This study summarises literature that describes the listener as an adaptive device by (a) identifying the processes involved in the perception of timbre within the paradigm of embodied cognition, and (b) formulating a solid ecological framework for the linking of the environment–human system to support further arguments on the influence of timbre in the music appreciation schemata. It offers an approach for the explanation of the dynamic relations emerging between the human body and specific timbral environments, presenting the idea that timbral features are relevant cues for the appreciation of music. The objective is to provide a theoretical framework for an empirical approach to timbre, aimed to characterise perceptual cues of (internal) musical representations and their interactions with diverse (external) environments.

**Interactionism, embodied cognition & timbre**

“...the human mind-brain is an organically embodied representational system, a system that enters into states that are systematically interrelated and that stand in intentional relations to the environment in which the human organism is embedded...” (Nussbaum, 2007, p. xi)

Music perception has been explained as an autopoietic system (i.e., a self-organised, self-structured and autonomous) in which dynamic interactions between the components of that system are the constituents of its existence more than the elements themselves (Zeleny, 1981; Chagas, 2005). Such a system can be used to model the cognition of musical phenomena by considering auditory schemata as a unit defined by its participative interaction with the environment (Kaipainen, 1996). The nature of these interactions is the central matter in this paper, which -as I will explain- are individuals’ intentional projections of their selves on the environment, meant to acquire an internal representation of specific fragments of the environment. The paradigm of embodied cognition (Leman, 2007) encapsulates these views. In a parallel approach taken from the computational theory of mind, music listeners can be considered as adaptive devices capable of organising their sensors and effectors to adapt themselves to the world and perform modifications on it (Reybrouck, 2005).

Through active interaction with the sonic environment, human beings develop a categorisation of musical phenomena (Dura, 2006), including finer variations of sounding qualities (Bregman, 1990); for instance, the ability of one month old babies to distinguish their mothers’ voice (Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978).

Currently these finer variations are compressed in the concept of timbre, which remains ill-defined due to several reasons (for a summary, see Donnadieu, 2007). One of these reasons is that the same word, timbre, has been used to describe quite different phenomena. I will pursue an argument that clarifies how this confusion in the semantic domain results in a limitation to the appreciation of music, and propose the term timbral environments to refer to a specific category of the general sounding phenomena that might influence the way we develop our musical preferences.

**Elusive concept of timbre**

**The origin of the problem**

History of knowledge narrates how existing things acquire new meanings from time to time. For example, the conceptual shift provoked by the invention of perspective (in
the visual arts) or polyphony (in music); the same elements organised in a different way leading to a whole new idea about depth in the visual and auditory domain. After the exposure and understanding of these inventions, it gestates a change in our minds, and as a consequence, changes the way we use our bodies to understand new characteristics about the things we already know. The way our body “makes sense” of its environment reflects our consciousness (Popper & Eccles, 1984). The concept of timbre is experiencing a similar history according to the claims of Fales (2005), who traced back to the Age of Enlightenment the modern acceptation for the term timbre. Through a historical review of the concept, Fales argues that one of the first westerners that became aware of timbre in the sense we use it today was Jean Phillipe Rameau (1863-1764), who proposed that the difference between “hearing and listening” would lead to an (un)effective understanding of the corpse sonore. This distinction between two tasks (hearing and listening) involving the auditory schemata, suggests that he intuited the need for a conscious effort to grasp particular qualities of sound.

“... One could say first that the phenomenon of the corps sonore is the first marvel that Nature has yet submitted to our reason. To believe, in effect, that one hears only one sound where one distinguishes three different ones; and to take it still for unique, though one knows it triple, whom could one persuade of this truth, if one couldn’t make him touch it with the finger and eye? (Jacobi, V: 351) ...

” (Fales, 2005, p. 9)

Nevertheless, scholars of his epoch failed to understand the idea and were also unable to give an empirical explanation for the phenomenon. This state remained for a century, until Herman von Helmholtz (1821-1894) started to relate the perceptual attributes of sound with its physical properties (Helmholtz, 1954). After him, the music psychologist Carl Seashore (1866-1949) proposed timbre as the most important and complex aspect of tone, over pitch, loudness and duration (Seashore, 1967). A new tradition on timbre research started within the paradigm of cognitive structuralism (based on similarity tests), which lead to an understanding of the multidimensional nature of timbre (Grey, 1977). This approach (that we will refer to as the classical approach), has been thoroughly used during the last 40 years and has also been used in the development of computational models with the goal to find representations of timbre that are “isomorphic with human perception” (Terasawa, Slaney, & Berger, 2006). But cognitive structuralism has a major drawback that it is not capable of capturing the dynamics of the perceptual system (Leman, 1995), so recent efforts are aimed at understanding timbre in dynamic (Hajda, 2007) and complex (Donnadieu, 2007) settings.

A matter of representation

According to the representational theory of mind (Nussbaum, 2007), individuals develop, throughout their life-span, a consciousness about themselves and the complexities of the environment, by making use of their bodies. A body is provided with perceptual capabilities that allow it to realise internal representations of the environment. This interaction with the environment (which include objects and other individuals) is the starting point of an interplay between action and perception.

The capability of an individual to afford (Gibson, 1986) an interaction depends on how aware she or he is about the contents of the environment. This awareness is constructed by the linking of three different realities: the first comprises the physical entities existing in the environment, the second is the mental state or states of consciousness, which involves the experience of thinking and perceiving, and a third reality composed of abstractions, and ideas or intuitions in the old platonic sense (Popper & Eccles, 1984). The same three realities can be referred to with alternative terminology that has been used in musicological research. First, there is music as a morphology existing as a physical entity, followed by, the so called isomorphisms (Terasawa et al., 2006; Dura, 2006), which are internalised morphologies, and finally, the isomorphisms of second order (Leman, 1995), which are abstractions that control the emergence and functionality of perception. The linkage of these three realities forms a loop of interactions, made of actions and perceptions that bring closure to the system in the autopoietic sense. But there is still the question about the nature of such
isomorphisms. The paradigm of embodied cognition sheds light on a possible answer that isomorphisms are embodied. In other words, if internalised representations of external objects use the body as a medium, they are most likely to be anthropomorphic projections (Godøy, 2006). Identity is what remains after the individual interacts with the environment and reorganises itself, preserving its unity, structure, and autonomy as a closed system. If this holds true, I propose that perceptual schemata is an (ontological) expression of the self, which can or cannot afford the retrieval of specific information contained in the environment. The affordances can be viewed as learnt methods used to apprehend specific characteristics of the environment.

Timbre can be explained in these three hypothetical worlds. In the first there are the physical attributes of sound, in the second the evidence that individuals can perceive, differentiate and experience emotions with diverse timbres (Eerola, Alluri, & Ferrer, 2008), and the third comprising all the possible descriptions, hypothesis and theories about it. In this third world it is safe (from an epistemological point of view) to speculate about the existence of an unembodied timbre. If we, as a culture, lack a theory of timbre it is because the specific embodiment remains unexplained. Hence suggesting that part of the problem has an origin in the way we conceive our projection onto timbre. For example, as timbre is a quality of sound, we could think that the auditory system is the main sense involved in the acquisition of an isomorphic representation. Nevertheless, by analysing two of the most used verbal descriptors of timbre such as brightness and roughness, it is evident that individuals’ embodiment of timbre is mostly visual and tactile.

**Beyond macrotimbre: timbral environments**

Macrotimbre (Sandell, 1998) is a term that bisects the traditional concept of timbre by referring to the set of qualities that remain invariant across several pitches at different loudness levels. The definition contrasts slightly with the old ANSI (American National Standards Institute, 1973) view, which alludes to the characteristic of sound that remains after loudness and pitch, and makes us distinguish one source from another. The difference is how broad the scope is in terms of time, and also that it is a better approach in terms of how timbre is internally represented. While the traditional approach is concerned with milliseconds and monophonically isolated sounds, macrotimbre seems concerned with a summary of characteristics that lets us differentiate one source from another, even if they are performed at different loudness levels and pitches. Thus macrotimbre is concerned with a greater temporal span, where memory plays a central role in the form of perceptual constancy (Sandell & Chronopoulos, 1997). For example, the spectral centroid (which has a high correlation with perceived brightness) of one second of music can be computed by analysing consecutive frames of 25 milliseconds of music. This computation will give a list of 40 numbers. We can observe the range and variance of these numbers and conclude that the mean is a meaningful perceptual measure if the standard deviation is not greater than the inter-quartile range, and by increasing the length of the analysis frame to 100 milliseconds, we would probably reach to a similar result. Nevertheless it is expected that if we apply the same operation to a more than one second, let’s say 30 seconds of music, the numerical result will be very different. Furthermore, if the music contains contrasting instrumentation during that time, our measurement might not be very representative of what we actually perceive, because “. . . listeners do not perceive the acoustical environment in terms of ‘phenomenological descriptions’ but as ‘ecological events’. . . ” (Reybrouck, 2005). The concept of macrotimbre seems to allude to an ecological perspective by including in a single category a wide variety of timbral events, suggesting that the schemata that controls the perception of timbre makes us able to understand that across pitches, loudness levels, attack types and articulations (i.e., sul ponticello, muted, staccato), the source remains the same.

The addition of “macro” to the word timbre is also useful to make a distinction between the classical studies (i.e., short excerpts of isolated sounds, timbral events happening
below the temporal span and complexity level described by macrotimbre) and a novel approach proposed here, aimed to characterise even longer temporal and more complex timbral events such as soundscapes (Schafer, 1977). The term is constructed by substituting the prefix of the word ‘landscape’ with ‘sound’ to transpose the concept from the visual to the auditory domain, used for example, to analyse the sonic environment of two geographical locations by contrasting its salient characteristics (Ge & Hokao, 2005).

The same principle of perceptual constancy used to define macrotimbre could be used to discriminate among prototypical soundscapes (i.e., predominant sounds surrounding a house in a city in contrast those surrounding a house near the sea). What I propose is to merge the two concepts, macrotimbre and soundscapes, to (a) take advantage of the slightly difference in perspective with respect to the classical approach to timbre, and (b) apply that body of knowledge (i.e., methodologies and experiences) to different temporal domains and complexity levels. The perceptual validity for musicological research would rest in our evident ability to discriminate not only between sources, but also between prototypical mixtures of sources (i.e., the overall sound of a rock ensemble versus the overall sound of a big-band orchestra).

If these prototypical descriptions of timbre can be characterised and systematically differentiated, we could call them timbral environments for two main reasons. First, to make a distinction between the classical approach to timbre and timbral spaces (Grey, 1977; Wessel, 1979; Krumhansl & Iverson, 1992; McAdams, Winsberg, Donnadieu, Soete, & Krimphoff, 1995; Terasawa et al., 2006), and second, to stress the ecological approach, linked with the paradigm of embodied cognition (Godøy, 2006; Leman, 2007) and evolutionary epistemology (Reybrouck, 2005).

Music appreciation as an example of interactionism

Interactionism is a view that mental events and physical events causally influence each other (Tye, 2008a), and proposes a solution for the mind-body problem (Grey, 1977). It has been pointed out that the word interaction evokes a neutral relation, and that for music interaction the term participation (Kaipainen, 1996) should be used instead, because with conscious (participative) interaction with the environment, we re-create the environment and promote changes in our internal organisation. Discovery of timbre provoked a shift in the way we listen to music because it opened a new aesthetic category to be aware of. This is more obvious in cross-cultural settings, for instance, when a westerner tries to describe the characteristics of a musical expression where the building blocks (i.e., form and rhythm) are not the main aesthetic resource (Fales, 2002). The understanding of novel qualities or expressions opens new ways of appreciation of the products already known. Music appreciation is a very good example of participative interaction with the environment. Musical activities play a central role in the construction of the self (DeNora, 1999), which is expressed through musical preferences. Connections between the personalities of individuals and their musical preferences for specific musical genres have been reported by Rentfrow and Gosling (2003). Research on musical preferences evoke social dynamics, where consensus about the set of beliefs and stereotypes (Gregory, 1999; Rentfrow & Gosling, 2007) with respect to the music are another relevant force in the construction of the self by means of social identity (Bakagiannis & Tarrant, 2006) and interpersonal perception (Rentfrow & Gosling, 2006).

Conclusions

The word timbre seems to fit various acceptations, which is an important part of the problem to define it. Despite that the categorisation of specific qualities of music as timbre started almost two hundred years ago, it seems nowadays that the term is only a weak reference to a physical attribute of sound. In order to reach an effective taxonomy for this probably transitional term, it would be advisable to learn which parts of our body (or sensory systems) are most involved when we attempt to grasp an image of it. Perception of timbre involves a complex
system of interactions between a listener and its environment. The individual has to intentionally focus on a particular sonic event, keeping in mind that she or he is going to acquire a mental representation of such an event. It would also be desirable for listeners to know that the objects in the environment cannot be represented as static morphologies, but as fluctuating anthropomorphic projections of the self.

A more inclusive perspective aimed to observe not only the components of the system (i.e., perception, acoustic features, society and culture) but also the interactions between them, would contribute to a better understanding of timbre and its role in music. Ultimately this effort is aimed to promote the awareness of timbre as an environmental issue that possibly influences our daily decisions about what to listen. For this reason the term timbral environments becomes necessary.

This theoretical foundation will be used in future research targeted at investigating the existing forces that shape the emergence and functionality of a perceptual schemata of timbre, and in better understanding the relations between timbre, musical preference and personality.

References


Pylkkänen, P., Pylkkö, A. Hakamäki (Eds.), Brain, Mind and Physics (pp. 266-277). Netherlands: IOS Press.


\(^{1}\) in Observations sur notre instinct pour la musique (1754).