending on the infant's degree of attention and participation), for now these findings provide new insight into the link between action experience and music cognition. There was a significant effect of group on the distribution of looking preference for the asynchronous or synchronous condition ($\chi^2(2) = 6.093, p < .05$).

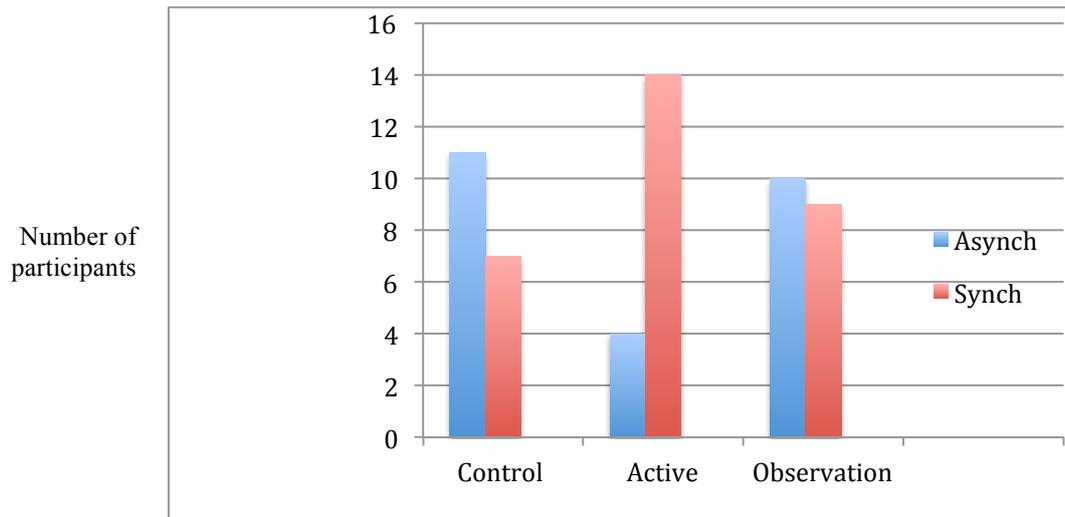


Fig. 9

The results show a clear effect of active training over observational and control condition. The infants who experienced an active engagement with the musical toy displayed a greater sensitivity to the action-effect contingency of the video, preferring to look at the synchronous video, when compared to the subjects of the other groups.

A more detailed data analysis is performed by dividing the experiment into two tests, where the first one compares control versus active condition, and the second compares observational versus active condition. This is also justified by the fact that one condition (observation) was added later to the experimental paradigm.

1) Control vs Active condition

In order to assess whether infants consistently looked at synchronous versus asynchronous videos across trials, each infant was given a binary score of 0 for each trial for which he or she looked longer at the asynchronous video and a score of 1 for each trial for which he or she looked longer to the synchronous video. The proportion of the eight trials for which each infant looked longer at synchronous videos was then used as a dependent variable. An independent samples t-test indicated that infants in the two conditions differed in the proportion of trials for which they looked longer at synchronous videos, $t(38) = 2.27, p = .029$ (cohen's $d = .72$). In order to follow up on this effect of condition, one-sample t-tests were conducted for each condition to determine whether the preferred proportion of trials using synchronous videos differed from chance (50%). In the pre-training condition, infants looked longer at synchronous videos on approximately 60% of the trials; this differed significantly from chance level: $t(19) = 2.73, p = .013$ (cohen's $d = 1.25$ —*bonferroni corrected: $p < .025$*). In the post-training condition, infants looked toward synchronous videos on approximately 49% of trials and did not differ from chance, $t(19) = -.29, p = .78$ (cohen's $d = .13$). Non-parametric, binomial tests confirmed these findings. More infants in the pre-training condition looked longer at the synchronous trials than asynchronous trials ($p = .022$), but the number of infants who looked longer at these different types of trials did not differ for infants in the post-training condition ($p = .79$). We also examined infants' differentiation between synchronous and asynchronous videos separately for 300ms and 600ms offsets. Independent samples t-tests indicated that infants in the two conditions differed in the proportion of trials for which they looked longer at

synchronous videos for the 600ms offset trials, $t(38) = 2.53, p = .016$ (cohen's $d = .80$), but not for the 300ms offset trials, $t(38) = .99, p = .33$. One sample t-tests for the 600ms offset trials confirmed that infants in the pre-training condition differed significantly from chance, $t(19) = 3.68, p = .002$ (cohen's $d = 1.69$), whereas infants in the post-training condition did not, $t(19) = -.30, p = .77$ (cohen's $d = -.14$). Non-parametric, binomial tests confirmed these findings. More infants in the pre-training condition looked longer at 600ms offset synchronous trials than asynchronous trials ($p = .006$), but the number of infants who looked longer at these different types of trials did not differ for infants in the post-training condition ($p = 1.00$).

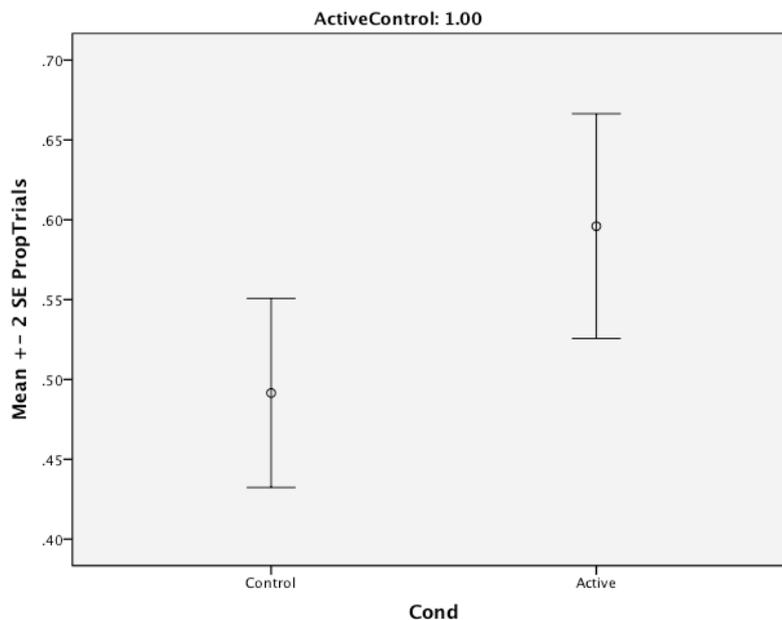


Fig. 10
 T-test comparing the means of control and active condition:
 $t(38) = 2.27, p = .029$.
 One samples T-test comparing the mean with chance level (.5):
 Control group: $t(19) = -.29, p = .78$
 Active group: $t(19) = 2.73, p = .013$

2) Observational vs Active Condition

As in the previous analysis, we first examined whether infants differed from chance in the number of trials in which they looked longer at synchronous versus asynchronous trials. Infants in Experiment Two looked longer at synchronous trials in approximately 51% of trials. A one-sample t-test confirmed that this did not differ from chance, $t(19) = .11, p = .92$ (cohen's $d = .05$). A binomial test further indicated that the number of infants who looked longer to more synchronous than asynchronous trials did not differ ($p = 1.0$). An independent-samples t-test was used to directly compare infants in the post-training condition from Experiment One with infants in Experiment Two. Across all trials, infants in these two conditions did not differ in the number of trials in which they looked longer at synchronous or asynchronous trials, $t(38) = 1.35, p = .19$ (cohen's $d = .43$). Following these analyses, we once again examined the trials with 600ms versus 300ms offsets

separately. One-samples t-tests indicated that infants did not differentiate between synchronous and asynchronous trials in either 600ms offset, $t(19) = -.69, p = .50$ (cohen's $d = -.32$) or 300ms offset trials, $t(19) = .78, p = .45$. A binomial test indicated that the number of infants who looked longer at more synchronous versus asynchronous trials did not differ in the 600ms offset trials ($p = .79$). We then compared infants' attention to synchronous versus asynchronous 600ms offset trials across post-training infants from Experiment One and observational infants from Experiment Two. An independent-samples t-test indicated that infants in these two conditions differed from each other in the 600ms offset trials, $t(38) = 2.42, p = .02$ (cohen's $d = .76$).

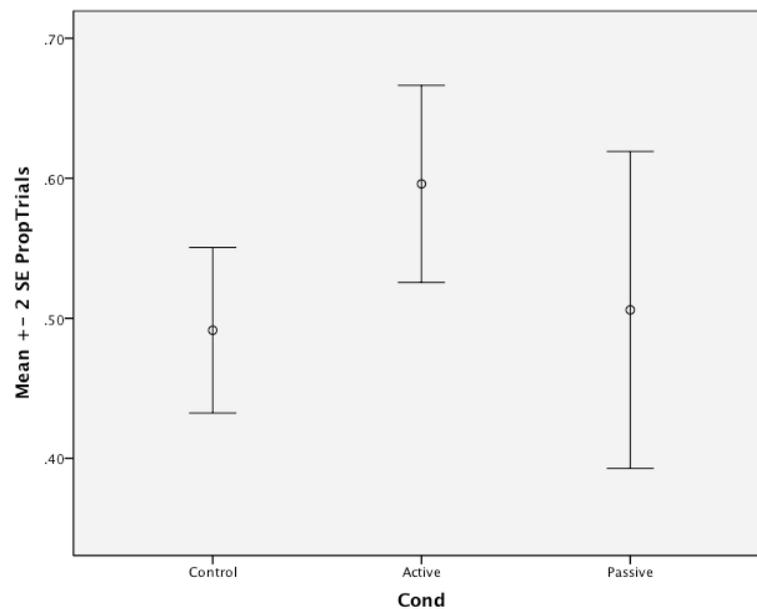


Fig. 11

T-test comparing the means of passive and active condition:
 $t(38) = 1.35, p = .19$ (not different)
 For 600 ms: $t(38) = 2.42, p = .02$ (different).
 One samples T-test comparing the mean with chance level (.5):
 Passive group: $t(19) = .11, p = .92$

These data, therefore, confirm our hypothesis. Infants who participated in the active training displayed higher sensitivity to the action-effect contingency of the videos, highlighting the key role of motor knowledge for musical perception. In other words, the acquired teleomusical acts represent the necessary condition to understand the musical object in a meaningful way, given the infants' ability to discern between the synchronous and the asynchronous condition. With regard to this point, it is worth noticing that the new motor knowledge acquired (in terms of playing with drumsticks on a drum) represents a good example of how CTAs could emerge from mutual interaction, active engagement and motivation. The experimenter helped the participant to employ strategies in order to meaningfully interact with the explored object. The sensorimotor integration at the basis of this process shows that 'the infant's nervous system has an endogenous, intrinsic dynamic, which generates transient patterns of activity' (Chemero, 2009, p. 202). These patterns of motor behaviour would alter something in the environment; in this case they generate sounds. And these sounds 'impact the infant's sensorimotor coupling to the niche [...], which in turn perturbs the

endogenous dynamics of the infant's nervous system [...] and again and again through the loop' (ibid.). De Jaegher and Di Paolo (2007) assume that sensorimotor coupling is determined by the self-organized structure of a living system. As they put it, in fact:

Such a view of cognitive systems as autonomous rejects the traditional poles of seeing cognizers as responding to environmental stimuli on the one hand, and as satisfying internal demands on the other - both of which fail to give the autonomous agent its proper ontological status and subordinate it to a passive role of obedience. A key principle of the enactive approach is that the organism is a centre of activity in the world. The relation of *emergence* between novel forms of identity (e.g., integrated sensorimotor engagements as emerging from neural, bodily and environmental dynamics) is one whereby the coupling between the emergent process and its context leads to constraints and modulation of the operation of the underlying levels (ibid., p. 488).

In other terms, to explain infants' musicality, we would need to integrate the entire system of inter(en)active sense-making, without only focusing on brain correlates, sonic environment, muscles and so on. Infants provided with adequate training display significant facilitation in detecting a synchronous stimulus because perception is a skilful activity involving the whole living creature. This perspective on learning and development might remind us of the influential work by Esther Thelen (i.e. 1989; 1994). Consider, for example, the following passage by Linda Smith, which perfectly describes Thelen's contribution:

According to Thelen, the processes that give rise to motor behaviour are [...] the repository of knowledge and the driver of developmental change. As phenomenon, they also provide the key to the nested dynamics of human development [...]. Traditionally, psychologists have considered action, learning, and development as distinct processes. Thelen argued and showed us in her work how this conceptualization is wrong (Smith, 2006, p. 89).

Moreover, the fact that - already at 6 months of age - infants are able to distinguish between synchronous and asynchronous stimulus because of their newly acquired repertoire, reflects the flexibility of the model of musical ontogeny provided in Chapter 5. It would be impossible to determine *exactly* how and when patterns of sensorimotor agency would constitute a truly enactive kind of intentionality. My model aimed at emphasising the active role of explorative behaviours in light of the 6-to-10-month attentive shifts, rather than providing a definitive structure of the nature of musicality. My aim was to provide a biological plausible account of early musical behaviours in light of phenomenological and empirical results without postulating any sort of "unified timetable" of developmental changes. As Thelen argues

at developmental transitions, one or several components of the complex system may act as control parameters, including variables in the context or in the environment [...]. Although all of the elements or subsystems are essential for the systems output, only one or a few of the subsystems will trigger transitions, which, in turn, will lead to system-wide reorganization. This principle helps explain the heterochronic, asynchronous, and often nonlinear character of behavioural ontogeny. We commonly observe "pieces" of a functional behaviour long before the performance of the mature behaviour. These pieces seem to be used out of sequence, in inappropriate or different functional contexts, only under certain experimental conditions, or otherwise not properly "connected" with the other elements needed for goal-directed activity. Theories that assume that developmental change is driven by a unified timetable in the form of maturational plans, neurological reorganizations, or cognitive structures have had difficulty accounting for both the anticipations of function and regressions. In this systems approach, we strongly emphasize that contributing components may mature at different rates. The component processes are thus developing in parallel, but not synchronously or symmetrically (Thelen, 1989, p. 88).

It is quite clear, therefore, that a new approach to music cognition should focus on the *processes* underlying musical behaviour rather than their mere *outcome*. Otherwise we will lose any chance to gain an insight into the phenomenological structures of human musicality. In particular, given that no ontogenetic change is pre-wired in the brain of the animal, it seems that what matters for musical behaviour is the intentional characterisation of the dynamic processes of sense-making, where goal directed patterns of action continuously implement and integrate self-organized properties of the system in its openness to the world. As Thelen puts it, in fact, 'developing systems are stable and predictable where their adaptive demands have constrained, through phylogenetic mechanisms, their range of solutions [...]. The physical and social context of the developing animal is more than just a supportive frame; it is an essential component of the assembled system. In such systems, new forms arise when the stability of the system is disrupted when random fluctuations are amplified by the scaling of a critical component' (ibid., p. 114).

In light of these considerations, it seems that the boundaries between brain, body and environment should be reconsidered, hence admitting a paradigm where the body-brain-environment nexus would orchestrate the inter-relational, active, and sensorimotor dynamics of human musicality. Music cognition, in these terms, is a distributed category that reaches out beyond the boundaries of skull and skin. After having established the core of my proposal with regards to infants' musicality and the basic forms of musical sense-making, it would be interesting to see how this model would work in adults. The next section, therefore, describes another experiment with a population of adult musicians and non-musicians, engaging in learning atonal melodies under different conditions.

6.4 - Cross-Modal facilitation in learning atonal melodies

In this study, Renee Timmers and I²⁹ investigated the role of motor learning in the memorization of four different musical excerpts for piano. Pianists, other musicians (non-pianists) and non-musicians have been asked to familiarize with four different short piano melodies under different conditions: “playing condition” (performing the melodies on a keyboard), “silent tapping condition” (performing the melodies on a piano without any auditory feedback) and “seeing condition” (watching a video with a performer playing the melodies). Our idea was to examine the role of motor learning in basic musical tasks.

The audio-motor integration underlying musical experience has increasingly received attention over the last few decades. As shown by Stefan and colleague (2005), the ‘mere observation of movements leads to the formation of a lasting specific memory trace in movement representations that resembled that elicited by physical training’ (ibid., p. 9344). The relationship between perception and action can be intended more as *perception-for-action*, considering, in the realm of music, the heard, read or played melody ‘in terms of the intentional, hierarchically organized sequences of expressive motor acts behind the signal’ (Overy & Molnar-Szakacs, 2009, p. 492). The implications for musical learning, however, remain still not fully clear. How this paradigm can improve our understanding on the intentional relationship between a musical subject and a musical object? Is the musicians’ memory *instrument-specific*? And how is non-musicians’ memory affected by music-directed actions’ observation?

It is already well known that auditory–motor associations acquired while learning to play a piece can be determinant for later auditory recognition (Brown & Palmer 2012), but most of the relevant literature focuses on a population of well trained musicians engaged in expertise-related tasks with their own instrument. In their study on musical learning, indeed, Brown and Palmer (2012) argued that ‘motor learning can aid performers’ auditory recognition of music beyond auditory learning alone, and that motor learning is influenced by individual abilities in mental imagery and by variation in acoustic features’ (ibid., p. 567). However, the authors considered only a population of expert pianists, thus referring to subjects whose musical-directed motor vocabulary is already partially constituted.

The present study, therefore, aimed to investigate the constitution of a musical-motor vocabulary of acts, comparing in the same piano-based task pianists, musicians (non-pianists) and non-musicians to see not only whether the motor system engagement can be similar in musicians and non-musicians when the task requirements are musically relevant (Zatorre et al., 2007), but also to understand in which terms the musicians’ intentionality can be intended as *cross-modal*. Considering that many empirical evidences (e.g. Lotze et al., 2003; Kristeva et al., 2003; Zatorre & Halpern, 2005) show that musical imaginary and perception are strictly linked, being underlied by the same phenomenological structure and neural correlates, a crucial issue concerns the *topography* of this cross-modal feature of musical intentionality. Namely, we wanted to investigate the concrete *modalities* in which musical subjects with different musical expertise engage with musical objects. We therefore asked the participants to familiarize with four atonal³⁰ melodies in three different

²⁹ Schiavio, A., & Timmers, R. (in preparation). Cross-modal facilitation in learning atonal melodies.

³⁰ We used atonal music in order to avoid any facilitation in remembering the melodic fragments with the help of the common tonal rules.

conditions with diverse degrees of motor engagement (playing, playing without any auditory feedback and only seeing the hand of a performer playing), testing their ability to recognize these melodies among different ones. We expected the subjects' motor expertise to show its advantages beyond other learning modalities. In particular, we hypothesised that the strong coupling between auditory and motor information would allow the participants who performed the melodies on the piano (with *both* auditory and motor feedback) to remember them with more accuracy than subjects involved in other learning modalities. For this reason, we also expected an advantage for pianists in familiarizing with the stimuli.

- Participants:

One hundred and twenty healthy adult volunteers (mean age = 26.725) were asked to take part in this study. In order to recruit participants, an announcement was placed in the Music Department of the University of Sheffield and on the Internet. The population examined was divided into three groups, according to their musical expertise. *Musicians who did not play the piano* (A, n=40, mean age: 27.75), and *pianists* (B, n=40, mean age: 27.75) were required to have at least 5 years of advanced musical training. *Non-musicians* (C, n=40, mean age: 24.675) were required to have no previous experience in musical training or performance. Musicians who took part into the experiment had a different background, having studied classic music, rock, metal, folk, or jazz. In particular, besides 40 pianists (group B), the population of musicians (group A) includes 40 subjects who studied the following instruments: guitar (n=16), voice (n=9), violin (n=3), clarinet (n=3), drums (n=2), sax (n=2), trumpet (n=1), cello (n=1), oboe (n=1), accordion (n=1), French horn (n=1). The mean of the years of training for all the musicians was 12.83.(pianists = 16.175; other musicians = 9.5).

- Materials

Subjects listened to the stimuli through headphones at a comfortable volume. Participants involved in playing and silent tapping conditions performed on an Evolution (model MK 44-9C) keyboard without weighted keys. Several videos have been recorded in order to assist the participants in the tasks: for the silent tapping and playing conditions two videos have been recorded (1 for right-handed participants and 1 for left-handed). The videos provided the volunteers with all the relevant instructions for placing their hand on the keyboard and perform the melodies following a representation of the keyboard with highlighted keys on the screen (see figure 12). For the seeing condition, volunteers have been only asked to watch a video with the performer's hand playing the melodies. The videos (one for right-handed subjects and one meant for left-handed participants) have been recorded with a video camera placed above the piano, giving an aerial view of the performer's hand (see figure 13).



Fig. 12
Playing and silent tapping condition; highlighted keys to assist the playing task



Fig. 13
Seeing condition: aerial view of the performer's hand

The stimuli of the experiment include (i) 4 short atonal piano melodies and (ii) their variations - for a total of 16 melodies (4 correct melodies / 12 variations). The stimuli have been performed by a professional pianist and recorded at the Department of Music of the University of Sheffield using Cubase[®]. The 4 original piano melodies have been taken (entirely or with slight modifications) from C. Ives compositions for voice and piano (1922) and last less than 15 seconds each. The melodies are made of 5 notes, allowing the participants to use a single hand position on the keyboard, pressing one key with each finger. The four main stimuli are the following:

♩ = 66

♩ = 66

♩ = 66



Fig. 14
The four stimuli used for the experiment

The other stimuli are melodic or rhythmic variations of the original melodies, providing a total of 16 stimuli (4 corrects and 12 incorrect). The next figure shows an example of how a melodic fragment has been varied.

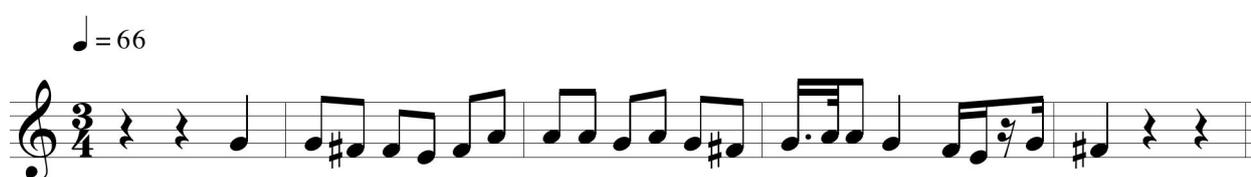


Fig. 15
first variation of Stimulus 4

- Procedure:

A specific learning trial was set up for all the conditions: firstly, all the participants listened once to all the stimuli at a comfortable volume. Afterwards they went through their specific learning condition. According to the relevant task, each group was divided into four subgroups, each of them made by 10 volunteers. *Subgroup 1* was involved in a playing task, *subgroup 2* in a silent tapping condition, participants of *subgroup 3* had to watch a video and *subgroup 4* was the control group. Participants were comfortably seated in front of a keyboard, which was placed by the screen of a computer. For conditions 1 and 2 (playing and silent tapping) participants listened to one melody before performing it on the keyboard. This process is repeated two times for each melody. For condition 3 (seeing), subjects listened to one melody before watching the performer playing (without any auditory feedback) it. Again, as for the other conditions, after two repetitions of the same stimulus the video presents another melody to familiarize with. For the control condition (4) participants are only asked to listen to the original melodies and to go straight to the recognition phase. The melodies were always presented randomly for a total of eight trials (two for each melody). In other words, the experiment is divided into two main parts. In the first one subjects are asked to listen to the four atonal piano melodies. After listening to each melody the volunteers will perform a different task, according to his/her subgroup:

- Sub-group 1 → every participant *plays* on the piano the melodies
- Sub-group 2 → every participant *imagines* to play the melodies, through a silent tapping performance on a mute keyboard.

- Sub-group 3 → every participant *watches* a video with a pianist's hand playing the same melodies.
- Sub-group 4 → participants are asked only to *listen to* the melodies.

After 10 minutes from the first part, participants were asked to recognize the same melodies from a set of different stimuli. A computer-based questionnaire was given for the volunteer to respond Y (yes) or N (no) after each stimulus.

6.5 - Results and discussion

A univariate analysis was conducted with learning condition and musical training as between-subjects variables. The dependent variable was the mean correct response. This analysis showed a significant main effect of condition ($F(3,108)=5.466, p=.002, r=.363$). There was no significant effect of musicianship ($F(2,108)=0.235, p=.791$), nor a significant interaction between musicianship and condition ($F(6,108)=0.091, p=.997$). A pairwise posthoc test was conducted using Bonferroni corrections for multiple comparisons to investigate the differences between the conditions. This analysis showed that the mean correct scores were lower for the control condition than the playing ($p = .05$) and silent tapping ($p = .002$) conditions. It also revealed that the scores in the silent tapping condition were marginally better than in the seeing condition ($p = .053$). The means per condition are given in Figure 16.

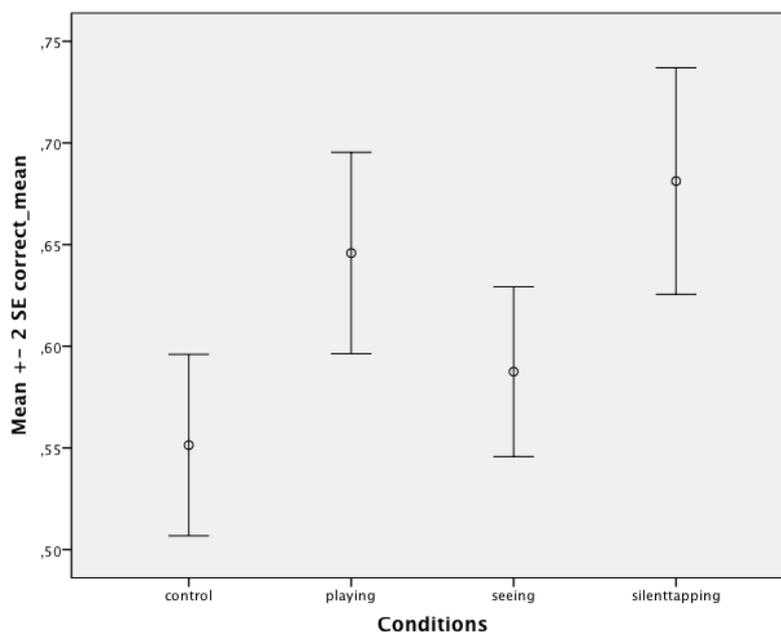


Fig. 16
Mean of the correct results per condition.
Note that no significance difference was recorder between groups.

The results show a difference between the subjects of subgroups 3 and 4 (seeing and control) and subgroups 1 and 2 (playing and silent tapping). It is interesting to note that the discrepancy between musicians and non-musicians in providing correct answers (recognizing the correct melodies and naming the wrong ones as incorrect) is not significant. These data suggest that playing music or silent tapping gives the participants a greater chance of learning the melodies, regardless of their musical expertise. A standard interpretation of these results would be in terms of memory's facilitation elicited by multi-modal engagement with the musical stimuli. Basically, the more modalities you are engaged with, the better your memory will perform when trying to recognize musical fragments previously learned. This is certainly true. However, a more speculative interpretation is possible, without any contradiction with this more traditional observation. What changes from the classical interpretation is the nature of the modalities involved. Instead of perceiving, coding, storing and representing, we could think of these results in more dynamical terms, where what truly matters for sense-making is the active sensorimotor engagement with the melodic fragments proposed. Indeed, according to the *intentional* account of musical experience previous provided, these results are not surprising. In order to make sense of the melodies, and then to recognise them among others, a subject needs to be *motorically linked* to the musical stimuli. The newly acquired motor knowledge, gained through playing and silent tapping conditions, provides the participants with a stronger modality to engage with the musical environment, above conceptual conscious reflection.

With this in mind, note that these data cannot fit into the traditional paradigms of musical cognition, where musicians, thanks to their knowledge of musical structures, will be significantly better than non-musicians - using strategies such as anticipation and grouping to facilitate their understanding of the excerpt. With regards to this point, someone would argue that musicians would not actually have any advantage, given the atonal nature of the stimuli. For example, it would be unclear how they would use the theoretical knowledge that comes from traditional Western music. However, the musicians' familiarity with musical structures in general, and with learning or listening strategies in particular, would have certainly played a key role in the task. On the contrary, despite the basic nature of the stimuli, what seems to matter for musical sense-making is the *motor intentionality* at the basis of the mutual co-determination between musical subjects and objects, and not abstract theoretical abilities.

At this point, however, a clarification seems necessary. Mark Rowlands (2006) assumes that the movement of the pianist's fingers, while playing, is to be considered as pre-intentional, as the performer is *not aware* of the actual position of each finger on the keyboard. According to Gallagher and Zahavi (2008), however, although an expert pianist will not display a detailed awareness of everything she is doing while playing, if we would stop her in the middle of the performance and we would ask her whether her middle finger was playing a C-Sharp, she would certainly reply correctly. So, as their argument goes, the pianist's movements are intentional acts. Although I do consider these movements as intentional, I disagree with both Rowlands and Gallagher and Zahavi's interpretations. There is no need to postulate any kind of intermediate level between non-intentional and intentional acts, given that not-being-aware cannot be a coherent category to distinguish among them. As previously shown, in fact, the Mirror Mechanism's literature clearly highlights pre-conceptual and unmediated resonances as the prerogative of motor intentionality. Therefore, the movements of pianist's hand on the keyboard are intentional not because she is aware of each finger's position (I highly doubt that a performer would be actually

aware of each finger's position) but because her actions are goal-directed, regardless of her awareness.

The goal of a musician's set of actions is not to put each finger in the right position. Rather it is playing the chord, using the right dynamic, the phrasing the motif, and so forth. Otherwise, the segmentation of actions in mere kinematic movements would substantially compromise the fluidity of executions that characterise most of the musicians. To my mind, this experiment shows that musicians and non-musicians are engaged with the musical object without recruiting high-level, theoretical knowledge. Otherwise musicians (pianists and non-pianists) would have performed much better, when compared to non-musicians. In other words, the layers of the sense's stratifications of the stimuli are best understood in motor rather than conceptual terms. Many non-musicians, while engaging in the playing or silent tapping condition, were completely unaware of the position of their hand on the keyboard (because of their lack of knowledge of the keyboard itself, and because their attention was more focused on performing correctly non-familiar movements). However, since their results were not significantly different from musicians, we could conclude stating that their intentional relationship with the musical object displayed the same structure and phenomenological character of the group of pianists and other performers.

Chapter 7

Conclusion. Towards a post-Cartesian cognitive science of music

Because we are in the world, we are condemned to meaning, and we cannot do or say anything without its acquiring a name in history.

- Maurice Merleau-Ponty, *Phenomenology of Perception*

The feeling of an unbridgeable gulf between consciousness and brain-process: how does it come about that this does not come into the considerations of our ordinary life?

- Ludwig Wittgenstein, *Philosophical Investigations*

Traditional music cognition, in its commitment to various forms of Cartesianism, has presupposed and defended the integrity of boundaries that may not hold up. This work, therefore, aimed to reconsider dichotomies and assumptions derived from the Cartesian metaphysics, proposing a different interpretation of human musicality, in light of the enactive approach to cognition. The openness of an embodied agent towards her social environment, the active, affective engagement underlining mutual forms of communications are seen in my dissertation as crucial points for criticizing these long term dichotomies (between *internal* and *external*, *input* and *output*, *brain* and *body*, *subjects* and *objects*). The notion of musical intentionality here offered intended to overcome such a dualistic stance by emphasising the direct, motor-based, coupling that links musical subjects and musical objects.

I began this dissertation in Chapter 1 highlighting the continuity between Cartesian tradition and disembodied cognitive science. This commitment gives rise to problematic perspectives about the nature of mentality and fails to integrate the role of the environment and of the body for cognition. Shortly, I have argued that cognitivism makes a double mistake. First, it does not provide any kind of accurate phenomenological explanation of musicality. Second, its presuppositions are flawed, being immersed in an isolationist analysis that cannot coherently reflect the active and articulated dynamics underlining music cognition.

Similarly, as I have pointed out in Chapter 2, current trends in music cognition suffer from the same “Cartesian anxiety”, postulating dichotomies between inner and outer, input and output that would eventually portray music and musical experience as something substantially avulsed from the concrete practices in which they emerge. I gave arguments against the common perspective in

musical research about the modularity of the musical mind, defending the claim the music, rather than being biologically pre-wired in brain modules, is ontogenetically and intentionally constituted.

The development of this alternative view, based on phenomenological and embodied traditions was provided in Chapter 3, where I also introduced the fundamental notion of musical intentionality in light of the discovery of a Mirror Mechanism. The same section also explored the conceptual geography of the enactive and embodied approaches to human cognition, dealing extensively with empirical perspectives (i.e. the application of dynamic system theory).

In Chapter 4 I provided a detailed analysis of one of the most famous books on embodied music cognition, showing that some leftovers from Cartesianism are still present in the current debate on embodiment. In this section I have argued that only a full integration of mirroring and co-constitutive processes could provide a biologically plausible interpretation of the basic forms of musical intentionality, providing a reconceptualization of musical affordances. In particular, considering the activation of mirror neurons as the enabling condition of enactive perception and sense-making allows us to avoid the problematic notion of *simulation*.

Chapter 5 faced the challenge to see in detail how this paradigm would work, considering infants' musical behaviours as the starting point for discussion. In this section I have emphasized the role of embodied and enactive dynamics at the basis of human musicality, showing that the active exploratory behaviours play a cardinal role in the dynamic interplay between cognizers and musical environment.

Finally, in Chapter 6, I presented two empirical studies, in order to corroborate the previous assumptions with regard to infants and adult musical learning. *The enactive approach to music cognition* I have presented, to put it shortly, holds that human musicality is the active engagement that links embodied agents and the world of sound. It emerges in circular, interactive, sensorimotor and goal-directed dynamics.

From these chapters it emerges that *musical sense-making* is not a passive representation of elements of the musical environment. Rather, it is a process of bringing forth, or *enacting*, a subject's own domain of meaning through sensorimotor knowledge. This view is best understood by considering the musical niche as a place of affordances, defined not as fully given categories, but as properties of the intentional relationship itself - hence avoiding any risk of objectifying the role of the environment. *Musical affordances*, being constituted by the circular interplay between musical subjects and musical objects, show the direct and intrinsically motor fashion of the mutuality of world and embodied subjectivity. Another very important issue is the analysis of the actual motor behaviours defining the sensorimotor patterns of interactions. From the detailed focus on infants' musicality, I have thus developed the notion of *teleomusical acts*, musical-directed sets of actions at the basis of musical expertise. These acts do play a fundamental role not only in music performance, but also during perception. In fact, given that mirroring processes represent the enabling condition of sense-making, *teleomusical acts* would constitute the subject's motor vocabulary, giving rise to the processes of mutual interplay at the basis of perceptual activity, intended as *exploration*.

In light of these considerations, my work has implications not only for music cognition but also for the broader field of cognitive science. While investigating human musicality in terms of embodied action and mutual inter(en)action is a significant step forward from more traditional methodological paradigms in musical research, my proposal to integrate the enactive standpoint with the study of goal-directed (chains of) actions, represents a new contribution for the studies of mind and subjectivity in general. Although many scholars have already proposed a paradigm shift

in the philosophy of mind and the cognitive science, focusing on the concrete activity of the agents in the world, a precise characterization of the *type* of action relevant for sense-making has rarely been addressed. This approach, broadly speaking, aims to mediate between the radical view of embodiment posited by *enaction* - by questioning the plausibility of neural correlates of musical consciousness - and the neuroscience of Mirror Neurons. Taking into account the direct, interactive, reciprocity of embodied agents, this standpoint advocates a reconsideration of musical learning and music cognition in general.

The *positive* part of the dissertation, can be summarized through the following points:

(i) Music is an *intentional object* of experience

The core of my argument is based on the phenomenological notion of “intentionality”, a term which refers to the property of consciousness to be always directed towards an object. Music, therefore, is an object of intentionality because there is always a subjectivity directed towards it. We can listen to music, imagining music, playing music, analysing music, and so forth, and we cannot have music, or a musical experience, without being in a relation with it. Considering music as an intentional object provides a significant step forward from traditional research, mainly because under this characterization, music becomes inseparable from the subjects who are experiencing music. In this sense, music is not an abstract, isolated, category but rather a complex phenomenon, which manifests itself only in its relation with human subjectivity. This relation, called “musical intentionality”, represents the main premise of my argument.

(ii) The main feature of musical intentionality is, for agents, to *give sense* to a (musical) object.

Musical intentionality defines the relation between music and experiencing subjects. But what is the nature of this relation? After analysing phenomenological, neuroscientific, and musical literature, my argument proposes that, as an intentional object of experience, music is an object of significance, value, meaning and emotion. But because we cannot consider music as an independent category from the subjects of experience (first premise) it seems that this imposition of meaning is part of the relation itself, and not a separate process. In other words, to experience something already means to give sense to that something. A subject is always attributing a meaning to the object of experience, as there is no experience without significance. Experience is sense-making.

(iii) This sense-giving ability is based on the *power of action of the body*.

Meaning attribution does not require any high-level, cognitive, subordination. As Merleau-Ponty (1945) famously asserted, in fact, the body is the ultimate source of significance. We can make sense of the world in which we are embedded and its objects thanks to our embodied ability to generate patterns of meaningful actions. We are always immediately open to the world without

postulating representative mechanisms at the basis of our meaningful engagement with it. Experiencing music, before and below any conceptualization, is an embodied experience, where bodies and modalities of action provide a horizon of possible ways of understanding and making sense of the intentional object, given the motor expertise of the subjects. This final point is corroborated by the neuroscientific evidence of a mirror mechanism for action understanding in humans, as it is involved in music experience (Overy & Molnar-Szacaks 2000). Although it still remains controversial how exactly, and to what extent, neuronal processes would participate in meaning attribution, it seems clear that the motor repertoire of a subject, in terms of her ability to play an instrument, underlies a basic imposition of meaning, prior to any cognitive or intellectual processing. Experiencing music, at a basic and automatic level, is a sense-giving process (second premise) immersed in the dynamics of action.

(iv) A *music-directed repertoire* of motor acts plays the key role for musical experience.

The enabling condition of musical sense-making is represented by the mirroring activity, which allows an agent to attribute a meaning to the musical object, thanks to the subject's degree of motor knowledge (third premise). But if brains play only a part of musical sense-making, then it seems we should investigate the nature of musical experience from a perspective that takes into account the embodied dynamism of the experiencing subjects. The decentralization of mental processes I proposed leads the study of human musicality towards the active modalities of embodied inter(en)action with the musical environment. I thus individuated two sets of musical-directed actions at the basis of our ability to make sense of music: original teleomusical acts (OTA) and constituted teleomusical acts (CTA). While OTAs can be easily witnessed in infants' exploratory behaviours, CTAs reflect the mastery of the specific goal-directed chains of action involved in the production of musical contexts, being constituent of the development of musical expertise. When an infant autonomously explores the environment and her attention is captured by the sound properties of an object, she might employ a specific goal-directed motor behaviour to generate (and to modify) sounds. She can, for example, continuously hit an object in order to play with the sounds and even generate basic rhythmic, dynamic, or melodic variations. When this kind of activity exhibits clear goal-directedness (that is: the infant is employing specific actions in order to generate and play with sounds), then the actions used by the infant can be named original teleomusical acts. As the infant grows older and the ability to master these actions improves, she could perform more sophisticated patterns of goal-directed sensorimotor activity, perhaps associating different OTAs without compromising the fluidity of her movements. This is for evident for adult musicians, who can easily display non-associative motor behaviours while playing. When for example a professional guitarist explores the dynamic possibilities of the instrument to increase the tension of a given theme, she will not focus on the position of each finger independently. Her motor expertise, rather, allows her to intervene and modify the piece's dynamics through a series of coordinated movements, without unnecessary high-level, intermediate, speculations. Similarly, an infant exploring an object, as she acquires enough motor expertise, develops greater ability to reach a particular musical goal without focusing on single movements. For example she could use a stick, or both hands, in order to create sounds experiencing *grasping the tool* or *coordinating the hands* as contributing components of the goal-directed musical actions and not mere kinematics per se.

Although at the beginning she might encounter difficulties and she might focus on single actions (holding the stick, individuating the surface to generate sounds, drumming) after some time she will be able to construe and perform basic sets of actions with greater accuracy and fluidity, individuating a musical goal (the goal of the set of action is “playing with sounds” rather than “holding the stick”). Both OTAs and CTAs, therefore, are constituent of human musicality by shaping the degree of meaning attribution, being part of the circular interplay between perception and action that is central for the enactive approach to cognition.

(v) Music cognition is *not in the head*.

Musical intentionality is dynamically modulated by the musical object (which provides the agent with a horizon of motor possibilities to interact with - defined as affordances) and the musical subject (which makes sense of the object through her sensorimotor knowledge). It follows that experiencing music is a process of codetermination between subjects and objects, where brains, bodies and environmental features dynamically interact through circular sensorimotor loops. In light of these considerations it seems that musical sense-making is a process that extends beyond the boundaries of skull and skin, being constituted by embodied and embedded forms of mutual determination.

Implications

- Musical education - from early infancy - should focus more on the *constitution* of a vocabulary of motor acts, by investigating motivations, participatory musical practices and exploratory behaviours. The development of musical expertise is an emergent phenomenon constituted by embedded and embodied forms of interactions between a self-organized living system and its (musical) environment and not a mere solipsistic event.
- Further research in music cognition should assume that the very bases of human musicality are *motor* in nature and the taken-for-granted division between internal mental representations and pre-given external world should be therefore ruled out, as it cannot coherently integrate the feedback of the organism's action to cognition.

Ending a thesis without the classic inflammatory claim, however, might be not entirely satisfactory. Hence, consider the second implication from the summary of my proposal. If further research in music cognition will seriously consider the very bases of human musicality as *motor* in nature, then music therapy (e.g. in Parkinson's disease) might exploit the role of the *chains of teleomusical acts* to design rehabilitation paradigms for improving patients' motility. In this sense, the processes of sensorimotor integration described will improve not only our understanding of human musicality or musical learning, but also other aspects of human mind and subjectivity. In particular, it would be interesting to see (i) whether musical training can enhance the ability to engage in fine motor skills' tasks and (ii) whether the training effect is, for instance, modulated by single or collective musical practice. The results obtained might be beneficial for implementing the

rehabilitative paradigms for patients with motor skills' deficit, providing further evidences for the sensorimotor primacy in musical perception (or, in general, musical sense-making).

The theory and explanatory strategies I have employed throughout this work come from different but partly overlapping frameworks: Enactivism, Phenomenology, Dynamic System Theory. The original interpretation of enactivism that I have endorsed derives not only from its application to the realm of musical experience, but also from the integration in the dynamics of sense-making of the Mirror Mechanism for action understanding, seen as its enabling condition. In this dissertation I have not demonstrated that the enactive approach is *the only* paradigm able to shed light upon the nature of human musicality. It is just a promising - and according to my insights, necessary - step forward towards a better biological and phenomenological description of musical behaviours, which is likely to provide greater explanatory value when compared to classical cognitivism.

Bibliography

- Acitores, A.P. (2011). Towards a theory of proprioception as a bodily basis for consciousness in music. In Clarke D., & Clarke E. (Eds.), *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 215-231). Oxford University Press, Oxford.
- Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, *14*, pp. 43-64.
- Adams, F., & Aizawa, K. (2010). The value of cognitivism in thinking about Extended Cognition. *Phenomenology and the Cognitive Sciences*, *9*(4), pp. 579-603.
- Amedi, A., Stern, W., Camprodon, J.A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C.C., Meijer, P., & Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*, *10*, pp. 687-689.
- Anisfeld, M., Turkewitz, G., Rose, S.A., Faigi, R., Rosenburg, F.R., Sheiber, F.J., Couturier-Fagan, D.A., Ger, J.S., & Sommer, I. (2001). No compelling evidence that newborns imitate oral gestures. *Infancy*, *2*(1), pp. 111-122.
- Arom, S. (2000). Prolegomena to Biomusicology. In Wallin, N., Merker, B., & Brown, S. (Eds.), *The origins of music* (pp. 27-30). MIT Press, Cambridge, MA.
- Auvray, M., Lenay, C., Hanneton, S., Gapenne, O., & Stewart, J. (2005). Spatial perception and sensory substitution. In *International Conference on Perception and Action*, 5-10 July, Monterey, USA.
- Bach-y-Rita, P. (1967). Sensory plasticity. *Acta Neurologica Scandinavica*, *43*, pp. 417-26.
- Bach-y-Rita, P. (1972). *Brain Mechanisms in Sensory Substitution*. Academic Press, New York.
- Bahrack, L.E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In Kail, R. (Ed.), *Advances in Child Development and Behavior*, Vol. 30, (pp. 153-187). Academic Press, New York.
- Baldwin, D.A., & Moses, L. J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, *10*(3), pp. 309-329.
- Bangert, M., & Altenmüller, E. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neuroscience*, *4*(26).
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.J., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage*, *30*, pp. 917-926.
- Bannan, N., & Woodward, S. (2009). Spontaneity in the Musicality and Music Learning of Children. In Malloch, S., & Trevarthen, C. (Eds.), *Communicative Musicality - Exploring the Basis of Human Companionship*. Oxford University Press, Oxford.

- Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioural and Brain Sciences*, 19(1), pp. 107-121.
- Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, pp. 617-645.
- Baum, W.M. (1994). *Understanding Behaviorism*. Harper Collins, USA.
- Beer, R.D. (1990). *Intelligence as Adaptive Behavior*. Academic Press, Boston.
- Beer, R.D. (1995a). A dynamical systems perspective on agent-environment interaction. *Artificial intelligence*, 72, pp. 173-215.
- Bernstein, L. (1976). *The Unanswered Question: Six Talks at Harvard*. Harvard University Press, Cambridge.
- Bernstein, R.J. (1983). *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis*. University of Pennsylvania Press, Philadelphia.
- Blacking, J. (1973). *How Musical is Man?*. University of Washington press, Seattle-London.
- Block, N. (1980). What intuitions about homunculi don't show. *Behavior and Brain Sciences*, 3, open peer commentary on Searle, pp. 425-426.
- Bloem, A. (2012). (E)motion in the XVIIIth Century. A closer look at the changing semantics of the French verbs émouvoir and mouvoir. In Foolen, A., Lüdtke, U., Racine, T., & Zlatev, J. (Eds.), *Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness and Language* (pp. 407-422). J. Benjamins, Amsterdam.
- Bloom, P. (2001). Word learning. *Current Biology*, 11, R5 - R6.
- Bloom, P., & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Sciences*, 2(2), pp. 67-73.
- Boesch, C. (2005). Joint cooperative hunting among wild chimpanzees. Taking natural observations seriously. *Behavioral and Brain Sciences*, 29, pp. 692-693.
- Bohlman, P.V. (2002). *World Music, an Introduction*. Oxford University press, New York.
- Bondioli, A. (1996). *Gioco ed Educazione*. Franco Angeli, Milano.
- Boria, S. (2009). *La comprensione dell'intenzionalità motoria nei bambini con spettro acustico* (PhD dissertation in Neuroscience). University of Parma, Supervisor Prof. Gallese, V.
- Bowlby, J. (1988). *A Secure Base: Parent-Child Attachment and Healthy Human Development*. Basic Books.
- Bowman, W. (2004). Cognition and the Body: Perspectives from Music Education. In Bresler, L. (Ed.), *Knowing Bodies, Moving Minds: Toward Embodied Teaching and Learning* (pp. 29-50). Kluwer Academic Press, Netherlands.

- Braddon-Mitchell, D., & Jackson, F. (1997). *Philosophy of Mind and Cognition*. Blackwell, Oxford.
- Branquinho, J. (2001). *The Foundations of Cognitive Science*. Oxford University Press, Oxford.
- Brentano, F. (1874). *Psychologie Vom Empirischen Standpunkte*. Dunker & Humblot, Leipzig.
- Bretherton, I. (1990). Communication patterns, internal working models, and the intergenerational transmission of attachment relationships, *Infant Mental Health Journal*, 11(3), pp. 237-252.
- Brincker, M. (2010). *Moving beyond mirroring* (PhD Dissertation in Philosophy). City University of New York, Supervisor Prof. Prinz, J.
- Brown, S., & Martinez, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain and Cognition*, 63, pp. 59-69.
- Brown, R.M., & Palmer, C. (2012). Auditory-motor learning influences auditory memory for music. *Memory & Cognition*, 40, pp. 567-578.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C.A., & Rizzolatti, G. (2004a). Neural circuits involved in the recognition of actions performed by non con-specifics: an fMRI study. *Journal of Cognitive Neuroscience*, 16(1), pp. 114-126
- Burge, T. (2010). *The Origins of Objectivity*. Oxford University Press, Oxford.
- Calabretta, R., Di ferdinando, A., Wagner, G.P., & Parisi, D. (2003). What does it take to evolve behaviourally complex organisms?. *BioSystems*, 69, pp. 245-262.
- Calabretta, R., & Parisi, D. (2005). Evolutionary Connectionism and Mind/Brain Modularity. In Callabaut, W., & Rasskin-Gutman D. (Eds.), *Modularity. Understanding the development and evolution of complex natural systems*, The MIT Press, Cambridge.
- Calvert, G.A., Brammer, M.J., & Iversen, S.D. (1998). Crossmodal identification. *Trends in Cognitive Science*, 2, pp. 247-253.
- Carman, T. (1999). The body in Husserl and Merleau-Ponty. *Philosophical Topics*, 27(20), pp. 205-226.
- Carroll-Phelan, B., & Hampson, P.J. (1996). Multiple Components of the Perception of Musical Sequences: A Cognitive Neuroscience Analysis and Some Implications for Auditory Imagery. *Music Perception* 13(4), pp. 517-561.
- Castellano, G., Mortillaro, M., Camurri, A., Volpi, G., & Sherer, K. (2006). Automade analysis of body movement in emotionally expressed piano performances. *Music Perception*, 26, pp. 103-120.
- Chalmers, D. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies* 2(3), pp. 200-219.
- Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, New York.

- Chalmers, D. (2000). What is a neural correlate of consciousness?. In Metzinger T. (Ed.), *Neural Correlates of Consciousness* (pp. 18-39). MIT Press, Cambridge.
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. the MIT Press, Cambridge.
- Chisholm, R.M. (1967). Intentionality. In Edwards, P. (Ed.), *Encyclopaedia of Philosophy*. MacMillan, Oxford.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton & Co, The Hague.
- Chomsky, N. (1965). *Cartesian Linguistics*. Harper and Row, New York.
- Chomsky, N. (1966). *Topics in the Theory of Generative Grammar*. Mouton & Co, The Hague.
- Churchland, P. (1986). *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. The MIT Press, Cambridge.
- Clark, A. (1989). *Microcognition. Philosophy, Cognitive Science and Parallel Distributed Processing*. Mit Press, Cambridge.
- Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. MIT Press, Cambridge.
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press, New York.
- Clark, A. (2009). Spreading the Joy? Why the Machinery of Consciousness is (Probably) Still in the Head. *Mind*, 118(472), pp. 963-993.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analyses*, 58, pp. 7-19.
- Clarke, E. (2005). *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press, Oxford.
- Clearfield, M.W., & Thelen, E. (2001). Stability and flexibility in the acquisition of skilled movement. In Nelson, C.A., & Luciana, M. (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 253-266). MIT Press, Cambridge.
- Cochrane, T. (2010). A Simulation Theory of Musical Expressivity. *Australasian Journal of Philosophy*, 88(2), pp. 191-207.
- Cohen, M.A., & Dennett, D. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, 15, pp. 358-364.
- Colby, C.L. (1998). Action-oriented spatial reference frames in cortex. *Neuron*, 20, pp. 15-24.
- Cole, D. (1991). Artificial Minds: Cam on Searle. *Australasian Journal of Philosophy*, 69, pp. 329-333.

- Colombetti, G. (2010). Enaction, sense-making and emotion. In Stewart, J., Gapenne, O., & Di Paolo, E.A. (Eds.), *Enaction: Toward a New Paradigm for Cognitive Science* (pp. 145-164). MIT Press, Cambridge.
- Colombetti, G. (2014). *The Feeling Body: Affective Science Meets the Enactive Mind*. MIT Press, Cambridge.
- Colombetti, G., & Thompson, E. (2008). The Feeling Body: towards an enactive approach to emotion. In Overton, W.F., Muller, U., & Newman, J.L. (Eds.), *Developmental Perspectives on Embodiment and Consciousness* (pp. 45-68). Lawrence Erlbaum Ass, New York.
- Colombetti, G., & Torrance, S. (2009). Emotion and ethics: an inter(en)active approach. *Phenomenology and the Cognitive Sciences*, 8(4), pp. 505-526.
- Cook, N. (1990). *Music, Imagination and Culture*, Clarendon Press, Oxford.
- Cordeschi, R., & Frixione, F. (2007). Computationalism under attack. In Marraffa, M., De Caro, M., & Ferretti, F. (Eds.), *Cartographies of the Mind: Philosophy and Psychology in Intersection*. Springer.
- Cosmelli, D., & Thompson, E. (2009). Embodiment or envatment? Reflections on the bodily basis of consciousness. In Stewart, J., Gapenne, O., & Di Paolo, E.A. (Eds.), (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press, Cambridge.
- Costantini, M., Ambrosini, E., Scorolli, C., & Borghi, A.M. (2011). When objects are close to me: affordances in the peripersonal space. *Psychonomic Bulletin & Review*, 18, pp. 302-308.
- Crammond, D.J. (1997). Motor imagery: never in your wildest dream. *Trends in Neurosciences*, 20(2), pp. 54-57.
- Crane, T. (1991). All the Difference in the World. *The Philosophical Quarterly*, 41, pp. 1-25.
- Crane, T. (2001). *Elements of mind. An introduction to the Philosophy of mind*, Oxford University Press, Oxford.
- Crane, T. (2006). Brentano's concept of intentional inexistence. In Textor, M. (Ed.), *The Austrian Contribution to Analytic Philosophy*. Routledge, New York.
- Cross, I. (2003). Music, Cognition, Culture and Evolution. In Peretz, I., & Zatorre, R.J. (Eds), *The Cognitive Neuroscience of Music*. Oxford University Press, Oxford.
- Cross, I. (2010). The evolutionary basis of meaning in music: Some neurological and neuroscientific implications. In Clifford F. (Ed.), *The Neurology of Music* (pp. 1-15). Imperial College Press, London.
- Csibra, G., & Gergely, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*, 7, pp. 287-292.
- Currie, G., & Ravenscroft, I. (2002). *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford University Press, Oxford.

- Custodero, L.A. (2009). Intimacy and reciprocity in improvisatory musical performance: Pedagogical lessons from adult artists and young children. In Malloch, S., & Trevarthen, C. (Eds.), *Communicative Musicality: Exploring the Basis of Human Companionship* (pp. 513-530). Oxford University Press, Oxford.
- D'Ausilio, A. (2007). The Role of the Mirror System in Mapping Complex Sounds into Actions. *Journal of Neuroscience*, 27(22), pp. 5847-5848.
- D'Ausilio, A. (2009). Mirror-Like Mechanisms and Music. *The Scientific World Journal*, 9, pp. 1415-1422.
- D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, 24, pp. 955-958.
- Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam, New York.
- Damasio, A. (2003). *Looking for Spinoza. Joy, Sorrow and the Feeling Brain*. Harcourt, Orlando
- Darwin, C. (1871). *The Descent of Man and the Selection in Relation to Sex*. Murray, London.
- Darwin, C. (1872). *Expression of the Emotions in Man and Animals*. Murray, London.
- Darwin, F. (Ed.), (1909). *The Foundation of the Origins of Species. Two essays in 1842 and 1844*. Cambridge University Press, London.
- Daum, M.M., Prinz, W., & Aschersleben, G. (2008). Encoding the goal of an object-directed but uncompleted reaching action in 6- and 9-month-old infants. *Developmental Science*, 11, pp. 607-619.
- Davies, M. (1994). The mental simulation debate. In Peacocke, C. (Ed.), *Objectivity, Simulation and the Unity of Consciousness, Current Issues in the Philosophy of Mind* (pp. 99-127). Oxford University Press, Oxford.
- De Jaegher, H. (2006). *Social interaction rhythm and participatory sense-making: An embodied, interactional approach to social understanding, with implications for autism* (PhD dissertation). University of Sussex, Brighton, UK.
- De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition*, 18(2), pp. 535-542.
- De Jaegher, H. (2010). Enaction versus representation: An opinion piece. In Fuchs, T., Sattel, H., & Henningsen, P. (Eds.), *The embodied self: Dimensions, coherence and disorders* (pp. 218-224). Schattauer, Stuttgart.
- De Jaegher, H., & Di Paolo, E.A. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), pp. 485-507.
- De Preester, H., & Tsakiris, M. (2009). *Body-extension versus body-incorporation: Is there a need for a body-model?*. *Phenomenology and the Cognitive Sciences*, 8(3), pp. 307-319.

- Delalande, F. (Ed.), (2009). *La Nascita della Musica*, FrancoAngeli, Milan.
- Dennett, D. (1980). The milk of human intentionality. *Behavior and Brain Sciences*, 3, open peer commentary on Searle, pp. 428-430.
- Dennett, D. (1987). *The Intentional Stance*. MIT Press, Cambridge.
- Dennett, D. (1988). Quining Qualia. In Marcel, A.J & Bisiach, E. (Eds), *Consciousness in Modern Science*. Oxford University Press, Oxford.
- Dennett, D. (1991). *Consciousness Explained*. Little Brown, Boston.
- Dennett, D. (1993). The Embodied Mind. In Varela, F., Thompson, E., & Rosch, E. (Eds.), *American Journal of Psychology*, 106, pp. 121-126.
- Dennett, D. (1994). Cognitive science as reverse engineering: Several meanings of “top-down” and “bottom- up”. In Prawitz, D., Skyrms, B., & Westerståhl, D. (Eds.), *Logic, Methodology and Philosophy of Science IX, D* (pp. 679-689). The Netherlands Elsevier Science, Amsterdam.
- Descartes, R. (1637 [1960]). Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences. [In Lafleur, L.J. (Ed.), *Discourse on Method and Meditations*. The Liberal Arts Press, New York].
- Descartes, R. (1641 [1961; 1980; 1982]). Meditations on First Philosophy. [In Cottingham, J., Stoothoff, R., & Murdoch, D. (Eds), *The Philosophical Writings of René Descartes, Vol. 2* (pp. 1-62). Cambridge University Press, Cambridge.
- Dewey, J. (1916). *Democracy and education. An introduction to the philosophy of education*. The MacMillan Company, New York.
- Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), pp. 429-452.
- Di Paolo, E. (2009). Extended life. *Topoi*, 28, pp. 9-21.
- Di Paolo, E., Rohde, M., & De Jaegher, H., (2010). Horizons for the Enactive Mind: Values, Social Interaction, and Play. In Stewart, J., Gapenne, O., & Di Paolo, E.A. (Eds.), *Enaction: Toward a New Paradigm for Cognitive Science* (pp. 33- 87). MIT Press, Cambridge.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain research*, 91, pp. 176-180.
- Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interactions. In Wallin, N., Merker, B., & Brown, S. (Eds.), *The origins of music*. MIT Press, Cambridge, MA.
- Dreyfus, H.L. (1995). The Current Relevance of Merleau-Ponty's Phenomenology of Embodiment. *Filozofska Istrazivanja*, 15(3), pp. 385-399.

- Dreyfus, H.L. (2002). Intelligence without representation - Merleau-Ponty's critique of mental representation. The relevance of phenomenology to scientific explanation. *Phenomenology and the Cognitive Sciences*, 1(4), pp. 367-383.
- Dunbar, R. (1996). *Grooming, Gossip and the Evolution of Language*. Harvard University Press, Cambridge.
- Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, 114, pp. 405-422.
- Ellis, R.D., & Newton, N. (2012). Could moving ourselves be the link between emotion and consciousness?. In Foolen, A., Lüdtke, U., Racine, T., & Zlatev, J. (Eds.), *Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness and Language* (pp. 57-80). J. Benjamins, Amsterdam.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor Facilitation During Action Observation: A Magnetic Stimulation Study. *Journal of Neurophysiology*, 73, pp. 2608-11.
- Feld, S. (1982). *Sound and Sentiment: Birds, Weeping, Poetics, and Song in Kaluli expression*. University of Pennsylvania Press, Philadelphia.
- Ferretti, F. (2007). *Perchè Non Siamo Speciali. Mente, Linguaggio e Natura Umana*. Laterza, Bari.
- Fisher, R.A. (1915). The evolution of sexual preference. *Eugenics Review*, 7, pp. 184-192.
- Fishkin, K.P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, 8(5), pp. 347-358.
- Fitch, W.T. (2006). The biology and evolution of music: a comparative perspective. *Cognition*, 100, pp. 173-215.
- Fodor, J. (1968). *Psychological Explanation*. Random House, New York.
- Fodor, J. (1980a). Methodological solipsism considered as a research strategy in Cognitive science. *Behavioral and Brain Sciences*, 3, pp. 63-73.
- Fodor, J. (1980b). Searle on what only brains can do. *Behavior and Brain Sciences*, 3, open peer commentary on Searle. pp. 431-432.
- Fodor, J. (1981 [1994]). The mind-body problem. In Szubka, T., & Warner, R. (Eds.), *Scientific American*, 244, pp. 114-125, [*The Mind-Body Problem*. (1994). Blackwell, Oxford].
- Fodor, J. (1983). *The Modularity of Mind*. MIT Press, Cambridge.
- Fogassi, L., & Luppino, G. (2005). Motor functions in the parietal lobe. *Current opinion in Neurobiology*, 15, pp. 626-631.
- Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Charsi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308, pp. 662-667.

- Foolen, A., Lüdtke, U., Racine, T., & Zlatev, J. (Eds.), (2012). *Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness and Language*. J. Benjamins, Amsterdam.
- Fraisse, P. (1974). *Psychologie du Ritme*. PUF, Parigi.
- Freeman, W. (1999). *How Brains Make up Their Minds*. Orion Press, London.
- Freeman, W. (2000). A Neurobiological Role of Music in Social Bonding. In Wallin, N., Merker, B., & Brown, S. (Eds.), *The origins of music*. MIT Press, Cambridge.
- Freschi, A.M. (2006). *Movimento e Misura. Esperienza e Didattica del Ritmo*. EDT, Torino.
- Froese, T. (2009). *Sociality and the Life-Mind Continuity Thesis: A Study in Evolutionary Robotics* (PhD Dissertation). University of Sussex, Brighton, UK.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Science*, 4(1), 14-21.
- Gallagher, S. (2005). Dynamic models of body schematic processes. In De Preester, H., & Knockaert, V. (Eds), *Body Image and Body Schema* (pp. 233-250). J. Benjamins, Amsterdam.
- Gallagher, S. (2011). Interpretations of embodied cognition. In Tschacher, W., & Bergomi, C. (Eds.), *The Implications of Embodiment: Cognition and Communication* (pp. 59-71). Imprint Academic, Exeter.
- Gallagher, S. (2012). Neurons, neonates and narrative: From embodied resonance to empathic understanding. In Foolen, A., Lüdtke, U., Racine, T., & Zlatev, J. (Eds.), *Moving Ourselves, Moving Others: Motion and Emotion in Intersubjectivity, Consciousness and Language* (pp. 167-196). J. Benjamins, Amsterdam.
- Gallagher, S., & Hutto, D. (2008). Understanding others through Primary Interaction and Narrative Practice. In Zlatev, T., Racine, T., Sinha, C., & Itkonen, E. (Eds.), *The Shared Mind: Perspectives on Intersubjectivity* (pp. 17-38). J. Benjamins, Amsterdam.
- Gallagher, S., Hutto, D., Slaby, J., & Cole, J. (2013). The brain as part of an enactive system. Commentary to appear in *Behavioral and Brain Science*.
- Gallagher, S., & Zahavi, D. (2008). *The Phenomenological Mind*. Routledge, London.
- Gallese, V. (2001). The “Shared Manifold” Hypothesis: from mirror neurons to empathy. *Journal of Consciousness Studies*, 8(5-7), pp. 33-50.
- Gallese, V. (2005). Embodied simulation: from neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, 4, pp.23-48.
- Gallese, V. (2009). Motor abstraction: a neuroscientific account of how action goals and intentions are mapped and understood. *Psychological Research*, pp.73: 486-498.
- Gallese, V. (2010). Neuroscienze e fenomenologia. In *Enciclopedia Treccani terzo Millennio*,

- Gallese, V. (2011). Embodied Simulation Theory: Imagination and narrative. A commentary on Siri Hustvedt. *Neuropsychoanalysis*, 13(2), pp. 196-200.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, pp. 593-609.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In Prinz, W., & Hommel, B. (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance*. Oxford University Press, Oxford.
- Gallese, V., Gernsbacher, M.A., Heyes, C., Hickok, G., & Iacoboni, M. (2011). Mirror neuron forum. *Perspective on Psychological Science*, 6(4), pp.369-407.
- Gallese, V., & Goldman, A.I. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493-501.
- Gallese, V., & Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Reason and Language. *Cognitive Neuropsychology*, 22, pp. 455-479.
- Gallese, V., Rochat, M., Cossu, G., & Sinigaglia, C. (2009). Motor cognition and its role in the phylogeny and ontogeny of intentional understanding. *Developmental Psychology*, 45, pp. 103-113.
- Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation?. *Trends in cognitive science*, 15(11), pp. 512-519.
- Garvey, A. & Fogel, A. (2008). Emotions and Communication as a Dynamic Developmental System. *Espaciotiempo 2: Dossier: Enfoques de la complejidad y el desarrollo en las Humanidades y las Ciencias Sociales*, pp. 62-73.
- Geeves, A., McIlwain, D., Sutton, J., & Christensen, W. (2009). Expanding Expertise: Investigating a Musician's Experience of Music Performance. In W. Christensen, E., Schier, & J. Sutton, (Eds.), *ASCS09: Proceedings of the 9th Conference of the Australasian Society for Cognitive Science* (pp. 106-113). Macquarie Centre for Cognitive Science, Sydney, Australia.
- Geissmann, T. (2000). Gibbon song and human music. In Wallin, N., Merker, B., & Brown, S. (Eds.), *The Origins of Music*. MIT Press, Cambridge, MA.
- Gerson, S. (in press). Sharing and comparing: How comparing shared goals broadens goal understanding in development. *Child Development Perspectives*.
- Gerson, S., & Woodward, A.L. (2012). A claw is like my hand: Comparison supports goal analysis. *infants Cognition*, 122, pp. 181-192.
- Gerson, S., & Woodward, A.L. (in press). The joint role of trained, untrained, and observed actions at the origins of goal recognition. *Infant Behaviour and Development*.
- Ghigi, N. (2007). *La Metafisica di Edmund Husserl*. FrancoAngeli, Milan.
- Giannattasio, F. (1998). *Il Concetto di Musica*. Bulzoni, Rome.

- Giannattasio, F. (2005). Il concetto di musica in una prospettiva culturale. In Nattiez, J.J. (Ed.) *Enciclopedia della Musica, in X Vol.* Einaudi, Torino.
- Gibson, E.J. (1982). The concept of affordances in development. The renaissance of functionalism. In Collins, A.W. (Ed.), *The concept of development. The Minnesota symposia on child psychology (15)*, Lawrence Erlbaum, Hillsdale, NJ.
- Gibson, E.J. (1988). Exploratory behaviour in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, 39, pp. 1-41.
- Gibson, E., & Gibson, J.J. (1955). Perceptual learning: differentiation or enrichment?. *Psychology Review*, 62, pp. 32-41.
- Gibson, J.J. (1977). The theory of affordances. In Shaw, R., & Bransford, J. (Eds.), *Perceiving, Acting and Knowing: Toward an Ecological Psychology* (pp. 67-82). Lawrence Erlbaum, Hillsdale.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin Company, Boston.
- Glaserfeld, E. Von (1988). The reluctance to change the way of thinking. *Irish Journal of Psychology*, 9(1), pp. 83-90. Available at <http://www.vonglasersfeld.com/111>
- Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press, Cambridge.
- Godøy, R.I. (2003). Motor-Mimetic Music Cognition. *Leonardo* 36(4), pp. 317-319.
- Godøy, R.I. (2009). Geometry and Effort in Gestural Renderings of Musical Sounds. In Sales Dias, M., Gybet, S., Wanderlay, M.M., & Bastos, R. (Eds.), *Gesture-Based Human-Computer Interaction and Simulation. 7th International Gesture Workshop, GW 2007*, Lisbon, Portugal, May 23-25, 2007, Revised Selected Papers (pp. 205-215). Springer.
- Goldberg, S., & Pessin, A. (1997). *Gray Matters. An Introduction to the Philosophy of Mind*. Sharpe, New York.
- Goldman, A.I. (2002). Simulation Theory and Mental Concepts. In Dokic, J., & Proust, J. (Eds.), *Simulation and Knowledge of Action* (pp. 1-19). J. Benjamins, Amsterdam.
- Goldman, A.I. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford University Press, New York.
- Goldman, A.I. (2012). A Moderate Approach to Embodied Cognitive Science. *Review of Philosophy and Psychology*, 3(1), pp. 71-88.
- Gordon, R.M. (1986). Folk psychology as simulation. In *Mind and Language*, 1, pp. 158-171 [In Davies, M., & Stone, T. (Eds.), *Folk Psychology* (pp. 60-73). Blackwell, Oxford, 1995]
- Gregoriou, G.G., Borra, E., Matelli, M., & Luppino, G. (2006). Architectonic organization of the inferior parietal convexity of the macaque monkey. In *Journal of Comparative Neurology*, 496, pp. 422-451.

- Griffiths, P. (1995). *Modern Music and After: Directions since 1945*. Oxford University Press, Oxford - New York.
- Gruhn, W. (2006). Music learning in schools: Perspectives of a new foundation for music teaching and learning. *Action, Criticism, and Theory for Music Education*, 5(2), pp. 2-27.
- Hanna, R., & Thompson, E. (2003). Neurophenomenology and the spontaneity of consciousness. *Canadian journal of Philosophy*, 29, pp. 133-162.
- Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behavior*, 59, pp. 771-785.
- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A.O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, 17(2), pp. 282-293.
- Hauelsen, J., & Knösche, T.R. (2001). Involuntary motor activity in pianist evoked by music perception. *Journal of Cognitive Neuroscience*, 1(3), pp. 235-241.
- Hauser, M.D., & McDermott, J. (2003). The evolution of the music faculty. A comparative perspective. *Nature Neuroscience*, 6, pp. 663-668.
- Heidegger, M. (1927). Sein und Zeit. In *Jahrbuch für Philosophie und phänomenologische Forschung*. Max Niemeyer. Halle. [*Being and time*. Harper and Raw, New York 1962].
- Held, R. (1965). Plasticity in the Sensorimotor System. *Scientific American*, 213(15), pp. 84-94.
- Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, 56, pp. 872-876.
- Helmholtz, H. Von (1863). *Die Lehre von der Tonempfindungen als psychologische Grundlage für die Theorie der Musik*. Olms, Hildensheim.
- Herholtz, S., Lippe C., Knief, A., & Pantev, C. (2008). Neural basis of musical imagery and the effect of musical expertise. *European Journal of Neuroscience*, 28, pp. 2352-2360.
- Hill, C. (2009). *Consciousness*. Cambridge University Press, Cambridge.
- Himberg, T. (2006). *Co-operative tapping and collective time-keeping. Differences of timing accuracy in duet performance with human or computer partner*. Paper presented at IX International Conference on Music Perception and Cognition. Bologna, Italy.
- Hogg, B. (2011). Enactive consciousness. Intertextuality, and musical free improvisation: Deconstructing mythologies and finding connections. In Clarke, D., & Clarke, E. (Eds.), *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 79-94). Oxford University Press, Oxford.
- Holmes, T. (1985). *Electronic and Experimental Music: Technology, Music, and Culture*. Scribner, New York.

- Honing, H., & Ploeger, A. (2012). Cognition and the Evolution of Music. *Topics in Cognitive Science*, pp. 1-12.
- Hurley, S. (1998). *Consciousness in Action*. Harvard University Press, London.
- Hurley, S. (2001). Perception and action. Alternative views. *Synthese*, 129, pp. 3-40.
- Hurley, S. (2010). The Varieties of Externalism. In Menary, R. (Ed.), *The Extended Mind*. MIT press, Cambridge, MA.
- Huron, D. (2003). Is Music an evolutionary adaptation?. In Peretz, I., & Zatorre, R.J. (Eds), *The Cognitive Neuroscience of Music* (pp. 57-74). Oxford University Press, Oxford.
- Huron, D. (2006). *Sweet Anticipation. Music and the Psychology of Expectation*. MIT Press, Cambridge, MA.
- Husserl, E. (1901). *Logische Untersuchungen. Zweiter Teil: Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*. Max Niemeyer, Halle. [*Logical investigations. Second part. Investigations concerning phenomenology and the theory of knowledge*. Martinus Nijhoff, The Hague, 1984].
- Husserl, E. (1907). *Die Idee der Phänomenologie. Fünf Vorlesungen*. [*The idea of phenomenology. Five lectures*. Martinus Nijhoff, The Hague, 1964].
- Husserl, E. (1912-1929). *Ideen zur einer reinen Phänomenologie und phänomenologischen Philosophie*. [*Ideas pertaining to a pure phenomenology and to a phenomenological philosophy*. [First Book: *General introduction to a pure phenomenology*; Second Book, *Phenomenological investigations of constitution*; Third Book, *Phenomenology and the foundations of the sciences*] Martinus Nijhoff, The Hague [respectively], 1988, 1952, 1971.
- Husserl, E. (1925). *Phänomenologische Psychologie*. Published post-mortem by Martinus Nijhoff, The Hague, 1962. [*Phenomenological Psychology*. Martinus Nijhoff, The Hague, 1977].
- Husserl, E. (1931). *Méditations Cartésiennes*. Armand Colin, Paris. [*Cartesian Meditations*. Kluwer, Dordrecht, 1988].
- Husserl, E. (1935). *The Vienna Lecture*. Lecture delivered by Edmund Husserl, Vienna, 10 May 1935.
- Husserl, E. (1936). *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie. Eine Einleitung in die phänomenologische Philosophie*. In *Philosophia I*. [The crisis of European sciences and transcendental philosophy. An Introduction to Phenomenology. Martinus Nijhoff, The Hague, 1976].
- Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science*, 2(4), pp. 705-715.
- Hutto, D. (2006). Turning hard problems on their heads. *Phenomenology and the Cognitive Sciences*, 5(1), pp. 75-88.
- Hutto, D. (2012). Truly Enactive Emotion. *Emotion Review*. 4(2), pp. 176-181.

- Hutto, D. (2013). Radically enactive cognition in our grasp. In Radman, Z. (Ed.), *The Hand: An Organ of the Mind* (pp. 227-252). MIT Press, Cambridge, MA.
- Hutto, D., & Myin, E. (2013). *Radicalizing Enactivism*. MIT Press, Cambridge, MA.
- Iacoboni, M. (2009). *Mirroring People: The Science of Empathy and How We Connect with Others*. Picador, US.
- Imberty, M. (1983). *Les écritures du temps. Sèmantique Psychologique de la Musique*. Bordas, Paris.
- Iyer, V. (2002). Embodied Mind, Situated Cognition, and Expressive Microtiming in African-American Music. *Music Perception*, 19(3), pp. 387-414.
- Iyer, V. (2004). Improvisation, temporality and embodied experience. *Journal of Consciousness Studies*, 11(3-4), pp. 159-173.
- Jackendoff, R. (1987). *Consciousness and the Computational Mind*. MIT Press, Cambridge.
- Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, 32(127), pp. 127-136.
- Jacob, P. (2008). What Do Mirror Neurons Contribute to Human Social Cognition?. *Mind and Language*, 23(2), pp. 190-223.
- James, W. (1890). *The Principles of Psychology*. Henry Holt, New York.
- Jänke, L. (2006). From cognition to action. In Altenmüller, E., Kesselring, J., & Wiesendanger, M. (Eds.), *Music, Motor control and the Brain* (pp. 25-36). Oxford University Press, New York.
- Janik, V.M., & Slater, P.B. (1997). Vocal learning in mammals. *Advanced in the Study of Behavior*, 26, pp. 59 - 99.
- Jeannerod, M., Arbib, M.A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neuroscience*, 18, pp. 314-320.
- Johnson, M. (1987). *The Body in the Mind. The bodily basis of meaning, imagination, and reason*. University of Chicago Press, Chicago.
- Juslin, P.N. (2001). Communicating emotion in music performance. A review and theoretical framework. In Juslin, P.N., & Sloboda, J.A. (Eds.), *Music and emotion: theory and research* (pp. 309-337). Oxford University Press, Oxford.
- Juslin, P.N., & Sloboda, J.A. (Eds.). (2001). *Music and Emotion: Theory and Research*. Oxford University Press, Oxford.
- Juslin P.N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, 8, pp. 668-683.
- Karmiloff-Smith, A., & Thomas, M.S.C. (2003). What can developmental disorders tell us about the neurocomputational constraints that shape development? The case of Williams syndrome. *Development and Psychopathology*, 15, pp. 969-990.

- Keller, P. E., & Janata, P. (2009). Review of Embodied Music Cognition and Mediation Technology. *Music Perception*, 26(3), pp.289-292.
- Kemler Nelson, D.G. (1999). Attention to functional properties in toddlers naming and problem-solving. *Cognitive Development*, 14, pp. 77-100.
- Keysers, C. (2007). Mirror neurons. *New Encyclopaedia of Neuroscience*. Elsevier Press, Amsterdam.
- Kirschner, S., & Tomasello, M. (2009a). Joint drumming. Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, 102, pp. 299-314.
- Kirschner, S., & Tomasello, M. (2009b). *Joint Music making promotes prosocial behavior among four-year-old*. Paper presented at II European Conference on the Developmental Psychology of Music. Roehampton University, London.
- Kivy, P. (1959). Charles Darwin on music. *Journal of the American Musicological Society*, 12, pp. 42-48.
- Kivy, P. (2001). *New Essays on Musical Understanding*. Clarendon Press, Oxford.
- Kivy, P. (2007). *Music, Language, and Cognition: And Other Essays in the Philosophy of Music, further collected essays of Peter Kivy*. Oxford university press, Oxford.
- Knudsen, C. (1982). Intentions and impositions. In Kretzmann, N., Kenny, A., & Pinborg, J. (Eds.), *The Cambridge History of Later Medieval Philosophy*. Cambridge University Press, Cambridge.
- Koelsch, S., Schroëger, E., & Tervaniemi, M. (1999). Superior attentive and pre-attentive auditory processing in musicians. *Neuroreport*, 10, pp. 1309-1313.
- Koeneke, S., Lutz, K., Wustenberg, T., & Jancke, L. (2004). Long-term training affects cerebellar processing in skilled keyboard players. *Neuroreport*, 15, pp. 1279–1282.
- Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297, pp. 846-848.
- Kristeva, R., Chakarov, V., Schulte-Mönting, J., & Spreer, J. (2003). Activation of cortical areas in music execution and imagining: a high-resolution EEG study. *Neuroimagine*, 20, pp. 1872-1883.
- Krueger, J. (2009). Enacting musical experience. *Journal of Consciousness Studies*, 16(2-3), pp. 98-123.
- Krueger, J. (2011a). Doing things with music. *Phenomenology and the Cognitive Sciences*, 10(1), pp. 1- 22.
- Krueger, J. (2011b). Enacting musical content. In Manzotti, R. (Ed.), *Situated Aesthetics: Art beyond the Skin*. Imprint Academic, Exeter.

- Krueger, J. (2013). Empathy, Enaction, and Shared Musical Experience. In Cochrane, T., Fantini, B., & Scherer, K. (Eds.), *The Emotional Power of Music: Multidisciplinary Perspectives on Musical Expression, Arousal, and Social Control* (pp. 177-196). Oxford University Press, Oxford.
- Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology*, 4, pp. 1-13.
- Kyselo, M & Di Paolo, E. (2013). Locked-in Syndrome. A Challenge for Embodied Cognitive Science. *Phenomenology and the Cognitive Science*, doi: 10.1007/s11097-013-9344-9.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, 27, pp. 308-314.
- Lakoff, G., & Johnson M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Basic Books, New York.
- Langmore, N.E. (1998). Function of duets and solo songs of female birds. *Trends in Ecology and Evolution*, 13, pp. 136-140.
- Langmore, N.E. (2000). Why female birds sing. In Espmark, Y., Amundsen, T., & Rosenqvist, G. (Eds.), *Signaling and signal design in animal communication*. Tapir Academic Press, Tapir.
- Laske, O. (1977). *Music, memory and thought*. University of Microfilms International. Ann Arbor MI.
- Leakey, R. (1994). *The origins of humankind*. Basic Books, New York.
- Legrand, D. (2010). Myself with no body? Body, bodily-consciousness and self-consciousness. In Schmicking, D., & Gallagher, S. (Eds.), *Handbook of Phenomenology and Cognitive Science* (pp. 181-200). Springer.
- Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. MIT Press, Cambridge, MA.
- Leman, M., Desmet, F., Styns, F., Van Noorden, L., & Moelants, D. (2009). Sharing musical expression through embodied listening: a case study based on Chinese Guqin Music. *Music Perception*, 26, pp. 263-268.
- Leman, M., & Schneider, A. (1997). Systematic, cognitive and historical approaches in musicology. In Leman, M. (Ed.), *Music, gestalt and computing* (pp. 13-29). Springer, Berlin.
- Lerdahl, F. (2009). Genesis and architecture of the GTTM project. *Music Perception*, 26, February, pp. 187-194.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press, Cambridge, MA.
- Levine, J. (1983). Materialism and the Explanatory Gap. *Pacific Philosophical Quarterly*, 64, pp. 354-361.

- Lewitin, D.J. (2006). *This is Your Brain on Music. The Science of a Human Obsession*. Dutton, New York.
- Lewis, C.W. (1929). *Mind and the World Order: Outline of a Theory of Knowledge*. [Dover reprint, 1956].
- Lewontin, R. (1983). The Organism as Subject and Object of Evolution. *Scientia*, 118(1), pp. 65-82.
- Lewontin, R. (1998). The evolution of cognition: Questions we will never answer. In Scarborough, D., & Sternberg, S. (Eds.), *An invitation to cognitive science, Volume 4: Methods, models, and conceptual issues*. MIT Press, Cambridge.
- Lewkowicz, D.J (1996). Perception of auditory-visual temporal synchrony in human infants. *Journal of experimental Psychology: Human Perception and Performance*, 22, pp. 1904-1106.
- Lidji P., Kolinsky R., Lochy A., & Morais, J. (2007). Spatial associations for musical stimuli: a piano in the head?. *Journal of Experimental Psychology: Human Perception and Performance*, 33, pp. 1189-1207.
- Lipps, T. (1903). Einfühlung, innere Nachahmung und Organempfindung. *Archiv für die gesamte Psychologie*, pp. 185-204.
- Locke, J. (1689). *An Essay Concerning Human Understanding*. Revised editions, 1694, 1700; edited by Peter Nidditch, 1975; Essay draft A edited by R.I. Aaron and Jocelyn Gibb, 1936; Essay draft B as An Essay Concerning the Understanding, Knowledge, Opinion, and Assent, edited by Benjamin Rand, 1931; drafts A and B edited by Nidditch, 2 vols. 1980-82.
- Lotze, M., Scheler, G., Tan, H.R.M., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, 20, pp. 1817-1829.
- Majewskam A., & Sur, M. (2006). Plasticity and specificity of cortical processing networks. *TINS*, 29, pp. 323-329.
- Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, 8(2), pp. 79-86.
- Marler, P. (1970b). A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Physiological Psychology Monographs*, 71, pp. 1-25.
- Maturana, H. (1990). *Biology of Cognition and Epistemology*. Ed Universidad de la Frontera, Temuco, Chile.
- Maturana, H., & Varela, F. (1980). Autopoiesis and Cognition: the Realization of the Living. In Cohen, S., & Wartofsky, M.W. (Eds.), *Boston Studies in the Philosophy of Science*. D. Reidel Publishing Co., Dordecht.
- Maturana, H., & Varela, F. (1992). *The tree of knowledge: The biological roots of human understanding* (rev. edition). Shambhala, Boston.

- McDermott, J., & Hauser, M.D. (2005). The origins of music: innateness, uniqueness and evolution. In *Music Perception*, 23, 1, pp. 29-59.
- McGee, K. (2005). Enactive Cognitive Science. Part I: Background and Research Themes. *Constructivist Foundation*, 1(1), pp. 19-34.
- McIntyre, R., & Woodroof Smith, D. (1989). Theory of intentionality. In Mohanty, J.N., & McKenna, W.R. (Eds.), *Husserl's Phenomenology: a textbook* (pp. 147-179). Center for Advanced Research in Phenomenology and University Press of America, Washington D.C.
- McMullen, E., & Saffran, J.R.(2004). Music and language: a developmental comparison. *Music Perception*, 21(3), pp. 289-311.
- Meijer, P. (1992). An experimental system for auditory image representations. In *IEEE Transaction on Biomedical Engineering*, 39 (2), pp. 112-121.
- Melis, A., Call, J., & Tomasello, M. (2006). Chimpanzees conceal visual and auditory informations from others. *Journal of Comparative Psychology*, 120, pp. 154-162.
- Meltzoff, A.N. (2005). Imitation and other minds: The 'like me' hypothesis. In Hurley, S., & Chater, N. (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science, Vol.2* (pp. 55-77). MIT Press, Cambridge.
- Meltzoff, A.N., & Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), pp. 75-78.
- Menin, D. (2009). *Cenni sull'intenzionalità musicale* (BA Dissertation in Philosophy of Science). University of Milan, Supervisor: Prof. Sinigaglia, C.
- Menin, D. (2011). *Affordances musicali* (MA Dissertation in Philosophy of Science). University of Milan. Supervisor: Prof. Sinigaglia, C.
- Merker, B. (2000). Synchronous Chorus and Human Origins. In Wallin, N., Merker, B., & Brown, S. (Eds.), *The origins of music*. MIT Press, Cambridge, MA.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Gallimard. Paris; [*Phenomenology of perception*. Routledge & Kegan. London 1962].
- Merriam, A.P.(1964). *The Anthropology of Music*. NorthWestern University Press, Evanston.
- Messiaen, O. (1944). *Tecniche de Mon Language Musical*. 2 Voll. Leduc. Paris.
- Metzinger, T. (2003). *Being No One: The Self-Model Theory of Subjectivity*. MIT Press, Cambridge.
- Meyer, L. (1956). *Music and Emotions*. University Press of Chicago. Chicago.
- Miller, G. (2000). Evolution of music through Sexual Selection. In Wallin, N., Merker, B., & Brown, S. (Eds.), *The origins of music*. MIT Press, Cambridge.

- Mithen, S. (2005). *The Singing Neanderthals. The Origins of Music, Language, Mind and Body*. Orion. London.
- Molino, J. (1975). Fait musical et sémiologie de la musique. *Musique en Jeu*, 17, pp. 37-62.
- Molnar-Szakacs, I., & Overy, K. (2006). Music and Mirror Neurons. From Motion to E'Motion. *Social Cognitive and Affective Neuroscience*, 1(3), pp. 235-241.
- Monelle, R. (1992). *Linguistics and Semiotics in Music*. Harwood Academic Publishers, Chur.
- Moran, D. (2000). *Introduction to Phenomenology*. Routledge. London - New York.
- Morgan, E. (1995). *The Descent of the Child. Human Evolution from a new Perspective*. Oxford University Press, New York.
- Münste, T.F., Altenmüller, E., & Jänke, L. (2002). The musician's brain as a model for neuroplasticity. *Nature Review. Neuroscience*, 3, pp. 473-478.
- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, 78, pp. 2226-2230.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of Neurophysiology*, 79, pp. 2580-2601.
- Myin, E., & O'Regan, K. (2009). Situated perception and sensation in vision and other modalities: form an active to a sensorimotor account. In Robbins, P., & Aydede, A. (Eds.), *Cambridge Handbook of Situated Cognition*. (pp. 185-200). Cambridge University Press, Cambridge.
- Nagel, T. (1974). What is it like to be a Bat?. *Philosophical Review*, 83, pp. 435-456.
- Nattiez, J.J. (1987). *Musicologie Générale et Sémiologie*. Christian Bourgois Editeur, Paris. [*Music and discourse: toward a semiology of music*. Princeton University Press, Princeton. 1990].
- Neisser, U. (1967). *Cognitive Psychology*. Appleton-Century-Crofts, New York.
- Nettl, B. (2005). *The Study of Ethnomusicology: 31 issues and concepts*. University of Illinois Press, Urbana-Chicago
- Newell, A. (1982). The knowledge level. *Artificial Intelligence*, 18, pp. 87-127.
- Noë, A. (2000). Experience and experiment in art. *Journal of Consciousness Studies*, 7(8-9), pp. 123-135.
- Noë, A. (2004). *Action in perception*. MIT Press, Cambridge.
- Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Hill and Wang, New York.
- Noë, A., & Thompson, E. (2004a). Sorting Out the Neural Basis of Consciousness. Authors' Reply

- to Commentators. *Journal of Consciousness Studies*, 11, pp. 87-98.
- Noë, A., & Thompson, E. (2004b). Are There Neural Correlates of Consciousness?. *Journal of Consciousness Studies*, 11, pp. 3-28. Special Target Article with Peer Commentary.
- Nonken, M. (2008). What Do Musical Chairs Afford? On Clarke's Ways of Listening and Sacks's Musicophilia. *Ecological Psychology*, 20, pp. 283-295.
- Nussbaum, C.O. (2007). *The Musical Representation: Meaning, Ontology, and Emotion*. MIT Press, Cambridge.
- O'Regan, J.K., & Noë, A. (2001a). A sensorimotor approach to vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), pp.939-973.
- O'Regan, J.K., & Noë, A. (2001b). What it is like to see: A sensorimotor theory of visual experience. *Synthèse*, 129(1), pp. 79-103.
- Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception*, 26(5), pp. 489-504.
- Oyama, S. (2000). *The Ontogeny of Information. Developmental Systems and Evolution*. Duke University Press, Durham, US.
- Palmer, C., & Meyer, R.K. (2000). Conceptual and motor learning in music performance. *Psychological Science*, 11, pp. 63-68.
- Parncutt, R. (2007). Systematic musicology and the history and future of Western musical scholarship. *Journal of Interdisciplinary Music Studies*, 1, pp. 1-32.
- Payne, K.B., & McVay, S. (1971). Songs of humpback whales. *Science*, 173, pp. 583-597.
- Pelinski, R. (2005).. Embodiment and musical experience. *Trans.Transcultural music review*, 9. Available at http://www.redalyc.uaemex.mx/redalyc/html/822/82200914/82200914_2.html
- Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, 10, pp. 21-56.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6(7), pp. 688-691.
- Peretz, I., Gaudreau, D., & Bonnel, A.M. (1998). Exposure effects on music preference and recognition. *Memory and Cognition*, 26, pp. 884-902.
- Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary Music Review*, 4, pp. 277-291.
- Peretz, I., & Zatorre, R.J. (Eds). (2003). *The Cognitive Neuroscience of Music*. Oxford University Press, Oxford.
- Perone, S., Madole, K.L., Ross-Sheehy, S., Carey, M., & Oakes, L.M. (2008). The relation between infants' activity with objects and attention to object appearance. *Developmental Psychology*, 44, pp. 1242-1248.

- Perone, S., & Oakes, L.M. (2006). It clicks when it is rolled and squeaks when it is squeezed: What 10-month-old infants learn about function. *Child Development*, 77, pp. 1608-1622.
- Philips-Silver, J., & Trainor, L.J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, pp. 533-546.
- Piana, G. (2007). *Barlumi per una Filosofia della Musica*. Available at <http://www.filosofia.unimi.it/piana/barlumi/pdf/barlumi.pdf>
- Piaget, J. (1964). *Six études de Psychologie*. Gonthier, Genève.
- Piccirilli, M., Sciarma, T., & Luzzi, S. (2000). Modularity of music: evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery & Psychiatry*, 69, pp. 541-545.
- Pierce, J.R. (1983). *The Science of Musical Sounds*. W.H. Freeman and Company, New York.
- Pinker, S. (1997). *How the mind works*. W.W. Norton, New York.
- Poincaré, J.H. (1908). *Science et method*. Edition Ernest Flammarion, Paris. [*Science and Method*. Routledge, London 1996].
- Poincaré, J.H. (1913). *Dernieres Pensees*. Edition Ernest Flammarion, Paris.
- Povinelli, D.J., & O'Neill, D. (2000). Do chimpanzee use their gestures to instruct each others?. In Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (Eds.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford University Press, Oxford.
- Prinz, J.J. (2006a). Is the mind really modular?. In Stainton, R.J. (Ed.), *Contemporary Debates in Cognitive Science* (pp. 22-36). Blackwell, Oxford.
- Prinz, J.J. (2006b). Putting the brakes on Enactive Perception, *Psyche*, 12(1), pp. 1-19.
- Proulx, M.J., Stoerig, P., Ludowig, E., & Knoll, I. (2008). Seeing 'where' through the ears: effects of learning-by-doing and long-term sensory deprivation on localization based on image-to-sound substitution. *PLoS One*, 3(3), e1840.
- Purwins, H., Perfecto, H., Gratchen, M., Hazan, A., Marxer, R., & Serra, X. (2008). Computational models of music perception and cognition I: The perceptual and cognitive processing chain. *Physics of Life Reviews*, 5, pp. 151-168.
- Putnam, H. (1981). *Reason, Truth, and History*. Cambridge University Press, Cambridge.
- Ratcliffe, M. (2009). Existential Feeling and Psychopathology (and response to commentaries 'Belonging to the World through the Feeling Body'). *Philosophy, Psychiatry & Psychology*, 16(2), pp. 179-211.
- Reddy, V. (2008). *How Infants Know Mind*. Harvard University Press, Cambridge.
- Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLOS One*, 8(6): e65289.

- Renier, L., & De Volder, A.G. (2005). Cognitive and brain mechanisms in sensory substitution of vision: A contribution to the study of human perception. Review. *Journal of Integrative Neuroscience, Special Edition in Honor of Paul Bach-y-Rita*, pp. 489-503.
- Reybrouck, M. (2005a). Body, mind and music: Musical semantics between experiential cognition and cognitive economy. *Trans: Transcultural Music Review*, 9. Available at <http://www.sibetrans.com/trans/a180/body-mind-and-music-musical-semantics-between-experiential-cognition-and-cognitive-economy>.
- Reybrouck, M. (2005b). A biosemiotic and ecological approach to music cognition: Event perception between auditory listening and cognitive economy. *Axiomates, An International Journal in Ontology and Cognitive Systems*, 15(2), pp. 391-409.
- Reybrouck, M. (2012). Musical sense-making and the concept of affordance: An ecosemiotic and experiential approach. *Biosemiotics*, 5(3), pp. 391-409.
- Riebel, K. (2003). The mute sex revisited: vocal production and perception learning in female songbirds. *Advances in the Study of Behavior*, 33, pp. 49-86.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71, pp. 491-507.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, pp. 169-192.
- Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science*, 277, pp. 190-191.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, pp. 131-141.
- Rizzolatti, G., & Gentilucci, M. (1988). Motor and visual-motor functions of the premotor cortex. In Rakic, P., & Singer, W. (Eds.), *Neurobiology of Neocortex* (pp. 269–284). Wiley, New York.
- Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, 31, pp. 889-901.
- Rizzolatti, G., & Matelli, M. (2003). Two Different Streams form the Dorsal Visual System: Anatomy and Functions. *Experimental Brain Research*, 153, pp. 146-157.
- Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the Brain. How our Minds Share Actions and Emotions*. Oxford University Press, Oxford.
- Rockwell, T. (2005). *Neither Brain nor Ghost: A Nondualist Alternative to the Mind- Brain Identity Theory*. MIT Press, Cambridge.
- Rockwell, T. (2008). Dynamic empathy: A formulation for the simulation theory of mind reading. *Cognitive Systems Research*, 9, pp.52-63.
- Romdenh-Romluc, K. (2011). *Routledge philosophy guidebook to Merleau-Ponty and Phenomenology of perception*. Routledge.

- Rouget, G. (1968). L'ethnomusicologie. In *Encyclopedie Fasquelle de la musique*. Fasquelle. Paris.
- Rowell, L. (1992). *Music and Musical Thought in Early India*. Chicago Studies in Ethnomusicology, University of Chicago Press, Chicago.
- Rowlands, M. (1999). *The Body in Mind: Understanding Cognitive Processes*. Cambridge University Press, Cambridge.
- Rowlands, M. (2003). *Externalism: Putting Mind and World Back Together Again*. Acumen/McGill-Queen's University Press.
- Rowlands, M. (2006). *Body Language: Representation in Action*. MIT Press, Cambridge.
- Rowlands, M. (2010). *The New Science of the Mind*. MIT Press, Cambridge.
- Ruff, H.A. (1982). Role of manipulation in infants' responses to invariant properties of objects. *Developmental Psychology*, 18, pp. 682-691.
- Ruiz-Mirazo, K., & Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artificial Life*, 10(3), pp. 235-259.
- Russo, F.A., & Thompson, W.F. (2005a). An interval-size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception & Psychophysics*, 67, pp. 559-568.
- Russo, F.A., & Thompson, W.F. (2005b). The subjective size of melodic intervals over a two-octave range. *Psychonomic Bulletin & Review*, 12, pp. 1068-1075.
- Ruwet, N. (1975). Théorie et méthodes dans les études musicales. *Musique en Jeu*, 17, pp. 11-35.
- Ryle, G. (1949). *The Concept of Mind*. University of Chicago Press, Chicago.
- Sacks, O. (2007). *Musicophilia. Tales of music and the brain*. Picador. London.
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., & Tanaka, Y. (1997). The TINS Lecture. The parietal association cortex in depth perception and visual control of hand action. *Trends in Neuroscience*, 20, pp. 350-357.
- Sakata, H., Taira, M., Murata, A., & Mine, S. (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cerebral Cortex*, 5, pp. 429-438.
- Sartre, J.P. (1957). *Being and Nothingness, An Essay on Phenomenological Ontogeny*. Methuen & Co., New York. [*L'être et la néant*. Gallimard, Paris.1943].
- Schantz, R. (Ed.), (2004). *The Externalist Challenge*. De Gruyter, Boston.
- Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural asymmetry in musicians. *Science*, 267, pp. 699-701.
- Scholes, P.A. (1938). *Bird Songs*. Oxford University Press, Oxford.

- Schubert, E. (2009-2010). The fundamental function of music. *Musica Scientiae, special issue on music and evolution*, pp. 63-76.
- Schurmann, M., Caetano, G., Hlushchuk, Y., Jousmaki, V., & Hari, R. (2006). Touch activates human auditory cortex. *Neuroimage*, 30, pp. 1325–1331.
- Searle, J. (1980). Minds, Brains, and Programs. *The Behavioral and Brain Sciences*, 3, pp. 417-457.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press, Cambridge.
- Searle, J. (1992). *The Rediscovery of the Mind*. MIT Press, Cambridge.
- Searle, J. (1999). The Chinese Room. In Wilson, R.A., & Keil, F. (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press, Cambridge.
- Searle, J. (2000a). Consciousness. *Annual Review of Neuroscience*, 23, pp. 557-578.
- Searle, J. (2002). Why I Am Not a Property Dualist. *Journal of Consciousness Studies*, 9(12), pp. 57-64.
- Searle, J. (2004). *Mind. A brief introduction*. Oxford University Press, New York.
- Serravezza, A. (1996). *Musica e Scienza nell'Età del Positivismo*. Il Mulino, Bologna.
- Shapiro, L. (2010). Embodied cognition. In Margolis E., Samuels R., & Stich S. (Eds.), *Oxford handbook of philosophy and cognitive science*. Oxford University Press, Oxford.
- Shapiro, L. (2011). *Embodied Cognition*. Routledge, New York.
- Sheets-Johnstone, M. (2010). Thinking in Movement. Further Analyses and Validations. In Stewart, J., Gapenne, O., & Di Paolo, E.A. (Eds.), *Enaction: Toward a New Paradigm for Cognitive Science* (pp. 165-182). MIT Press, Cambridge.
- Sheets-Johnstone, M. (2012). Fundamental and inherently interrelated aspects of animation. In Foolen, A., Lüdtke, U., Racine, T., & Zlatev, J. (Eds.), *Moving Ourselves, Moving Others: Motion and emotion in intersubjectivity, consciousness and language* (pp. 27-56). J. Benjamins, Amsterdam.
- Sinigaglia, C. (2008a). Enactive understanding and motor intentionality. In Morganti F., Carassa A., & Riva G. (Eds.), *Enactive Intersubjectivity: a cognitive and social perspective on the study of interaction*. IOS press, Amsterdam, 2008.
- Sinigaglia, C. (2008b). Mirror neurons: This is the question. *Journal of Consciousness Studies*, 15(10-11), pp.70-92.
- Sinigaglia, C., & Sparaci, L. (2010). Emotions in action through the looking glass. *Journal of Analytical Psychology*, 55, pp. 3-29.

- Sittiprapaporn, W., Chindaduangratn, C., Tervaniemi, M., & Kotchabhakdi, N. (2003). Preattentive processing of lexical tone perception by the human brain as indexed by the mismatch negativity paradigm. In *Annals of the New York Academy of Science*, 999, pp. 199-203.
- Sloboda, J.A. (1985). *The Musical Mind*. Oxford University press, Oxford.
- Sloboda, J.A. (2005). *Exploring the Musical Mind: Cognition, Emotion, Ability, Function*. Oxford University Press, Oxford.
- Sloboda, J.A., Davidson, J.W., & Howe, M.J.A. (1994). Is everyone musical?. *Psychologist*, 7(8), pp. 349-355.
- Smith, A.D. (2003). *Husserl and the Cartesian Meditations*. Routledge, London.
- Smith, L.B. (2006). Movement Matters. The contributions of Esther Thelen. *Biological Theory*, 1(1), pp. 87-89.
- Solomon, R. (1976). *The passions: Emotions and the meaning of life*. Hackett Publishing Co., Indianapolis.
- Sommerville, J.A., Hildebrand, E.A., & Crane, C.C. (2008). Experience matters. The impact of doing versus watching on infants' subsequent perception of tool use events. *Developmental Psychology*, 44(5), pp. 1249-1256.
- Sommerville, J.A., & Woodward, A.L. (2005). Infants' sensitivity to casual features of means-end support relations in action and perception. *Infancy*, 8(2), pp. 119-145.
- Sommerville, J.A., Woodward, A.L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96, b1-b11.
- Spencer, H. (1857). The origin and function of music. *Frazer's Magazine*, October.
- Stanley, J., & Williamson, T. (2001). Knowing How. *Journal of Philosophy*, 98(8), pp. 411-444.
- Stefan, K., Cohen, L.G., Duque, L., Mazzocchio, R., Celnik, P., Sawaki, L., Ungerleider, L., & Classen, J. (2005). Formation of a motor memory by action observation. *Journal of Neuroscience*, 25(4), pp. 9339-9346.
- Steinbeis, N., Koelsch, S., & Sloboda, J.A. (2006). The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological and neural responses. *Journal of Cognitive Neuroscience*, 18, pp. 1380-1393.
- Stern, D.N. (1974). Mother and infant at play: the dyadic interaction involving facial, vocal and gaze behaviours. In Lewis, M., & Roseblum, L. *The effect of the infant on its Caregiver*. Wiley, New York.
- Stern, D.N. (1985). *The Interpersonal World of the Infant. A View from Psychoanalysis and Developmental Psychology*. Basic Books, New York.
- Stern, D.N. (2004). *The Present Moment in Psychotherapy and Everyday Life*. Norton & Company, New York.

- Stewart, J., Gapenne, O., & Di Paolo, E.A. (Eds.), (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press, Cambridge.
- Stich, S., & Nichols, S. (1992). Folk Psychology: Simulation or Tacit Theory. *Mind & Language*, 7(1), pp. 35-71.
- Suomi, S.J. (2008). Attachment in rhesus monkeys. In Cassidy, J., & Shaver, P.R. (Eds.), *Handbook of attachment: Theory, research, and clinical applications*, 2nd ed. (pp. 173-191). Guilford Press, New York.
- Tan, S.L., Pfordrescher, P., & Harrè, R. (2010). *Psychology of Music. From Sound to Significance*. Psychology Press, Hove - New York.
- Thelen, E. (1989). Self-organization in developmental processes: Can systems approaches work?. In Gunnar, M., & Thelen, E. (Eds.), *Systems and Development: The Minnesota Symposia on Child Psychology, Vol. 22* (pp. 17-171). Erlbaum, Hillsdale.
- Thelen, E. (1994). Three-month-old infants can learn task-specific patterns of interlimb coordination. *Psychological Science*, 5, pp. 280–285.
- Thelen, E. (2000). Grounded in the world: Developmental origins of the embodied mind. *Infancy* 1(1), pp. 3-28.
- Thelen E., Schoner, G., Scheier, C., & Smith, L.B. (2001). The dynamics of embodiment: A field theory of infant preservative reaching. *Behavioral and Brain Sciences*, 24, pp. 1-86.
- Thelen, E., & Smith, L.B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press, Cambridge.
- Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences*, 4(4), pp. 407-427.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard University Press, Cambridge and London.
- Thompson, E., & Stapleton, M. (2009). Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories. *Topoi*, 28(1), pp. 23-30.
- Thompson E., & Varela, F. (2001). Radical embodiment: neural dynamics and consciousness. *Trends in Cognitive Sciences*, 5, pp. 418-425.
- Ticini, L.F., Schütz-Bosbach, S., Weiss, C., Casile, A., & Waszak, F. (2011). *When Sounds Become Actions: Higher-order representation of newly learnt action sounds in the human motor system*. Poster presented at 11th International Conference on Cognitive Neuroscience (ICON XI). Mallorca, Spain.
- Titon, J.T., Koetting, J.T., McAllester, D.P., Reck, D.B., & Slobin, M. (Eds.), (1984). *An Introduction to the Music of the World's People*. Schirmer Books, New York.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press,

Cambridge, MA.

Tomasello, M. (2008). *Origins of Human Communication*. MIT Press, Cambridge.

Torrance, S. (2005). In Search of the Enactive. Introduction to the special issue on Enactive Experience. *Phenomenology and the Cognitive Science*, 4(4), pp. 357-368.

Trevarthen, C. (1997). *Empatia e Biologia*. Raffello Cortina, Milano.

Trevarthen, C. (1999). Musicality and the Intrinsic Motive Pulse: Evidence from Human Psychobiology and Infant Communication. *Musicae Scientiae, Special Issue on Rhythms, Musical Narrative, and the Origins of Human Communication*, pp. 155-215.

Trevarthen, C. (2012). Epilogue. In Foalen, A., Lüdtke, U., Racine, T., & Zlatev, J. (Eds.), *Moving Ourselves, Moving Others: Motion and emotion in Intersubjectivity, consciousness and language* (pp.453-461). J. Benjamins, Amsterdam.

Truitt, F.E., Clifton, C., Pollatsek, A., & Rayner, K. (1997). The perceptual-span and eye-span in sight-reading music. *Visual cognition*, 4, pp. 143-161.

Tye, M. (1995). *Ten Problems of Consciousness. A Representational Theory of the Phenomenal Mind*. MIT Press, Cambridge.

Uttal, W.R. (2003). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge University Press, Cambridge.

Van Gelder, T. (1995). What might cognition be, if not computation?. *Journal of Philosophy*, 92(7), pp. 345–381.

Van Elk, M., Van Schie, H.T., Hunnius, S., Vesper, C., & Bekkering, H. (2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage*, 43(4), pp. 808-814.

Varela, F. (1979). *Principles of Biological Autonomy*. Kluwer Academic, Boston.

Varela, F. (1987). Cognition without representations. In *Biology Forum*, 80(1), pp. 241-242.

Varela, F. (1991). Organism: A meshwork of Selfless Selves. In Tauber, A.I. (Ed.), *Organism and the origins of self*. Kluwer, Dordrecht.

Varela, F. (1992). Autopoiesis and a biology of intentionality. In McMullin, B., & Murphy, N. (Eds.), *Autopoiesis and Perception: A Workshop with ESPRIT BRA 3352 (ADDENDUM)*. Addendum to the print proceedings distributed during the workshop. Dublin City University, Dublin.

Varela, F. (1996). Neurophenomenology: A methodological remedy for the hard Problem. *Journal of Consciousness, Studies* 3(4), pp. 330-349.

Varela, F. (1997). The Naturalization of Phenomenology as the Transcendence of Nature: Searching for Generative Mutual Constraints. *Alter: Revue de Phénoménologie*, 5, pp. 355-385.

- Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind*. MIT Press, Cambridge, MA.
- Von Hofsten, C. (2007). Action in Development. *Developmental Science*, 10(1), pp. 54-60.
- Wallin, N. (2000). Prolegomenon for a hypothesis on music as expression of an evolutionary early homeostatic feedback-mechanism. A biomusicological proposal. *Human Evolution*, 15, (3-4), pp. 199-242.
- Wallin, N., Merker, B., & Brown, S. (Eds.), (2000). *The origins of music*. MIT Press, Cambridge, MA.
- Ward, D., & Stapleton, M. (2012). Es are good. Cognition as enacted, embodied, embedded, affective and extended. In Paglieri, F. (Ed.), *Consciousness in Interaction: The role of the natural and social context in shaping consciousness* (pp. 89-104). J. Benjamins, Amsterdam.
- Weber, A. (2001). Turning the inside out: natural forms as expression of intentionality. *Sign Systems Studies*, 29(1), pp. 153-168.
- Weber, A., & Varela, F. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1, pp. 97-125.
- Weiskopf, D. (2008). Patrolling the mind's boundaries. *Erkenntnis*, 68, pp. 265–276.
- Weng, G., Bhalla, U.S., & Iyengar, R. (1999). Complexity in biological signaling systems. *Science*, 284, pp. 92-96.
- Wheeler, M. (2005). *Reconstructing the Cognitive World*. MIT Press, Cambridge, MA.
- Wheeler, M. (2008). Cognition in Context: Phenomenology, Situated Robotics, and the Frame Problem. *International Journal of Philosophical Studies*, 16(3), pp. 323–349.
- Wheeler, M. (2010). In Defense of Extended Functionalism. In Menary, R. (Ed.), *The Extended Mind* (pp. 246-270). MIT Press, Cambridge.
- Wilson, R.A. (2004). *Boundaries of the Mind: The Individual in the Fragile Sciences: Cognition*. Cambridge University Press, New York.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell, Oxford.
- Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, pp. 1-34.
- Woodward, A.L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, 6(3), pp. 297-311
- Woodward, A.L., & Gerson, S. (in press). Mirroring and the development of actions understanding. *Philosophical Transactions of the Royal Society B*.
- Zahavi, D. (2003). *Husserl's Phenomenology*. Stanford University Press, Stanford.

Zatorre, R.J. (2001). Neural Specialization for Tonal Processing. In Zatorre, R.J., & Peretz, I. (Eds.), *The Biological Foundations of Music* (pp. 193-210). New York Academy of Science, New York.

Zatorre, R.J., Chen, L.J., Penhune, V.B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), pp. 547-558.

Zatorre, R.J., & Halpern, A.J. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron*, 47, pp. 9-12.

Zipoli Caiani, S. (2011). *The ecological meaning of embodiment*. Phenomenology & Mind. Available at <http://www.phenomenologylab.eu/public/uploads/2011/05/caiani.pdf>