Timbral Environments: An Ecological Approach to the Cognition of Timbre

RAFAEL FERRER
Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, University of Jyväskylä

ABSTRACT: This study formulates an ecological framework that links the environment and human systems, to support further arguments on the influence of timbre in the music appreciation schemata. At the core of the framework is the notion of *timbral environments*, which is introduced as an epistemological foundation to characterize perceptual cues of internalized representations of music, and to explore how these are expressed in the dynamics of diverse external environments. The proposed notion merges the concepts of *macrotimbre* (Sandell, 1998) and *soundscape* (Schafer, 1977) to distinguish between the formulated framework and traditional approaches to timbre, which are mainly concerned with short-term temporal auditory events. The notion of timbral environments enables the focus of timbre research to be shifted from isolated events to socially relevant sounding objects, hence facilitating the identification of connections between semantic descriptors and the physical properties of sounds. [1]

Submitted 2009 January 26; accepted 2011 April 7.

KEYWORDS: *timbral environments, macrotimbre, soundscape*

THE word timbre is arbitrarily used to refer to multiple qualities of sound, which is an important part of the problem to define it. Despite the fact that categorisation of specific qualities of music, in terms of timbre, began almost two hundred years ago, psychologically inspired descriptions of timbre and existent research have not attained sufficient strength for generating a general theory of timbre (Huron, 2001).

Introductory paragraphs of this paper present selected ideas conceiving entities in a continuum of interactions (taken from different fields e.g., biology, cybernetics, music cognition) as epistemological basis to explore the cognition of timbre from an ecological perspective. Then, a review of major concerns with respect to timbre, such as history and categorization issues, precedes a description of how the ideas presented in the introduction could be used in the particular case of timbre. Here a distinction is made between the classical empirical approach to timbre that has been concerned with short, isolated sound events and a concept of global timbre, which covers a longer time-span, such as macrotimbre and soundscape (Sandell, 1998; Schafer, 1977). This exposition ends with the introduction of a new term: *timbral environments*, which merges the concepts of macrotimbre and soundscape to shift the focus away from the traditional approach to timbre. The closing part of the paper includes pragmatisms concerning empirical possibilities for the introduced term.

AN ECOLOGICAL PERSPECTIVE

Music perception and cognition could be modelled as an *autopoietic* (i.e., self-organised, self-structured and autonomous) system if music is considered a psychological construct and the sonic environment a continuum of information in which the individual exists. Such a framework serves to focus on the dynamic interactions between the components of that system rather than in the components themselves. Autopoietic theory has been used before to explain the cognition of polyphonic music (Chagas, 2005). In this paper, autopoiesis is used as a framework to present a theoretical model of timbre cognition

by considering an individual and her musical schemata (i.e., mental structures reflecting the relations between the perceived objects, see Leman, 1995) as an autonomous unit defined by its participative *interaction* with the environment.

An individual must intentionally project herself into the environment in order to internally represent specific fragments of the environment. According to this view, the human body is an autonomous unit interacting with the environment through its *sensors* and *effectors* (Godøy, 2006), which is in concordance with the paradigm of *embodied cognition* (Leman, 2007). Analogically, music listeners can be considered as *adaptive devices* in that they organize their sensors and effectors to adapt themselves to the world, while simultaneously modifying it (Reybrouck, 2005). According to the representational theory of mind (Nussbaum, 2007), individuals use their bodies, throughout their lives, to develop a consciousness about themselves and the complexities of the surrounding environment. A body has perceptual capabilities that allow it to internally represent the outside environment. This interaction with the environment, which includes both objects and other individuals, is the point at which the interplay between action and perception begins.

Interactionism maintains that mental and physical events "…causally influence each other" (Tye, 2008), and proposes a solution to the mind-body problem (Popper & Eccles, 1984) if we agree to extend the boundaries of the mind as an object of study beyond the individual into society in terms of *augmented* (Chi, 2009) *or distributed* (Barnier, Sutton, Harris, & Wilson, 2008) *social cognition*. However, the word *interaction* evokes a neutral relation, and for music perception, a term reflecting a more active role should be used. For that purpose, Kaipainen (1996) proposes the use of *participation*, arguing that through a conscious and participative interaction with the environment, we generate fluctuations in the system and at the same time promote changes in our internal structures (e.g., *neural plasticity)*. The dynamic interplay where structures and their internal organisations are mutually deformed is termed *structural coupling* (Maturana, 2002), (see Figure 1). In addition to the body acting as a mechanical medium, language - in a multimodal sense - is considered by Maturana (1988) as the subject matter of reciprocal coupling, hence, a social dimension is implied. In the views presented in this paper, the social domain is the fabric composing the environment, as necessary as unavoidable in a musical context.

By participating with the environment human beings, develop a categorization of musical phenomena (Dura, 2006), including finer variations of sounding qualities (Bregman, 1990). An example is the ability of a one month old baby to distinguish its mother's voice (Mehler, Bertoncini, Barriere, & Jassik-Gerschenfeld, 1978), which later in life develops into the ability to discriminate subtle timbral variations such as phonemes (Hinton, Nichols, & Ohala, 1995; McMullen & Saffran, 2004; Patel & Iversen, 2003). Clarke (2005) identifies this process of recurrent categorization as *perceptual learning*, which in an ecological context and from the perspective of *information foraging* (Held & Cress, 2009) theory, might be an expression of an externalized rather than internalized form of knowledge. At the core of the participation with the environment could be the *transactional memory*, i.e., people not having to know everything if they can use other people's knowledge (Hesse, 2009); what escapes one individual's perceptual capabilities, is captured by another, thus allowing the exchange of memory cues used to make transactions (Chi, 2009). The purpose of these transactions could be the *adaptation* (Maturana, 2002) of the individual to a given environment.

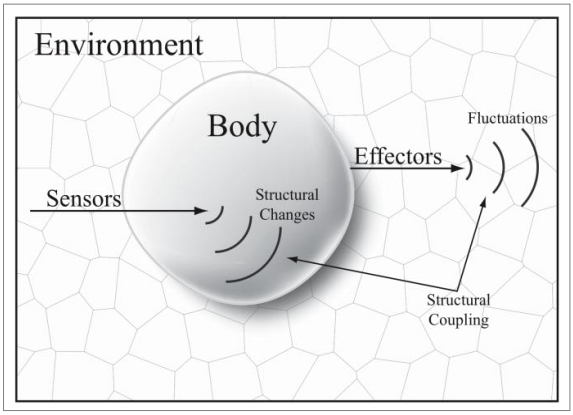

Fig 1. Diagram of the embodied-participative model

This model might well aid in the exploration of a wide variety of musical phenomena. For instance, individuals' musical preferences (Delsing, ter Bogt, Engels, & Meeus, 2008), developed in the dynamics of a social context (Gregory, 1999; Rentfrow & Gosling, 2007), contribute in the construction of the self (DeNora, 1999), by means of social identity (Bakagiannis & Tarrant, 2006) or interpersonal perception (Rentfrow & Gosling, 2006). In this example, the social dimension corresponds to the environment. The dynamic organisation, as the social identity and interpersonal perception, are fluctuations in the environment, and the self could be considered as the embodied entity experiencing constant changes in its structure.

I will now turn to the more difficult concept, namely timbre, and provide an exposition into the history of the concept before applying the interactional/embodied/autopoietic view to it. Finally, an attempt will be made to advance such embodied framework for timbre studies and provide the reader with lists of benefits associated with this perspective applied to timbre.

TIMBRE AS A CATEGORY OF SOUND

Many kinds of sound phenomena fall into the category of *timbre*, which remains ill defined for several reasons (cf. Donnadieu, 2007). One of the most problematic is that timbre is an umbrella term that has been used to describe many categories of sound, which have since been differentiated by empirical methods. To overcome this polysemic conflict I propose to use the term *timbral environments* to refer to a very specific category of the general sounding phenomena. Timbral environments are concerned with reminiscences of music pieces that roughly share the same range of acoustic features but have the property of being meaningfully grouped by noticeable distances in the perceptual space. An example of this meaningful grouping is the taxonomy of musical genres (Pachet & Cazaly, 2000), however, the term timbral environments is created with the purpose of generating alternative taxonomies that extend beyond musical genre into different layers of the music ontology. Ideally, timbral environments will aid the investigation of the emergence and functionality of musical schemata, particularly in relation to the music preferences. An elaboration of past ideas relating to timbre, the current proposal and connections with the general framework of the paper are presented next.

Highlights in the History of Timbre

Words tend to acquire new meanings from time to time as the concept they refer becomes more complex. For example, the conceptual shift provoked by the invention of perspective (in the visual arts) or polyphony (in music); the same elements in a given representational space but organised in a different way led to a whole new idea about depth in the visual and auditory domain respectively. After being exposed to these inventions a change in our minds emerges, and consequently, the way we use our bodies to perceive new characteristics about the things we already know is also transformed.

The invention of timbre as a novel category of sound wrought similar changes. According to Fales (2005), the modern meaning for the term timbre can be traced back to the Age of Enlightenment. Through a historical review of the concept, Fales argues that one of the first Westerners that became aware of timbre, in the sense that we use it today, was Jean Phillipe Rameau (1683-1764). He proposed that the difference between "hearing and listening"[2] posed problems for an effective understanding of the *corps sonore*. The distinction between the two tasks of hearing and listening, suggests that he intuitively recognised the need to make a conscious effort to grasp the particular qualities of sound he was capable of perceiving. For him, the *corps sonore* was an agglomerate of sound attributes that needed to be dissected - and perhaps this is the closest analogy to our present understanding of timbre. Nevertheless, scholars of his epoch failed to grasp the idea and nobody was able to give an empirical explanation for the phenomenon. This remained the state of affairs for a century, until Hermann 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 in timbre research started with the advent of *cognitive structuralism*, which was largely based on similarity tests (i.e., assessment of subjective similarity relations between audible stimuli). It led to an understanding of the multidimensional nature of timbre (Grey, 1977; McAdams, Winsberg, Donnadieu, Soete, & Krimphoff, 1995; Wessel, 1979), and has remained the classical approach in timbre research for the last 40 years. It has also been used in the development of computational models, with the goal being to find representations of timbre that are "*isomorphic* with human perception" (Terasawa, Slaney, & Berger, 2006). Nevertheless, cognitive structuralism has a major drawback, as explained by Leman (1995): it cannot capture the dynamics of the perceptual system. Therefore, recent efforts have been aimed at understanding timbre in both dynamic (Hajda, 2007) and complex settings (Donnadieu, 2007).

An embodied view of timbre embedded in an autopoietic framework might be able to provide a solution to the explanatory deficiency of cognitive structuralism because it extends the focus of research to an ecological context, where the dynamic relations between the environment and the elements coexisting within are at the core of the model.

Embodiment of Timbre

Direct and inferential theories of perception are often presented as antagonists (Chemero, 2003); the difference between the two views is related with the localization of *meaning*, whether it is in the environment (direct) or in the individuals (inferential). In this work, these views are rather presented as complementary, but in order to do that we need a common ground, which can be the theory of affordances (Gibson, 1986).

The capability of an individual to *afford* a participative interaction with the environment depends on how aware she is about the contents of that environment. This awareness is constructed by linking three different kinds of reality: the first comprises physical entities that exist in the environment; the second kind is the mental state or states of consciousness associated with thinking and perceiving; and a third reality is composed of abstractions and ideas, or intuitions in the old Platonic sense (Popper & Eccles, 1984). The same three realities also translate into musicological research, albeit with alternative terminology. The first sees music as a morphology that consists of physical entities; the second is an internal and isomorphic representation of those morphologies (Dura, 2006; Terasawa et al., 2006), and the third consists of *isomorphisms of second order*, which are abstractions that control the emergence and functionality of perception (Leman, 1995). These three realities are linked by loops of action-perception (i.e., structural coupling) that bring closure to the system in the autopoietic sense. However there is still a question regarding the nature of the isomorphisms. The paradigm of embodied cognition sheds light on a possible explanation, which maintains that such isomorphisms are encoded by, and in the body. 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). It is probable that these projections reflect a unique part of the individual's self, as well as a fingerprint of the cultural environment embedded on her self as a result of her development. Such a reflection can be identified as identity, at an individual and at a social level. Identity is what remains after the individual participates with the environment and reorganises itself, preserving its unity, structure, and autonomy as a closed system. If this holds true, perceptual schemata are an ontological expression of the adaptive self, which potentially afford any information contained in the environment. Affordances can be viewed as learning methods, developed to apprehend specific characteristics of the environment, distributed among individuals in the environment and possibly taking the form of transactional memories.

Timbre can be explained in these three hypothetical worlds. For instance, the first where the physical attributes of sound exist (e.g., acoustic descriptors such as Mel Frequency Cepstral Coefficients and other statistical descriptions of the sound spectrum), the second, where individuals perceive (e.g., as in the studies of John M. Grey, David Wessel and Stephen McAdams among others), and the third comprising all the possible descriptions, hypotheses 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; created and reserved only to be empirically tested. It is also on this third world where internalized representations of sound, imagery, and words to describe the sound experience, converge (e.g., internalized experiences uttered as onomatopoeias).

The specific embodiment of timbre remains unexplained due the particularities of the phenomena such as its multidimensionality, and perhaps because of a failure in the way we conceive the abstraction of

our bodies in projection onto different categories of sound. For example, as timbre is a quality of sound, we could assert that the auditory system is the main sensory apparatus involved in the acquisition of an isomorphic representation. Nevertheless that can only be part of the truth, because by regarding two of the most used verbal descriptors of timbre such as *colourful - colourless* and *dull - sharp* (Sethares, 1999), it is evident that individuals' embodiment of timbre is mostly visual and tactile. Efforts in the vein of this example, where free verbal descriptions are paired with acoustic descriptors have shed light on this issue (Sarkar, Lan, Diaz, & Vercoe, 2009), nevertheless, the vocabulary has never been filtered and processed to obtain an anthropomorphic ontology. An example of this is provided at the end of the following section.

Defining Timbral Environments

The idea of timbral environments can be regarded as an extension of the work of Sandell (1998), who proposed the term *Macrotimbre*. This term challenges the traditional concept of timbre by referring to the set of qualities that remain invariant across several pitches at different loudness levels. Sandell's notion differs from the classical ANSI definition (American National Standards Institute, 1973), which also alludes to the characteristic of sound that allow us distinguish one source from another, but conceptually separated from loudness and pitch. The difference—apart from the inclusion/exclusion of obvious dimensions such as loudness and pitch—resides in how broad the scope is in terms of time. The ANSI view—that has permeated most of the corpus of research—presents a fragmented auditory object of a short duration in the order of seconds and milliseconds. In contrast, macrotimbre refers to events beyond such time restrictions, in the order of minutes or hours. Therefore it provides a better approach in terms of how timbre is internally represented in a holistic way, closer to the popular expression "it sounds like…." While the classical empirical approach is mostly concerned with short and monophonic isolated sounds (Grey, 1977; Krumhansl & Iverson, 1992; McAdams et al., 1995; Terasawa et al., 2006; Wessel, 1979), Sandell's view is concerned with a summary of characteristics that makes us able to differentiate one source from the other even if they are performed at different loudness levels and pitches. The key to temporal span considerations resides in memory, which plays a central role in the form of *perceptual constancy* (Sandell & Chronopoulos, 1997).

The whole idea of microtimbre suggests that the schemata controling the perception of timbre enables us to understand that across pitches, loudness levels, attack types and articulations (i.e., *sul ponticello*, muted, *staccato*), the sound source remains the same. Such an interpretation has a high ecological quality and validity, since "…listeners do not perceive the acoustical environment in terms of 'phenomenological descriptions' but as 'ecological events'…" (Reybrouck, 2005, p. 234); reminiscences of musical events represent a global impression of past events. For instance, it can be argued that musical genres are characterized by their prototypical macrotimbres. Furthermore, the addition of the prefix "macro" to the word timbre is useful to make an epistemological distinction between the classical studies investigating the perceptual correlates of short excerpts of isolated sounds and further explorations that extend beyond such conceptual and methodological constraints.

The novel approach proposed here is aimed at characterizing longer temporal (i.e., beyond the lifespan of an individual's reminiscence about auditory events and complex timbral events such as *soundscapes* (Schafer, 1977). Schafer's term was constructed by substituting the prefix of the word 'landscape' with 'sound' to transpose the concept from the visual to the auditory domain. The term has inspired a host of publications within the field of acoustic ecology, where for example the sonic environment of two geographical locations is analyzed by contrasting their salient acoustic characteristics (Ge & Hokao, 2005).

What I propose is to merge the two notions of macrotimbre and soundscapes into *timbral environments*. To take advantage of the different perspective that macrotimbre affords with respect to the classical interpretations of timbre, and apply the methodologies and experiences that have been developed over the past forty years of soundscape research into different temporal domains and levels of complexity. Within the notion of timbral environments, the principle of perceptual constancy supporting macrotimbre, could be used to discriminate among prototypical soundscapes (e.g., predominant sounds surrounding a house in a city, in contrast to the predominant sounds surrounding a house near the sea, or distinguishing the differences between salient perceptual characteristics of musical genres). The perceptual validity for

musicological research would be to provide a means to better explain 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). Listeners are able to recognize that such sources belong to well-defined categories of sound despite the huge numerical variance in objectively measured timbral descriptors. If these categories can be empirically characterized and systematically differentiated, we could call them timbral environments. This would make them distinct from the classical approach to timbre and *timbral spaces*, and stress the ecological approach (Godøy, 2006; Leman, 2007; Reybrouck, 2005). For a visual comparison between the three different definitions, see Figure 2. As illustrated in the figure, the role of memory is one of the crucial differences between the definitions, since the first one (classical timbre) does not have any direct relation, the second one (macrotimbre) is based on individuals' memory capacities for handling perceptual constancy and recognition and the third definition, timbral environments, is based on social, collective memory.

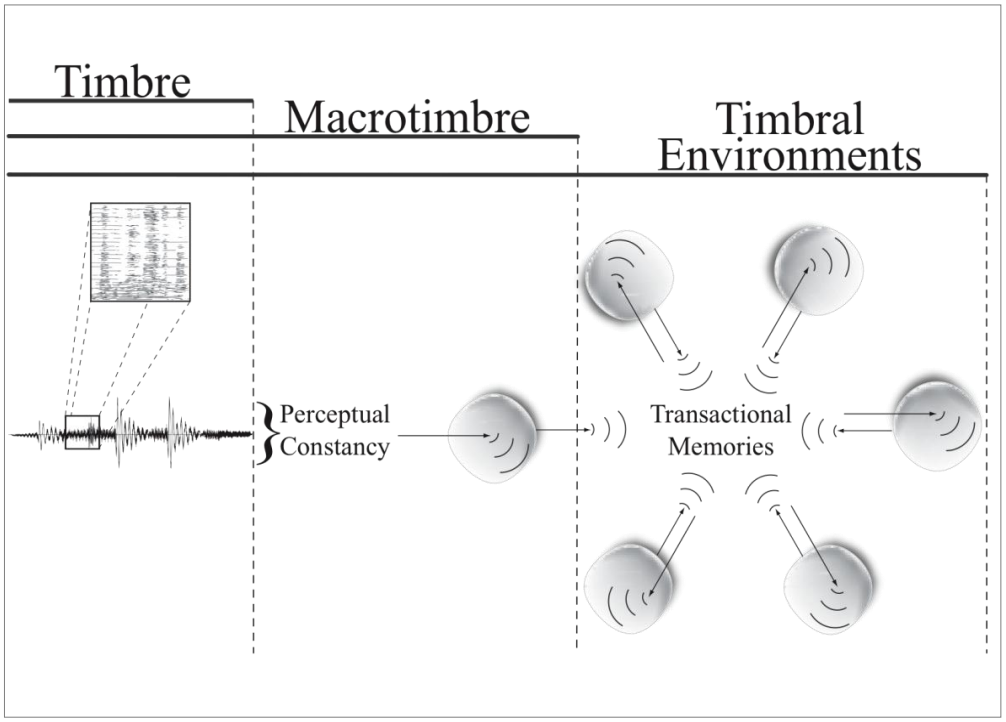

Fig 2. Comparison of the focus of research between the different definitions of timbre.

The notion of timbral environments could be used to represent a convergence between semantic and acoustic spaces empirically, for instance by filtering verbal descriptions of music to an anthropomorphic ontology and correlating such structure with the acoustic descriptors of the described music. Take for example the emotional attributes of a piece of music. A song is said to be sad or happy, although it can be argued that there is no such a thing contained in the song, or that there is no consensus about it. What is certain is the interpretation of the listener, or in the context of this paper, sadness or happiness are projections of the listener's self, who judges the piece and attaches a particular label to it. Such labelling does exist in the everyday chain of consumption-distribution of music, and it is called tagging (Lamere, 2008); users of social media (e.g., Last.fm) tag their music according to their own projections in the most varied semantic categories. The corpus of verbal descriptions can be classified and filtered according to categories related to the body or attributes inherently human such as the emotions [3]. Such an analysis has been carried out by Laurier, Sordo, Serrà and Herrera (2009), who derived a mood space by filtering 6,814,068 tags attributed to 575,149 pieces of music. Moreover, such a semantic structure can be connected with the music to establish a correlation between the semantic and acoustic domains (Ferrer & Eerola, 2010). The set of qualities describing the connections between the semantic and the acoustic domain could be considered as timbral environments, thus allowing us to refer to the set of acoustic descriptors that define, for example, sadness or happiness, or any other category related with the perception and cognition of auditory events.

CONCLUSIONS

The term *timbral environments*, is presented as a theoretical solution to further investigate the perceived general sounding quality of music in an ecologically valid fashion. However, it remains to be validitated by empirical means. Therefore, this theoretical foundation will be used in future research that is targeted at investigating the existing forces that shape the emergence and functionality of perceptual schemata of timbre.

Perception of timbre involves a complex system of interactions between listeners and their environment. Therefore, in order to extend our knowledge to reach an effective ontology of musical timbre, it would be advisable to learn about which parts of our body (or sensory systems) are involved when we attempt to grasp an internal representation of it. Objects in the environment cannot be represented as static morphologies, but as fluctuating anthropomorphic projections of the self. The ideas presented in this paper represent an effort to promote the awareness of timbre as an environmental issue that, for example, may have a possible influence on our daily decisions about what to listen to.

If the notion of timbral environments survives the process of empirical validation, it will have various implications for future studies by extending the focus of research beyond the traditional views displayed for instance in monophonic and polyphonic timbre research, or by contributing with empirical evidence to the definition of timbre as an aeshetic resource in Western and non-Western traditions. It will also be useful to derive the sounding objects and their features from conceptual units and sources that are meaningful and common for the listeners (everyday sounds, speech, typical instrument combinations), allowing for a better connection between semantic descriptors and acoustic features. Timbral environments could be studied using an array of behavioural methods (similarity ratings, priming tasks, semantic rating scales) as has been done in the past, but perhaps using richer sets of sound categories to keep the comparisons at a meaningful level. This will result in the creation of sets of stimuli in a bottom-up fashion, in which listeners' natural sound categories (e.g., musical genres, associations of sounds) are taken as the meaningful units.

NOTES

[1] Part of this work was presented in the SysMus08 conference in Graz, Austria, and was selected for publication in the British Postgraduate Musicology on-line under the title of "Embodied Cognition Applied to Timbre and Musical Appreciation: Theoretical Foundation."

[2] In *Observations sur notre instinct pour la musique* (1754).

[3] Note that in the interpretation made here, emotions elicited by music are considered as anthropomorphic attributions of music with the purpose of extending Godøy's (2006) term, anthropomorphic projection, beyond physical appearance.

REFERENCES

American National Standards Institute. (1973). Psychoacoustical terminology. In *S3.20-1973.* New York: American National Standards Institute.

Bakagiannis, S., & Tarrant, M. (2006). Can music bring people together? Effects of shared musical preference on intergroup bias in adolescence. *Scandinavian Journal of Psychology*, Vol. 47, pp. 129-136.

Barnier, A., Sutton, J., Harris, C., & Wilson, R. (2008). A conceptual and empirical framework for the social distribution of cognition: the case of memory. *Cognitive Systems Research*, Vol. 9, No. 1-2, pp. 33-51.

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge: MIT Press.

Chagas, P. (2005). Polyphony and embodiment: a critical approach to the theory of autopoiesis. *TRANS-Transcultural Music Review*, Vol. 9, Article 15. Retrieved September 6, 2010, from http://www.sibetrans.com/trans/trans9/chagas.htm

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, Vol. 15, No. 2, pp. 181-195.

Chi, E. (2009). Augmented social cognition: using social web technology to enhance the ability of groups to remember, think, and reason. In: C. Binnig & B. Dageville (Eds.), *Proceedings of the 35th SIGMOD International Conference on Management of Data*, Providence, Rhode Island, USA, pp. 973-984.

Clarke, E. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. USA: Oxford University Press.

Delsing, M., ter Bogt, T., Engels, R., & Meeus, W. (2008). Adolescents' music preferences and personality characteristics. *European Journal of Personality*, Vol. 22, No. 2, pp. 109-130.

DeNora, T. (1999). Music as a technology of the self. *Poetics*, Vol. 27, No. 1, pp. 31-56.

Donnadieu, S. (2007). Mental representation of the timbre of complex sounds. In: J. W. Beauchamp (Ed.), *Analysis, synthesis, and perception of musical sounds: The sound of music.* New York: Springer.

Dura, M. (2006). The phenomenology of the music-listening experience. *Arts Education Policy Review*, Vol. 107, No. 3, pp. 25-32.

Eerola, T., Alluri, V., & Ferrer, R. (2008). Emotional connotations of isolated instruments sounds. In: *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC)*. Sapporo, Japan: University of Hokkaido, pp. 483-489.

Fales, C. (2002). The paradox of timbre. *Ethnomusicology*, Vol. 46, No. 1, pp. 56-95.

Fales, C. (2005). Listening to timbre during the French enlightenment. In: C. Traube & S. Lacasse (Eds.), *Proceedings of the 2005 conference on interdisciplinary musicology (CIM).* Montréal, Québec, Canada: Centre for Interdisciplinary Research in Music Media and Technology.
Available from http://www.oicm.umontreal.ca/doc/cim05/articles/FALES C CIM05.pdf

Ferrer, R. & Eerola, T. (2010) Timbral qualities of semantic structures of music. In: *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR)*. Utrecht, Netherlands, pp. 571-576.

Ge, J., & Hokao, K. (2005). Applying the methods of image evaluation and spatial analysis to study the sound environment of urban street areas. *Journal of Environmental Psychology*, Vol. 25, No. 4, pp. 455-466.

Gibson, J. (1986). *The ecological approach to visual perception*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Godøy, R. (2006). Gestural-sonorous objects: embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound*, Vol. 11, No. 2, pp. 149-157.

Gregory, C. (1999). Stereotypes and personalities of musicians. *Journal of Psychology*, Vol. 133, No. 1, pp. 104-114.

Grey, J. (1977). Multidimensional perceptual scaling of musical timbres. *The Journal of the Acoustical Society of America*, Vol. 61, No. 5, pp. 1270-1277.

Hajda, J. (2007). The effect of dynamic acoustical features on musical timbre. In J. W. Beauchamp (Ed.), *Analysis, synthesis, and perception of musical sounds: The sound of music.* New York: Springer.

Held, C. & Cress, U. (2009). Learning by foraging: The impact of social tags on knowledge acquisition. In: *Learning in the Synergy of Multiple Disciplines: 4th European Conference on Technology Enhanced Learning (EC-TEL*). Nice, France: Springer, pp. 254-266.

Helmholtz, H. (1954). *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover Publications.

Hesse, F. (2009). Use and acquisition of externalized knowledge. In: *Learning in the Synergy of Multiple Disciplines: 4th European Conference on Technology Enhanced Learning (EC-TEL).* Nice, France. Springer, p. 5.

Hinton, L., Nichols, J., & Ohala, J. J. (Eds.). (1995). *Sound symbolism.* Cambridge: Cambridge University Press.

Huron, D. (2001). Toward a theory of timbre. Paper presented at the 12th Annual Conference of Music Theory Midwest. Abstract retrieved September 6, 2010, from http://www.musiccog.ohio-state.edu/Huron/Talks/SMTmidwest.2001/talk.01.html

Kaipainen, M. (1996). Prospects for ecomusicology: Inner and outer loops of the musical mind-environment system. In: P. Pylkkänen, P. Pylkkö, & A. Hautamäki (Eds.), *Brain, Mind and Physics*. Netherlands: IOS Press, pp. 266-277.

Krumhansl, C., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 18, No. 3, pp. 739-751.

Laurier, C., Sordo, M., Serrà, J., & Herrera, P. (2009). Music mood representations from social tags. In: *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR)*, Kobe, Japan, pp. 381-386.

Lamere, P. (2008). Social tagging and music information retrieval. *Journal of New Music Research*, Vol. 37, No. 2, pp. 101-114.

Leman, M. (1995). *Music and schema theory: Cognitive foundations of systematic musicology*. Berlin, Heidelberg: Springer.

Leman, M. (2007). *Embodied music cognition and mediation technology*. Cambridge: MIT Press.

Maturana, H. (1988). Ontología del conversar. *Revista Terapia Psicológica*, Vol. 10, pp. 15-23.

Maturana, H. (2002). Autopoiesis, structural coupling and cognition: A history of these and other notions in the biology of cognition. *Cybernetics & Human Knowing*, Vol. 9, No. 3-4, pp. 5-34.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, Vol. 58, No. 3, pp. 177-192.

McMullen, E., & Saffran, J. (2004). Music and language: A developmental comparison. *Music Perception*, Vol. 21, No. 3, pp. 289-311.

Mehler, J., Bertoncini, J., Barriere, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception*, Vol. 7, No. 5, pp. 491-497.

Nussbaum, C. (2007). *The musical representation: Meaning, ontology, and emotion*. Cambridge: MIT Press.

Pachet, F. & Cazaly, D. (2000). A taxonomy of musical genres. In: *Proc. Content-Based Multimedia Information Access* (RIAO), pp. 1238-1245.

Patel, A., & Iversen, J. (2003). Acoustic and perceptual comparison of speech and drum sounds in the north indian tabla tradition: An empirical study of sound symbolism. In: M. J. Solé, D. Recasens, & J. Romero (Eds.), In *Proceedings of the 15th international congress of phonetic sciences (ICPhS).* Barcelona, Spain: ICPhS, pp. *925-928.*

Popper, K., & Eccles, J. (1984). *The self and its brain*. New York, USA: Routledge.

Rentfrow, P., & Gosling, S. (2003). The do re mi's of everyday life: Examining the structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, Vol. 84, pp. 1236-56.

Rentfrow, P., & Gosling, S. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science*, Vol. 17, No. 3, pp. 236-242.

Rentfrow, P., & Gosling, S. (2007). The content and validity of music-genre stereotypes among college students. *Psychology of Music*, Vol. 35, No. 2, pp. 306-326.

Reybrouck, M. (2005). A biosemiotic and ecological approach to music cognition: Event perception between auditory listening and cognitive economy. *Axiomathes*, Vol. 15, No. 2, pp. 229-266.

Sandell, G. (1998). Macrotimbre: Contribution of attack, steady state, and verbal attributes. *The Journal of the Acoustical Society of America*, Vol. 103, p. 2966.

Sandell, G., & Chronopoulos, M. (1997). Perceptual constancy of musical instrument timbres; generalizing timbre knowledge across registers. In: A. Gabrielsson (Ed.), *Proceedings of the third triennial ESCOM conference*. Uppsala: Uppsala Unversity, pp. 222-227.

Sarkar, M., Lan, C., Diaz, J., & Vercoe, B. (2009). The effect of musical experience on describing sounds with everyday words. *The Journal of the Acoustical Society of America*, Vol. 125, p. 2683.

Schafer, R. M. (1977). *The tuning of the world*. Toronto, Canada: McClelland & Stewart.

Seashore, C. (1967). *Psychology of music*. New York: Dover Publications.

Sethares, W. (1999). *Tuning, timbre, spectrum, scale*. Great Britain: Springer.

Terasawa, H., Slaney, M., & Berger, J. (2006). Determining the euclidean distance between two steady state sounds. In: M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th international conference on music perception & cognition, international conference on music perception &*

cognition. Bologna, Italy: ICMPC-ESCOM. Retrieved September 6, 2010, from https://www-ccrma.stanford.edu/~hiroko/timbre/Terasawa2006_ICMPC9.pdf

Tye, M. (2008). Dualism. In: *Stanford Encyclopedia of Philosophy*. Stanford, CA: The Metaphysics Research Lab. Retrieved October 1, 2008, from http://plato.stanford.edu/entries/dualism/

Wessel, D. (1979). Timbre space as a musical control structure. *Computer Music Journal*, Vol. 3, No. 2, pp. 45-52.