Marc Leman

Foundations of Musicology as Content Processing Science

Summary

This paper describes musicology as a science of musical content processing. It is argued that evidence-based research and computational modeling provides a methodological basis for understanding the human information processing mechanisms that transform physical information streams into musical content and meaningful musical usage in different contexts. The focus on musical universals leads to an instrumental theory with applications in interactive music systems and audio-mining.

Keywords: music, meaning, content, musicology

3.1. Introduction

Science, technology, and industry have a profound impact on society, and it is quite natural that many aspects of musical culture, including the meaning of music, undergo changes concurrently with these developments. Several stages in the development of our society have been characterized by particular infrastructures, and musical culture has, in its turn, been largely influenced by these infrastructural activities.

In origin, music almost certainly had a cultural-evolutionary function displaying reproductive fitness and fostering group identification. Later on, it shaped the affective framework for joint labor, popular entertainment and religious services.

Music still has these functions today, but additional layers of meaning have been added. In the *industrial* area, music became a commodity in the hands of concert organizers and the

printing industry, paving the path for more attention to the abstract functionalities of music, and for listening to music as an art form, with particular attention to meanings intrinsic to the music itself, apart from its functional purposes. The distribution and consumption of sheet music and recorded music by the end of the nineteenth century enhanced the conditions for the materialization of music as an audio phenomenon and for the mass commercialization of the associated commodities.

In our *post-industrial* age, the interactions between culture, commerce, industry, and research have evolved to a point where music is a commodity associated with technology in a multi-media-embedded environment. We live in an era during which electronic delivery, human-machine interaction, and data-mining exert profound influences on music and its meaning. The listener has become an active consumer, choosing parts of the music (s)he wants to hear, making music with fragments from other music (Sabbe, 1996). The multi-media-embedded environment offers many new challenges and opportunities for music composition, performance, recording, distribution, and consumption. All these contexts call for *musical content processing applications* as an integrated part of the way in which people, through technology, engage with music.

It is worth mentioning from the outset, however, that the present paper is not about musical culture as such, nor about technology or an analysis of the music industry. Rather it is about the nature and role of modern (systematic/cognitive) musicology in relation to content and meaning. In what follows, I provide a picture of musicology as a *science of musical content processing*. Contrary to post-modernist verbal gymnastics and endless introspective argumentation about the justification of personal musical experiences, I claim that a study of musical meaning can be grounded in a broad scientific approach to content processing. I see musical meaning as the usage of content in particular contexts of creation, production, distribution, and consumption, and, furthermore, modern musicology as a science of musical content processing that fosters the development of instrumental theories of meaning and important applications in domains such as musical audio-mining and interactive multimedia systems for theater. The science of musical content processing aims at explaining and modeling the mechanisms that transform information streams into meaningful musical units (both cognitive and emotional).

The structure of the paper is as follows. First (Sect.2), music research is considered from a societal point of view, with an emphasis on the historical and social-economical background. Then (Sect.3), a broad scientific approach to musical content processing is presented, based on a distinction between different orientations in music research, and examples are given of concrete research results and possible future directions in musicology. The epistemological and methodological issues involved are handled in a next section (Sect.4). A final section (Sect.5) is devoted to applications in two different areas, in particular interactive music systems and musical audio-mining.

3.2. Historical and Socio-Economical Perspectives

To understand musicology as a science of musical content processing, I start with a sketch

of the historical trajectory. A brief section on the socio-economical perspectives serves to highlight the mission and motivation of many interdisciplinary scientific projects in this research domain.

The historical perspective concerns the trajectory that music research followed over time and that ultimately led to the idea that a physical stream of information can be processed into entities that convey content and meaning. Without going into too many historical details it is worth mentioning that this trajectory starts with the ancient Greeks, in particular their fascination with acoustics and the discovery of particular numerical relationships between sounds. The roots are reflected in the traditions of Pythagoras and Aristoxenos. The first focused attention on the mathematical order underlying harmonic musical relations, while the second concerned itself with perception and musical experience (Barker, 1989). This distinction between acoustics and practice, is still relevant today. What modern musicology offers is, in fact, a connection between acoustics and musical practice. Starting from the Ancient Greeks, the trajectory can be followed throughout different stages.

The pre-industrial stage spans from ancient times over medieval times to modern times until the late 18th century. The research interests culminated in a scientific revolution (Koyré, 1957) initiated by towering figures such as J. Kepler, S. Steven, G. Galilei, M. Mersenne, I. Beeckman, R. Descartes, C. Huygens, L. Euler, and others, who transformed the Pythagorean music tradition into a genuine acoustical science of music (Cohen, 1984). The tradition initiated by Aristoxenos culminated in recipes for musical practice such as G. Zarlino's, and later J. Rameau's (Rameau, 1722/1965). Today we can say that the main contribution of this period was a *theoretical foundation* for music research.

The industrial stage, initiated by the French revolution and the industrial revolution in England (Hobsbawn, 1962), is characterized by an extensive use of the experimental methodology. Landmarks are Lord Rayleigh (1877/1976), a major work in acoustics, and H.v.Helmholtz (1863/1968) and C. Stumpf (1883/90), which provide the foundations of psychoacoustics, and Gestalt psychology, respectively. Acoustics, psychoacoustics and music psychology form the *empirical basis* of modern musicology (Leman & Schneider, 1997).

The post-industrial stage is characterized by electronic communication services and information processing technologies that are grounded in older agricultural and industrial layers of technology (Bell, 1973). Many new possibilities for enhanced musical culture were created in the wake of the establishment of radio broadcasting companies in the late 1940's. These institutions became engaged in artistic experiments that aimed at extending the musical soundscape for radiophonic applications. New electronic equipment capable of generating and manipulating musical sound objects was needed and appropriate infrastructures were set up to realize this goal (Veitl, 1997). This was the start of what is

known today as *music engineering*, an area that emerged from the close collaboration between engineers and composers (e.g. MeyerEppler, 1949; Moles, 1952; Feldtkeller & Zwicker, 1956; Winckel, 1960).

The 1960s were also a starting point for new developments in music research (e.g. Schaeffer, 1966, Rösing, 1969), culminating in a great interest in semiotics in the 1970s (see Monelle, 1992; Leman, 1999b for overviews) and which led to the development of cognitive musicology in the 1980ies and 1990ies (Leman, 1997; Zannos, 1999; Godøy & Jørgensen, 2001). The new paradigm is grounded on evidence-based research and computational processing, forming an elaborate testing environment for the development of theories that, ultimately, aim at the processes that transform information streams into musical content and meaning, thus closing the old gap between sound and musical practice. This approach will be described in more detail in the following sections.

Apart from theory, experiment and technology, there are signs that a new fundamental component will be added to music research, namely biology (Zatorre & Peretz, 2001). Biology indeed has the potential to revolutionize our thinking about music in a rather profound way, through the introduction of an evolutionary perspective, and the focus on physiology. Motor mimesis theory, physical modeling, and evolutionary neurobiology provide many new ideas for how to look at memory and musical engagement. For example, how do we remember a musical piece? Do we store the acoustical traits of the music, or perhaps only the parameters that allow its synthesis? The latter is far more efficient in terms of memory storage consumption. Also body movement may have a profound influence on our understanding of musical affect processing and expressiveness, as well as Darwinist theories that study the human auditory system from the viewpoint of evolution. All of this is most exiting for the study of musical content processing and meaning formation.

To summarize: Based on ancient Greek groundwork, mathematicians and musicians of the pre-industrial area pioneered the foundations for a scientific theory of music that was given an empirical foundation in the industrial area, and a technological foundation in the post-industrial area. All that culminated in the last quarter of the 20th Century in a research paradigm for musical content processing that is grounded on evidence-based computational modeling and biological foundations.

Music, as one of the most important cultural phenomena in modern society, is strongly embedded in commercial activities, which, as is well known, is dominated by a few major companies, followed by a large number of smaller companies (Pichevin, 1997). Although recorded music sales worldwide fell to 32.000 milion US$ (or €) in 2002 (compared to 2001, sales of CD albums fell globally by 6%, and there were continued declines in sales of singles (-16%) and cassettes (-36%)) (IFPI, 2002), the music industry is still a very large commercial enterprise.

But despite this huge consumer market and the large network of interconnected music industrial activities, musical content has long been underestimated as a potential source of

combined academic, cultural, industrial, technological and economical activity. This attitude is now rapidly changing as many people start to be aware of the opportunities for arts and industry. I cannot otherwise explain the large-scale interest, nowadays, displayed by psychologists and engineers in these matters. A good theory of musical content processing is not only of scientific interest, but will have a large impact on culture and economy.

3.3. Musical Content Processing: Intuitive-Speculative and Empirical-Experimental Orientation

Musicology as a content processing science has a strong tradition, and its socio-economical perspectives look good. But what is musical content processing? What are its basic objectives and methods?

An answer to this question should be given along three axes, that I call (i) intuitive-speculative, (ii) empirical-experimental, and (iii) computational-modeling. Together, these axes define the scientific workspace for musical content processing. This section aims at capturing the essence of this workspace together with some concrete examples of research results. Whenever possible, examples will be taken from my own research activities.

In the 1960's, when musicology became an established academic discipline in many universities all over the world, music theory, semiotics, and hermeneutics brought into the foreground problems related to imagination, natural language, and the aesthetic quality of musical products. The questions pertained to a broad range of topics such as: How are musical concepts emerging in the human mind? How can we describe non-verbal objects of music in terms of a natural language? What is the mutual relationship between music, affect and emotion? What are the sociological factors that contribute to the reception of music? How can we understand the music of other cultures? What is the meaning of musical expressiveness?

These questions were considered to be the central questions that could lead to a better understanding of music as a cultural phenomenon. The focus was on musical meaning, seen as the outcome of a complex web of different content layers that function in particular social and historical contexts.

The questions have been handled from quite different points of view (see Leman, 1999b), including phenomenology and hermeneutics (Schaeffer, 1966; Faltin & Reinecke, 1973; Stefano, 1975), and sometimes viewpoints that were inspired by cybernetics and the up-and-coming area of brain science (Posner & Reinecke, 1977). These approaches have been very useful and necessary for the development of musicology as a content processing science. Unfortunately, some good insights lacked empirical foundation and many theories remained metaphors.

Serious work in aesthetics, semiotics and philosophy of music should always try to

understand the implications of empirical-experimental and computational-modeling approaches. When the empirical foundations are neglected, musicology has the tendency to quickly end up in endless discussions and speculations about personal experiences. If musicology wants to be a science, rather than an art, however, it cannot turn a deaf ear to those approaches that study music as something that is largely constrained by physical information streams and biological processes. For a field of endeavor to be a science, indeed, means that reliable knowledge can be obtained. The motivation for doing that is provided by the socio-economical context.

In what follows, I will argue that an exclusive focus on culture for understanding music, is too restricted, as is an exclusive focus on nature. Both components are interconnected and theories should take this connection into account. I argue, therefore, that there is an urgent need for cultural research based on quantitative descriptions, data gathering, and statistical analysis to map out cultural tendencies and regularities. This type of cultural research differs from psychology in that it is less focused on sensory, perceptive and cognitive mechanisms and therefore, perhaps, less apt to involve short-term memory experiments, testing and prediction. Culture-based music research will be about social aspects and traditions, in other words, about an understanding of music as a phenomenon that is expressive of particular cultural dependencies.

To summarize: the study of music as a cultural phenomenon is more than just a series of discussions about personal experiences with music. A scientifically justified study of the musical phenotype has to be put into relationship with empirical-experimental and computational-modeling approaches. The request for more rigorous methodologies in the cultural sciences forms part of my epistemological view, which I will introduce in the following sections. Intuition and speculation nevertheless remain necessary as starting points for research, for example, to define work strategies, to "feel" what to research, and to formulate hypothesis.

In my framework, I connect the empirical-experimental approach with research in psychoacoustics, music psychology, and neuromusicology. This type of research is typically concerned with the position of an individual or group of individuals in its *embedded* physical and cultural environment. Interactions between the subject and the environment, however, are determined by constraints that operate at different time-scales and that have their effects on the subject's content processing abilities. While cultural research of music can of course be evidence-based, I will highlight the psychological aspect.

*** Figure 1 [see internet version of this article] ***

Figure 1 gives a schematic picture of the physical and cultural interactions of an individual with its environment. The individual (small circle) is genetically shaped through evolution (arrow from left to right). The interactions are constrained by social and cultural conditions that undergo changes at medium time-scales (upper circle) and sensory, cognitive, affective and motorical interactions that are situated at relatively short time-scales (lower circle). The figure shows that the empirical-experimental orientation has to take into

account a rather complex layer of processes that are never completely under control.

Evidence-based research is most often the study of content processing in short-term variability under (sub-optimal) controlled conditions (lower circle). It is concerned with sensory, cognitive, affective and motorical interactions, rather than effects of natural selection, or cultural interactions. In contrast, the cultural approach discussed in the previous section has a focus on the upper cycle, whereas music archeology and evolutionary neurobiology would have a focus on the horizontal time scale. Strictly speaking, therefore, the research focus in psychology is necessarily limited, and its results should be evaluated within the broader cultural and evolutionary framework.

Psychoacoustics, music psychology, and brain research are the main research areas of the empirical-experimental orientation.

3.4. Psychoacoustics and Music Psychology

Psychoacoustics and music psychology aim at mapping out musical behavior in terms of psycho-physical mappings. Psychoacoustics is concerned with the relationship between physical reality and sensory processing (Moore, 1989; Zwicker & Fastl, 1990), while music psychology is concerned with musically relevant representational structures, both cognitive and affective, that are extracted from reality on the basis of the senses (Deutsch, 1982). The main difference, perhaps, is the time scale of information processing: less than about 3 seconds (the subjective awareness of the present) in psychoacoustics, usually more than 3 seconds in music psychology. Obviously, the shift to longer time scales in music psychology may involve the use of different methods to uncover content processing. The longer the time scales, the more impact we may expect from cultural constraints. Very short reactions and responses rely more on natural constraints.

Experiments aim at isolating, defining, and quantifying the musical functions through which a subject transforms an input (the physical information stream or musical signal) into an output (musical content). Explanation is then the interpretation of observed data in terms of such a mapping. The focus on input/output allows a link with computational modeling where functional aspects can be studied from the viewpoint of dynamics and representation, using so-called functional equivalent models, that is, models that are functionally equivalent to the physical cause/effect relationships (which is often unknown).

Research has been focusing on *syntactical* properties that can be extracted in an objective way from musical audio, such as pitch, harmony, tonality, rhythm, beat, texture. These properties are to be distinguished from *semantical* properties that express the music's subjective emotions and affects.

A main motivation for doing this type of research is that it may lead to the discovery of *musical universals* or regularities of sensory, cognitive, affective and motorical content processing. Universals are key components in a theory of musical content processing because they are assumed to be invariant to the local or instantaneous context. In doing

so, they can operate as functions for automated content extraction from musical audio streams. Universals give the theory an *instrumental* basis. Contrary to style regularities, universals do not pertain to regularities at the phenomenological level (such as regularities between musical compositions of the Romantic area in Central Europe), but to regularities that determine these cultural phenomena. Universals are rooted in nature, but have effects in culture. This is similar to the genotype that determines the form and structure of a plant, while the conditions in which it grows define the phenotype. The interactions between nature and culture, obviously, are very complex, which make it hard to predict music at the level of phenotypes. Yet knowledge about the musical genotype is feasible and very much at the core of any theory of musical content processing. A few examples should clarify the meaning and implication of musical universals.

3.4.1. Roughness

Roughness is experienced as that instance of sound texture having an unpleasant character. It occurs when the sound contains beating frequencies at a rate of about 25-100 times a second. v. Helmholtz (1883/1968), Plomp (1966) and many others since then (e.g. Kameoka & Kuriyagawa, 1969; Pressnitzer, 1999) report measured sensations of sounds having these beating properties: roughness is maximal at about one fourth of the critical bandwidth (Plomp, 1976) and corresponds more or less with 70 Hz.

While roughness is a determining factor in the affective judgment of tone intervals, the basic reasoning is that, of all possible pitch intervals between two sounds, the intervals with a lowest beating character will be preferred because of their most *pleasant* texture. A link can then be formed between the sound properties of a sound, and the pitch scale. When *harmonic* tone complexes are used, for instance, the intervals of the diatonic scale pop out, see Fig.2.

*** Figure 2 [see internet version of this article] ***
*** Figure 3 [see internet version of this article] ***
*** Figure 4 [see internet version of this article] ***

Whether this is a genuine universal depends on its applicability to new contexts. For example, when *inharmonic* tone complexes are used, what is, then, the implied pitch scales that people would prefer? Sethares (1997) has tested this idea to the study of scales in Indonesian culture and found that the combination of a bonang with a harmonic sound generates the slendro-scale (Fig.3). Also pelog scales (Fig.4) can be predicted. The theory furthermore predicts that the tunings between different Gamelan orchestras depends on the timbres of the instruments used in a particular Gamelan orchestra. When timbres differ from each other, they will lead to slightly different tunings.

The relationship between sound spectrum and pitch scale is a most fascinating phenomenon that points to the intricate connection between physical constraints (sound properties), physiological constraints (auditory system), and structures (the resulting scale) that get established in a culture.

Hence the hypothesis that *minimalisation of roughness* is a universal, allowing predictive

statements concerning the relationship between the physical aspects of sound and musical practice. Whether this is a *true* context-free universal remains to be seen because units of cultural information might exert top-down influences on what aspect of the sound is actually perceived (Schneider, 1997a). It might be that such universals are conditioned by very strong cultural tendencies, known under the concept of *memes* (Dawkins, 1976). Universals, in other words, might be culturally conditioned and unfold their operation only in favorable circumstances.

I favor the idea of adopting a data-driven (genotypic) research strategy for the understanding of musical phenomena, basically for pragmatical reasons. Everything that cannot be explained by this bottom-up strategy could then be hypothesized to be an effect of cultural influences and would then need further investigation relative to particular contexts.

3.4.2. Pitch Perception

Results in pitch perception offer a similar picture. What has been aimed at is a better understanding of (i) the perception of isolated pitches, (ii) the perception of context-dependent pitch, and (iii) the perception of stylistic aspects.

(i) Theories of pitch perception have gradually been refined due to a growing insight in the underlying auditory principles (Terhardt, Stoll, & Seewann, 1982; Meddis & Hewitt, 1991; Langner, 1992). The result is a modern conception of pitch that is very different from the traditional (musicological) concept of pitch as a fixed parameter of music. The modern concept is more dynamic, and considers pitch as an emerging quality from underlying auditory principles. The key notion is that the perceived pitch of a given tone complex may be *induced* given the relationships among the frequencies of the tone complex. The perceived pitch is not necessarily a frequency of the tone complex itself. Roughly speaking (though only usable as a rule of thumb), the perceived pitch is the smallest common period of the sound signal (Langner, 1997). If a tone complex of 600, 800 and 1000 Hz is presented, the periods have a length of 1.667, 1.25, and 1 millisecond. The greatest common period is 5 milliseconds because the period of 1.667 milliseconds fits 3 times, the one of 1.25 milliseconds 4 times, and the one of 1 millisecond 5 times. The pitch heard corresponds with the most salient period, which in this case is 5 milliseconds or 200 Hz. The salience of the induced pitches has been related to the tonal stability of musical chords (Terhardt, 1974). The theory explains for instance why a major triad chord is more stable than a diminished triad chord (Parncutt, 1997). It is because the induced pitch is less ambiguous in terms of common periods.

When a culture focuses on stable tone complexes, we might conjecture that *minimalisation of the ambiguity of induced pitch* is a candidate for universality. But obviously, pitch is in fact determined by a context of other pitches that should be taken into account as well. Cultural values may play a role in the perception of pitch and force the listener's attention to a particular aspect of tone perception, which is not necessarily the pitch.

A typical example is the tone of a carillon. The tuning of most of the carillons would be unbearable and not acceptable compared to a string quartet, for example. Yet, when playing the bells, the deviances are accepted, even appreciated (Houtsma & Tholen, 1987). The music displays qualities that listeners associate with interesting sound timbres and social meanings such as the pride of a city. Subjects that focus on pride might hear the inharmonic pitches and chords of the carillon as acceptable music. Unfortunately, these top-down influences of context-dependent cultural values remain badly understood phenomena of pitch perception (Burns & Ward, 1982; Schneider, 1997a). In this context, therefore, it is important to see how far we can go in explaining the working of carillon music from a bottom-up manner, in order to better comprehend the impact of culture.

(ii) Music psychology has contributed to a better understanding of context-dependent pitch perception of Western tonal music, that is, how pitch perception behaves in a musical context of other pitches. Krumhansl and colleagues (1979, 1982) developed an experimental technique which clearly demonstrated that pitch perceptions are hierarchically structured, for example, that in the context of C-major, the notes C, E, F and G are more prominent than D and A, and that D and A are even more prominent than C#, Eb, F#, Ab, Bb. The procedure through which it was obtained set the stage for a number of controlled experimental designs for studies in music perception (Vos & Leman, 2000).

All these studies draw on the idea that memory structures that underlay behavior can be monitored and represented as maps. One finding is that stable memory structures for pitch may determine perception in a top-down manner. Memories thus provide a binding factor between what is physiologically constrained and what is culturally exemplified. This explains the enormous appeal of Romantic music, or music that exploits the schemata of the tonal system like pop, rock and jazz music. The building of tension and relaxation using pitch, melody, harmony, and tonality draws on regularities in pitch syntax that optimally fit with both short-term and long-term pitch processing mechanisms. For tonal music we can put forward the hypothesis that *tension is the degree of fit between locally induced pitch and globally induced pitch patterns.* This offers an instrumental definition of the important concept of tonal tension, ready to be tested in a computational model (Leman, 1990; Huron & Parncutt, 1993; Leman, 1995, 2000).

(iii) Memories for styles typically involve traces of temporal deployments of certain musical parameters. These traces may be captured by statistical methods as well, provided that the statistics relates to distributions of transitions of musical characteristics, such as distributions of pitch transitions, or distributions of timbre transitions, or particular rhythmic configurations. Pieces in different styles such as Debussy and bebop jazz music may have similar pitch distributions on the whole, yet most listeners are able to distinguish the different styles. Style effects have to do with distributions of pitch transitions, and their stylistic grammar may be defined in terms of distributions of two, three or even four pitches (Eerola, Jarvinen, Louhivuori, & Toiviainen, 2001).

3.4.3. Beat, Tempo and Rhythm Perception Experiments

A beat is defined as the subjective tapping to the musical pulse. The meter refers to a level at which beats are structured, for example into groups of two or three. The tempo is expressed as the number of beats per minute. A rhythm, then, can be defined as a pattern that evokes a sense of pulse. Rhythm research has focused on perception and performance ranges as well as on the differences and deviations from nominal values (Fraisse, 1982). As in pitch perception, rhythm perception involves the *induction* of subjective responses to certain acoustical properties of the music (Desain & Windsor, 2000). A common paradigm is to ask listeners to tap the beat of synthesized patterns or music and then to measure the beat in terms of grouping and inter-onset-intervals (Essens & Povel, 1985; Parncutt, 1994). Van Noorden and Moelants (1999) found that data fit rather well with the hypothesis of a resonator forcing beat perception to be structured by a period of about 500 milliseconds (Fig.5).

*** Figure 5 [see internet version of this article] ***

Measurements of tempi in different musical genres lead to a distribution that comes very close to this resonance curve. As in pitch research, rhythm is highly context-dependent and stylistic aspects play a similar important role.

3.4.4. Affect Research

Many people think that true meanings in music are about affects and emotions (Panksepp & Bernatzky, 2002). Finding relationships between properties of music and affect perception is important for musical content processing even if the mechanism through which musical audio gets connected with affect and emotion are unknown at present. A main conjecture is that music has certain syntactical properties whose form and dynamics allow us either to recognize the form and dynamics of experienced affects and emotions (=affect attribution), or to experience certain affects and emotions (=affect induction). Sometimes, our focus is more on the cognitive aspect; sometimes we are more apt to engage ourselves in aesthetic experiences. Much depends on the context in which affect processing is taking place (Scherer & Zentner, 2001). That is why the affective meaning of music is often difficult to predict: it often depends on personal histories that science can never fully take into account.

Yet it is possible to uncover certain trends in musical affect processing and the related musical appreciation and aesthetic experience. People can communicate about expressiveness in art using descriptors that relate to affects and emotions. There exists an affect semantics grounded on inter-subjective regularities of affect perception that can be mapped out. By adopting our pragmatic strategy, we can see how far we can follow this inter-subjective path before we are forced to take into account the subjective component.

*** Figure 6 [see internet version of this article] ***

Figure 6 shows a typical example of an affect semantics obtained after Factor Analysis of 100 subjects describing 60 musical excerpts with 30 adjectives (Leman et al., 2003). The

first axis reflects the Valence or the distinction between favorable and unfavorable affect adjectives, the second axis reflects the Activity or the tender-bold aspect of the adjectives, while the third axis reflects the Interest, boring or exciting. The first two axes are inter-subjectively rather stable and they work well in cases of affect attribution. The Interest-axis, however, is inter-subjectively unstable. The Interest adjectives probe responses from listeners in different ways for different reasons.

*** Figure 7 [see internet version of this article] ***

The interesting thing is that affect semantics can be related with the syntactical features, so that regularities can be studied between syntactical cues (audio structures) and semantical cues (affect adjectives) (See Fig. 7). Several cues have been identified and related to affect processing (see Gabrielsson and Juslin, 2003). For example, tempo is considered to be a most important factor affecting emotional expression in music. Fast tempo is associated with various expressions of activity/excitement, happiness, potency, anger and fear and slow tempo with various expressions of sadness, calmness, dignity, solemnity, dignity. Loud music may be determinant for the perception of expressions of intensity, power, anger and joy whereas soft music may be associated with tenderness, sadness, solemnity, and fear. High pitch may be associated with expressions such as happy, graceful, exciting, anger, fear and activity and low pitch may suggest sadness, dignity, excitement as well as boredom and pleasantness.

Is the above list of connections between syntax and semantics a musical universal? At this moment it is difficult to say because this list looks more like the surface structure of more fundamental biological mechanisms. Moreover, cross-cultural research is needed to better test its validity outside the western culture. It may be that this semantics, as well as, its physical determinants, is to some extend dependent on populations and cultural properties.

To summarize: the above examples illustrate the relevance of the behavioral empirical-experimental approach for music research. The evidence-based explanations lead to interesting data about mappings between sound and musical behavior. These can be modeled with a computer and perhaps further refined or used in applications.

3.5. Music and brain science

I find it necessary to address very briefly the topic of music and brain science because this approach fits very well within the global picture of modern musicology. Music and brain research profited from spectacular progress in non-invasive brain imaging technology (Toga & Mazziotta, 1996) that allowed the study of brain regions that are actively involved in musical content processing (Zatorre & Peretz, 2001; Tervaniemi & Leman, 1999). fMRI is nowadays one of the most powerful methods to study the human brain and Zatorre has used fMRI in a paradigm that focuses on observed brain activation when a prescribed perceptive task (e.g. listening to a melody or imaging a melody) is performed (Zatorre, Halpern, Perry, Meyer, & Evans, 1996; Zatorre, 1997).

The combination of fMRI, behavioral responses, and computer modeling seems to be a straightforward approach (Zatorre & Krumhansl, 2002). The interpretation of brain activity as it evolves over time is a major issue, and computational models whose output can be related to a clear semantic interpretation of the brain activity can cause breakthroughs in this area. Janata et al. (2003) used a methodology that allowed a meaningful interpretation of fMRI-data obtained from listeners involved in a tonality perception task (taking several minutes). A computer model (Leman, 1995, 2000; Leman, Lesaffre, & Tanghe, 2001) processed the same audio stimuli and the output was correlated with brain units showing activation. The techniques allow the study of music perception, using natural musical stimuli rather than short artificial stimuli. Although the methodology needs to be further refined, this methodology seems have great potential as a means for studying brain activity in response to high-level affective and technical cues, including the brain localization of affective cues.

Summarizing this brief section, music and brain research (or: neuromusicology) offers a rather strict evidence-based approach to musical content processing. Its relevance is to be situated in connection with epistemological issues such as the problem of localization of content processing, whether pre-attentive processing takes place, or what kind of memory is involved. Neuromusicology is a young field and its technological basis is rapidly evolving so that many new results may be expected in the future.

3.6. Sound Computation and Modeling

The third pillar of musicology is the computational-modeling approach. This field witnessed an explosive growth in the 1980ies and it can be roughly divided into (i) sound computation, and (ii) modeling.

With the development of electronic equipment and its application in music production studios since the 1950s, sound analysis became part of the musicological toolbox. Music studios in the 1950s applied electronics for music generation (electro-acoustic music, 'musique concrête'). This was possible due to collaboration between engineers and composers (Veitl, 1997), see Meyer-Eppler (1949), Moles (1952), Feldtkeller and Zwicker (1956), Winckel (1960). Musicology followed with new music descriptions and sound-based analysis (Schaeffer, 1966; Rösing, 1969), which culminated in a general interest in nature-grounded semiotics and cybernetics. Digital electronic equipment and the spread of personal computers has made sound analysis fast and accessible. Sound analysis is now a well-established tool in the hands of musicologists. Typical areas in which sound analysis is applied are analysis of timbre and of sound qualities, tonometry, studies of musical expression, and content annotation, using techniques of spectral and physical modeling (Serra, 1997; Karljalainen et al., 2001). General introductions to music and sound computation are provided in Roads et al. (1997) and Zölzer (2002).

When the aim is to build an instrumental theory of content processing, then the modeling of certain aspects of content processing becomes necessary. My feeling is that in order to be effective, content processing models should rely on knowledge of the human

information processing system. My argument, however, is an epistemological one.

Modeling of musical perception and cognition began in the 1960s and was associated with sound modeling (Risset & Mathews, 1969) and artificial intelligence (Winograd, 1968; Longuet-Higgings, 1987; Laske, 1977; Kunst, 1976). Modeling is nowadays considered to be an essential component of music research and has already been used to test a particular notion of universal content processing (such as minimalisation of roughness, see Sethares, 1997), to refine an evidence-based hypothesis (e.g. Krumhansl's 1990 tonal hierarchy hypothesis, see Leman, 2000, and below), or to develop applications (e.g. interactive music system, Camurri et al., 2001), or to help understanding brain data (e.g. Janata et al., 2002).

In all these cases, modeling had a focus on sensory, perceptive, cognitive, motoric, and affective/emotional processes. Modeling turns out to be an excellent tool in the hands of those that study musical content processing from bottom-up. Using a model, a theory can be cast in operational terms, providing a basis for testing and (through testing) a better understanding of the content representations and the dynamics of content processing.

One example in the domain of context-dependent pitch processing may illustrate the role of computer modeling. As mentioned above, results in tonal music perception have led to an interesting discussion concerning the status of cognitive pitch processing, in particular the role and functioning of short-term memory and long-term memory. A good starting point for the discussion is Krumhansl and Kessler (1982), where it is stated that the subject's estimated profiles exhibit a torus structure and relative ordering in terms of the circle of fifths (Fig. 8).

*** Figure 8 [see internet version of this article] ***

Given these results, it is indeed straightforward to conjecture that listeners familiar with Western music would have abstracted tonal hierarchies in a long-term memory (or schema) for pitch, as reflected in the data from probe-tone experiments.

Computer modeling, based on models of virtual pitch perception and statistical long-term memories, indeed seemed to confirm this hypothesis at first sight. Simulations proved that schemata for tonal distributions may be built up by mere exposure to tonal music on the basis of self-organization (Leman, 1995). Given the sound signals of three different cadences played in all keys as input during learning, a computer memory could be trained by simple self-organization principles to hold the tonal distributions in terms of ordered structures (circles of fifths), providing results (see Fig.9) that are very similar to topographies of the key profiles (Fig.8).

*** Figure 9 [see internet version of this article] ***

Yet, Krumhansl and Kessler's data have been obtained by asking listeners to compare two sounds, namely the probe tone and the context. And when doing this, the listener in fact has to compare a memory image of the context with a memory image of the probe tone.

Hence the conjecture that the obtained results might perhaps be a mere effect of short-term memory, rather than long-term memory (Butler, 1989). At this point, adequate modeling was helpful in the discussion. If a short-term memory model would account for the probe tone rating data, then there would be an alternative explanation for the key profiles based on short-term memory, which differs from the tonal hierarchy theory that is based on long-term memory. Such a short-term model has been presented in Leman (2000). One particular useful, though very simple, feature of the model is shown in Fig.10.

*** Figure 10 [see internet version of this article] ***

The model relies on the comparison of pitch images of two echoic memories, each having slightly different half-time decays. The top figure represents a sound wave. The two middle figures represent pitch maps. A pitch image is defined as the activity of an array of neurons at a certain time instance. The maps in the figure are thus obtained by deploying the activity along two pitch image arrays in time. The difference between the two maps is due to the echo, which is short (0.1 seconds) in the upper figure, and long (1.5 seconds) in the lower figure. A longer echo has the effect that the pitch images are smeared out, so that the images reflect a context or global pitch, rather than a local pitch. By measuring the difference between the two images at each time instance, one obtains the curve in the figure at the bottom. It gives an indication of how well the local pitch image fits with the given context. The curve, in other words, represents tonal tension.

A simulation of the famous probe-tone experiments of Krumhansl & Kessler using this short-term memory model gives a good account of the probe tone ratings. The result implies that on the basis of the probe tone experiments alone, the conclusion that listeners have abstracted tonal hierarchies in a long-term memory for pitch may be unjustified, because if the short-term memory would have been insufficient, then one should have observed a much larger difference between the stimulus-driven inference of the simulation and the probe tone ratings of the experiments. Moreover, the short-term memory model is largely *adequate* (in the sense to be defined below).

The results, obviously, do not imply that long-term memory is not at all relevant for music perception. To be sure, there can be little doubt that long-term memories are necessary in tasks where recognition of a particular musical style (even tonal style), and certainly individual musical pieces and even single performances, is requested (Bigand, Poulin, Tillmann, Madurell, & D'Adamo, 2003).

3.7. Epistemological and Methodological Foundations

The above excursion into different branches of musicology suggest a common scientific program whose main features can be understood in the light of a few epistemological and methodological principles.

Epistemological questions are concerned with the acquisition of *reliable* knowledge about musical content processing. Two principles relevant in this context are gestaltism and

ecologism.

Gestaltism. Gestalt theory assumes that perception involves holistic processing. Perception, in other words, is determined by interactions between local and more global levels of perception (Leman & Schneider, 1997). In the pitch domain, for example, tonal tension is determined by the context in which pitch appears, but the pitch context is itself built up by the pitches. In rhythm perception, the beat is determined by the meter, but the meter is built up by the beat. The interactions between local and global levels is a remarkable and recurrent phenomenon in music that can be understood in terms of the different time-scales and memory systems involved during content processing.

The modern Gestalt approach subsumes the view that musical content is emerging from a number of processes, such as segregation and integration, context-dependency, and self-organized learning (Leman, 1997). It implies that musical content is more than just adaptation to the pre-established structures in the environment (Gibson's idea), and recognizes that content is also partly a construction depending on sound structures, physiological constraints, and cultural heritage.

Ecologism. The awareness that human musical content processing is taking place within an interactive framework defined by physical and cultural constraints (Fig.1), forms the basis for a theory of perception according to *ecological* principles. Any response to music, be it behavioral or physiological, should be seen as the result of a complex interaction of the subject in its social-cultural and physical environment. The notion of *embedded environment* embodies the view that a subject may pick up and store in memory particular knowledge structures that may be used for subsequent information processing. What is called *musical content* - and by extrapolation, any meaning - is defined by the interaction with the embedded environment.

In acquiring knowledge about music perception, it is important to understand the emphasis on three important features of this interaction: (i) a cyclic feature which sees interaction in terms of controlled feedback mechanisms, (ii) a multiple time-scale in which interaction is embedded within different time scales, and (iii) a evolutionary context in which the interaction of an individual subject is seen in terms of Darwinian adaptations in generations of individuals.

These interactions, obviously, should be understood as being interconnected. With regard to the long term, the information processing mechanisms, as wired in the physiology of an individual person, can be said to be formed by adaptation to the physical environment through different generations of individuals, while with respect to short term, this processing is used to establish an interaction between the individual and its environment. This interaction, in turn, may have an impact on the future development of generations, and so on. In this framework, culture is operating as the intermediate field between direct interactions between individual and the environment, and the long term genetic evolution.

This view somehow reconciles Gibson's (1966) *realism* with Fodor's (1981) *mentalism,* in that picking up the invariances offered by a structured physical reality is mediated by

constraints of the physiological processes that underlay sensory and cognitive processing. But invariances may be selected by cultural processes as well.

Patterns of information, thus structured by the embedded environment (that is: shaped by nature and by cultural evolution) can be stored in memories. As such, the structures get transformed into information content. The ecological point of view thus allows one to define musical content as the emergence of spatial-temporal forms (related to pitch, rhythm, timbre, affect, emotion and so on) that are constrained by processes at different time-scales relating to physics, human physiology, and social-cultural choices.

Carillon Culture. The carillon culture of the Low Countries, already mentioned before, provides an illustration of the complex interactions between nature and culture. Bells have been shaped since the 13th century (Lehr, 1981, 2000). The sound of bowed metallic plates could be used for music making, provided that the resonance properties of the bell were conform with certain expectations of harmonic stability. This focus on stability, rather than the alternative of focusing on instabilities, is an interesting effect of culture. It probably reveals fundamental choices in conceptions of what is essential in the world. But once the choice has been made, lawful interactions may unfold themselves.

The effort of finding a proper equilibrium between the inharmonic resonance properties of metallic plates (nature) and the harmonic foundations of music theory (culture) subsumes a cyclic interaction (Fig. 11) that is determined by sensory perception and cognitive evaluation. The cyclic interaction is itself a proper part of the embedded environment as shown in Fig. 1. In this musical action-reaction cycle, the physical environment is molded such as to give it the properties that best fit with a cultural criterion.

*** Figure 11 [see internet version of this article] ***

Cognitive evaluations are strongly influenced by societal considerations and tradition, such as the belief that stability is the optimal good, reflecting a property of divinity. In earlier times there were expert committees who evaluated the quality of the bell sounds: if the bells did not fulfill proper acoustical constraints they were destroyed, the metal purified and re-founded (Leman, 2002a). In the development of a culture, this dynamics may easily lead to a certain focal point, which can be called a cultural attractor. Harmonicity is one such example of a cultural attractor. Once it is attained, it is imposed on interactions with the physical world.

The conception of cultural dynamics in relation to the physical/physiological constraints offers a powerful paradigm for the science of musical content processing. It is a big challenge for musicology to use information processing technology in unraveling these interactions in more detail. The model of roughness, for example, deals with the sensory perception aspect of this interaction cycle. Pitch models may deal with the perception of individual bells and with bells in a musical context. Yet, cultural constraints that determine a wide range of pitch tolerance for bells are the most difficult to unravel. Nevertheless, by modeling the `hard-core' aspects, it is possible to clarify the `soft-core' aspects and hence

come closer to a better understanding of the problem.

The meaning of carillon music is related to these different aspects of content processing. When focused on sound properties it may apply to the harmonic quality. When focused on musical properties it relates to melody, harmony, and the performance of the carillon player. Its affective function may be to evoke chills when people recognize a song. When focused on social function it may reflect the pride of the city, and in economical terms its meaning is to attract tourists.

To summarize: Using modern technology, it becomes possible to develop intuition, evidence, and computation into an instrumental theory of musical content processing. Such a theory allows one to process a given input into a particular output and to explore the subtle dependencies between nature and culture. Gestalt theory should be conceived of in combination with the ecological framework. The joint combination of gestaltism and ecologism (which entails evolutionism) provides an epistemological paradigm on which a scientific theory of musical content processing is grounded.

3.8. Methodological Foundations

The methodological principles that I consider to be rather central for a science of musical content processing are naturalism, factualization, convergence, cross-culturalism, and mutual fertilization.

Naturalism embodies the belief that musical content can be studied using the methods of the natural sciences. Naturalism aims at constructing theories that explain and predict empirical data on the basis of causal principles to be described in functional mathematical terms. Naturalism, in my opinion, should be adopted for strategic reasons. The experimental methods used in psychoacoustics and music psychology can be used for the acquisition of reliable knowledge concerning musical content processing, but this does not imply, however, that naturalism is the only method by which processes of musical content should be studied. It is a reliable method, however, because it is testable and repeatable.

Factualization (Krajewski, 1977) is a process by which a theory, as abstract concept, is made more concrete and thus related to the complexities of the real world. Applied to harmony theory, for instance, factualization would imply that a theoretical construct such as *fundamental bas* can be reified in terms of pitch processing in the auditory system, and ultimately, even, to the flow of neural activity and its underlying physical properties.

*** Figure 12 [see internet version of this article] ***

The adequacy of a model (Leman, 1999a) can be evaluated in terms of a dual model (Fig.12) that relates to data from evidence-based approaches. Ideally, input and output of the model should converge with the input and output of the human subject carrying out a particular task in content processing. For example, a human subject taps a certain beat given a certain musical piece. A computer model may simulate this performance and tap

at a similar beat rate as the human subject. Such a model can be called *intelligent* when no distinction can be made between human tapping and computer tapping. Yet, this model is not necessarily an *adequate* model of the physical/physiological reality. In order to be adequate, its concept and structure should reflect the known physiological mechanisms and structures of the human processing. The model is more adequate if it satisfies deeper levels of factualization. At this point, sensory and cognitive neuromusicology may become partners in the attempt to work out adequate modeling.

In my view, adequate modeling is the most scientific and long-lasting approach because it contributes to a *theory* of content processing rather than the development of a *tool*. A tool is just something that can be practically useful, while a theory offers a view on how to understand the phenomenon. In practical applications, however, one may consider a trade-off between adequacy (and hence an ultimate epistemological justification) and efficiency.

Convergence points to the correspondence between input and output, even when different research strategies are used. In combination with the method of factualization, one may analyze the adequacy of a model in terms of global and local convergence. *Global* convergence is just the input/output convergence at the most global level. For instance, music is given as input, and the machine produces the beat as output. Global convergence points to a correspondence between the input-output relationships in humans and machines. *Local* convergence splits the model into subparts that may have their proper convergence properties. For instance, the signal is split up in different sub-bands similar to the sub-band decomposition of the human auditory system. Bandwidths correspond with physiological zones of interacting activities, but there is no simulation of the physical properties of basilar membrane resonance. In that case, further factualization might be needed to go into the physical properties of basilar membrane resonance.

Convergence always implies correlations between different research strategies. Convergent correlations between behavioral data, data from computer modeling, and data from brain research have a high epistemological relevance because they show that different research methods lead to similar results.

A major task of musicology is to figure out whether the found regularities of the evidence-based approach are real universals or merely stylistic regularities. Universal regularities should be found in all cultures, while stylistic regularities are highly context dependent. Yet research in other cultures poses its own problems. Cross-cultural understanding depends on acquaintance with tradition and with the societal aspects of a culture (Schneider, 1997b).

It's a real challenge for musicology to try to open up the traditional cultural viewpoint to naturalistic approaches, and vice versa, to connect the naturalistic approach to the cultural approach. As a matter of fact, mutual fertilization of different approaches is not only interesting, but it becomes a real necessity in modern musicology. The use of one research domain to improve or enhance the results of another domain is a very important principle that I like to relate to the idea of emergence of scientific knowledge due to local

interactions between subparts of a discipline. It has been mentioned above that the intuitive-speculative and empirical-experimental orientation contribute to the computational-modeling orientation. The latter in turn, however, may serve to sharpen hypothesis formation and to design novel experiments. Mutual fertilization presupposes a certain maturity of the research domains. The information processing technology became mature in the late 20th century, and given the powerful computers of today, modeling and simulation can provide useful feedback to the intuitive-speculative and empirical-experimental orientations. Needless to say, there is no straightforward automated interaction between the different research orientations. Much depends on the creativity of the researchers and the willingness to take into consideration the results and approaches of neighboring research areas.

3.9. Task of Musicology

The task of musicology, as I see it, is to proceed in working out an instrumental theory of perception-based music analysis. Such a theory will be based on the understanding of content processing mechanisms at different levels, including different time-scales and different memories. Understanding musical meaning relates to the employment of content in different contexts (creation, production, distribution, consumption). I see meaning in music as the outcome of a network of relationships between objective/syntactical, emotional/affective and motoric experiences of musical content within a particular socio-cultural setting. A theory of meaning should take into account the global picture, as sketched in Fig. 1, whereas an instrumental theory of content processing will be methodologically confined to study mechanisms of sensory, perceptual and cognitive processing. Again, I would stress that a bottom-up strategy in our understanding of musical meaning provides a strong working paradigm and an alternative to postmodern subjectivism.

Summarizing, I would characterize musicology in the following terms:

- It uses instruments to look at content and meaning-formation. Instruments, such as sound analysis tools or tools for monitoring human behavior of brain activity, are extensions of the human senses. They are needed because we deal with microscopic levels of musical phenomena.

- It looks for universals, that is: rules and regularities in musical content processing.

- It adopts a naturalistic approach, that is: techniques from natural sciences (e.g. statistics) are used to uncover the dependencies that underlay musical content processing.

- Its focus is on perception-based modeling and it seeks to be predictive using these models. To be predictive is the ultimate goal, but this is not always possible.

- Its models should be adequate, according to the principles of factualization and

convergence.

- It adopts a non-cognitivistic research strategy for pragmatic reasons. That means that it tries to explain cultural phenomena from nature in a bottom-up fashion, as much as possible.

- The results it produces should be accumulative, reliable and testable.

- Progress in musicology can only be made within interdisciplinary groups of researchers

-

3.10. Applicative Framework

With the workspace of musical content processing as a point of departure, this section addresses the issue of applications, first in general terms, then in terms of two concrete examples.

The interconnectedness of culture, commerce, and technology suggests a distinction between different types of contexts where musical content processing applications may be useful. It makes sense, I believe, to distinguish between: (i) creation, (ii) reification, (iii) consumption, and (iv) manufacturing. The creation context is the context in which art and music are made. The reification context is concerned with ways in which music is transformed into a product and, ultimately, a commodity. It is about recording studios, but also content description, documenting, and archiving. The consumption context is concerned with how reified musical content is brought into contact with a population of consumers. Distribution plays an important role in it. Finally, the manufacturing context provides the infrastructure for the creative, reification and consumption contexts.

Context of Creation. The study of the relationship between musicology and composition processes was initiated by Laske in the 1970s and 1980s (Laske, 1977). More recently, Rowe (2001) contributed to an integration of music cognition research into compositional systems by offering concrete tools for use in a compositional system environment. Such tools aim at exploiting evidence-based results to constrain environments for automated composition. Knowledge of the limits and possibilities of human perception and cognition can be useful in constraining the unlimited (but not always musically interesting) possibilities offered by automated composition systems. Technologies such as plug-ins/host or client/server applications offer the means through which musical content processing tools may be distributed and connected with existing software so that a personalized computational environment can be built.

Context of Reification. Reification is about making a thing out of an abstract issue. Applied to music it is about the processes that ultimately transform sound into a commodity. It involves the recording of performed music and its storage on a medium. Music can be handled as a commodity once it is put onto a concrete medium. The medium may include a textual description of the content of the music, the performers, but also a description of

the acoustical material, as well as information about the reification process itself. There is a need for content processing tools that automate this process.

Context of Consumption. The context of consumption is about the supply and demand of musical content by different users. Access to sound sources of all kinds, including speech and music are very useful for researchers, for TV-production houses (e.g. finding appropriate musical sequences, or spoken or sung film/video fragments), as well as for individual consumers who want to get direct access to particular forms of music. The capabilities of such systems can be greatly enhanced by more sophisticated tools for musical audio-mining. The musical audio-mining technology, based on content processing theory, provides a technology for the development of new types of indexing and searching in multimedia archives, thus giving more easy access to larger user groups (Leman, 2002b). In connection with a film archive, for example, users may be interested in retrieving film fragments in which similar musical sequences have been used. Or, in an archive of ethnic music, one may be interested in finding music that resembles a given melody.

Context of Manufacture. Manufacturing provides consumers with the necessary technical equipment for music consumption, as well as music production. With the advent of terabyte mass storage hard disks, one's music collection will no longer consist of a library of CDs, but rather of a library of audio on hard disk. Users will need to access this library using musical content. As to musical instruments, there is a growing market for manufactured software plug-ins and client/server related services for music production. As mentioned before, tools for musical content processing will become increasingly important in this area as well.

3.11. Application Examples

Given this general framework, I address two concrete examples of applications of musical content processing. The first example is about a system for interactive dance, music, and video, to be used for artistic performances in a theater environment. The second example is about a musical audio-mining project that aims at extracting musical content in an automatic way.

MEGA: Multisensory Expressive Gesture Applications. Until recently, most interactive music systems were built from the point of view of hyper-instruments or cyber-instruments (Hamman, 1999; Paradiso, 1999). A hyper-instrument can be conceived of as an extended version of the traditional musical instrument. The difference between the two is that the traditional instrument has its control and sound generation parts tightly coupled with a solid physical interface, whereas the hyper-instruments uncouple the input representation (physical manipulation) from the output representation (physical auditory and visual stimuli of a performance) (Ungvary & Vertegaal, 2000). The connection between input and output is based on controllers and control mappings which allow great freedom with respect to input/output controls, yet most instrument designers try to overcome mismatches of information flow through cognitive constraints imposed on the

instrument's user modalities. Ungvary and Vertegaal (2000) conceive of cyber-instruments not in the sense of the traditional dichotomy between musician and instrument, but in terms of a holistic human-machine concept based on an assembly of modular input/output devices and musical software that are mapped and arranged according to essential human musical content processing capabilities. This concept of cyber-instrument comes close to the concept of virtual and multi-modal environments (Camurri & Leman, 1997) that aim at introducing more *intelligence* and *flexibility*, so that action sequences or gestures, as well as forms of affect and emotion, can be dealt with.

Unfortunately, the current state-of-the-art suffers from a serious lack of advanced content processing capabilities in the cognitive, as well as in the affective/emotive, and motorical domains. Although advances have been made in the processing of musical pitch, timbre, texture, and rhythm, for example, most models operate on low-level features. The same can be said with respect to movement gestures. It remains hard to characterize a musical object, or to specify the typical characteristics of a movement gesture. If these systems have to interact in an intelligent and spontaneous way with users, then it is evident that their communication capabilities should rely on a set of advanced musical and gestural content processing tools. There are indeed many occasions where users may want to interact in a spontaneous and even expressive way with these systems, using descriptions of perceived qualities, or making expressive movements. Progress in the domain of content processing, therefore, is urgently needed.

Interactive systems for dance, music, and video performance explore the human-machine interaction concept within an artistic setting with content processing as a central issue. From a technological point of view, these multi-modal interactive systems rely on advanced software engineering approaches because they deal with movement analysis, audio analysis and synthesis, video synthesis. etc., in a real-time environment, utilizing the possibility of distributed processing in a networked environment. As mentioned above, the conceptual issues are challenging and such projects can only succeed if they rely on an interdisciplinary research group with a strong bias towards content processing.

The MEGA-IST project, running from 2001 to 2003 is an example of a European project having a focus on (M)ultisensory (E)xpressive (G)esture (A)pplications (www.megaproject.org). MEGA is focused on the exploration of expressive gesture applications in theater applications. The objectives of MEGA are (i) to model and communicate affective, emotive, and expressive content in non-verbal interaction by multi-sensory interfaces, (ii) to design novel paradigms of interaction analysis, synthesis and communication of expressive content in networked environments, (iii) to implement the MEGA System Environment (MEGASE) for cultural and artistic applications, and (iv) to develop of artistic public events showing the usefulness, flexibility and power of the system. Figure 13 show an example of a dancer from whom the body movements are captured by a video recorder and then analyzed in terms of expressiveness parameters. These parameters are then used to drive the synthesis of a musical performance.

*** Figure 13 [see internet version of this article] ***

MAMI: Musical Audio-Mining. The MAMI-IWT project (www.ipem.UGent.be/MAMI) (2002-2005), is an example of an audio-mining project at Ghent University that is based upon the notion of a perception-based instrumental theory of musical content processing.

*** Figure 14 [see internet version of this article] ***

Figure 14 shows the general architecture of an audio-based music retrieval system. It consists of a target audio-database (left), a query interface (right), and a similarity-matching engine with optional parts that account for users' profiling. The task is to retrieve the wanted musical audio files using information provided by the query.

Music content descriptors provide ways to access the information by means of queries. In that perspective, the *query-by-humming* is the most popular paradigm. The idea is that users can retrieve a musical piece by singing or humming in a microphone, or by playing a melody as input and asking for similar melodies as output. The system then analyzes the input and finds the intended piece. Humming or singing may be exchanged or completed with specifications at different levels of the description. For example, one may search for music that is expressive, has an energetic rhythm, and contains no human voice. These are three variables specified at the concept level, but the request involving them may constrain lower-level variables, such that an appropriate connection with the database becomes possible. Similarity measurement techniques are then used to find an appropriate match between the content extracted (and stored) from the available musical audio in the database, given the particular query.

3.12. Conclusion

The information society offers a number of challenging opportunities for musicology. One such opportunity is the development of a well-founded theory to transform musical signals into musical content and meaningful entities. This endeavor grew out of ancient interests in acoustics and musical practice. Initially, these disciplines developed as separate research areas, but gradually musicology became engaged with bridging the gap between acoustics and practice on the basis of information processing psychology and related computational modeling. This implied a broadening of the initial orientation in cultural and intuitive-speculative orientation, towards the inclusion of empirical-experimental (in the late 19th, and beginning of the 20th Century), computational-modeling (in the late 20th Century), and biology (nowadays: the early 21th Century).

The epistemological and methodological basis for musicology is grounded in intuition, evidence-based research, information processing technology, biology and, above all, their interdisciplinary integration. The central idea is that musical universals are grounded in nature and biological evolution, and that musical content and meaning, as phenomenon, emerges within a cultural context. Both nature and culture should be taken into account.

Musicology, with its focus on the development of a perception-based instrumental theory of content processing, may play a central role in the musical culture of the information society. In this society, music can no longer be considered a manifestation of pure artistic

qualities, but rather as a manifestation of qualities that are technologically embedded and contained within an enlarged social environment. A theory of musical content processing can be tested in applications such as musical audio-mining and interactive multimedia where the usage of content in terms of meaning formation can be studied. The feedback from these applications to theory formation has become an essential part of the adopted methodology in modern musicological research.

Figure Captions

[The figures are shown in the internet version of this article]
Figure 1. Schematic representation of the embedded environment in which the study of musical meaning formation is situated. The small circle represents an individual subject as the result of a long-term evolutionary process, interacting with a physical environment using sensory, perceptual and cognitive capabilities. This interaction is connected with another interaction cycle by which the subject engages in culture and society.

Figure 2. Roughness curve obtained from two harmonic tone complexes, each containing six partials. One tone complex with 6 harmonics is kept constant at a fundamental frequency of 250 Hz. The other tone complex is shifted, so that the fundamental goes from 250 Hz to 500 Hz, i.e. one octave. At each time instance, the roughness is calculated. The peaks show points with low roughness.

Figure 3. The tone complex consists of the combination of a bonang and a harmonic sound. By keeping this tone complex constant, and gradually shifting a copy over two octaves, one obtains the scale steps as minima of a roughness curve (Sethares, 1997, p. 184).

Figure 4. The tone complex consists of the combination of a saron and a harmonic sound. By keeping this tone complex constant, and gradually shifting a copy over two octaves, one obtains the scale steps as minima of a roughness curve (Sethares, 1997, p. 185).

Figure5. Results of tapping experiments compared to a resonance curve that models the data (Van Noorden & Moelants, 1999). The gray (fluent) line is the prediction of the model, the black line is the sum of tempi found in six separate style categories (dance, 16th Century polyphony, baroque, romantic, jazz, hitparade music), the dotted line is the distribution of tempi found in radio listening experiments.

Figure 6. Affect descriptors are mapped onto a three factors (Leman et al., 2003). The first factor (Valence) contains affect adjectives with positive and negative connotations. The second factor (Activity) contains adjectives that measure calm and dynamics. The third factor (Interest) contains adjectives that probe the subjective interest such as exciting or boring.

Figure 7. The individual semantics of eight subjects can be related with syntactical cues (harmony, roughness, tempo, etc.) and general trends can be observed. In this figure, the affect attributions of 8 subjects have been projected onto the Valence-Activity-Interest space of Fig. 6. Then for each factor and each subject, syntactical cues are found that best account for the data. When looking at trends among subjects, it appears that Activity correlates with loudness, while Valence is accounted for by harmonicity and tempo. With a high tempo and harmonic sound, chances are good that people will perceive this music as pleasant.

Figure 8. Structure of keys (major keys in capital letters) on a torus: left and right sides connect, as well as bottom and top sides. The map is topographical in that perceptually proximate keys are located close to each other. Major and minor keys form a circle-of-fifth structure on the map.

Figure 9. Top: Neural network with torus structure as in Fig. 8. The ordering has been obtained by self-organization of pitch images obtained from an auditory model processing musical signals. The network was trained with 72 chord-cadences, using Shepard-timbres for tones. The gray values represent the amount of activation of the neurons to a musical signal in the key of C-major. Bottom: same network but seen from the viewpoint of clustering. The figure shows boundaries between the neurons such that the clusters become visible.

Figure 10. Auditory-based pitch images of a musical signal (top panel) are smeared over time using different half-decay values. Local pitch images (second panel from top) use a 0.1 second value, global pitch images (third panel) 1.5 seconds. The global pitch images provide a context for the meaning of the local pitch images and their comparison results in a value that represents the degree of tonal tension (bottom panel).

Figure 11. The action-reaction cycle is about interaction with the physical environment. When an instrument is played (Play), physical vibrations are generated that are received and processed in the auditory system (Listen). Then a judgment is made based on physiological and cultural constraints (Judge). Then the subject can act and change the physical conditions to optimize the judgment (Act), play the instrument (Play), etc. The whole process, though constrained by nature and physiology, can be greatly influenced by cultural constraints, such as searching for purity in tone.

Figure 12. Dual validation model. The sound stimulus is the input to the auditory system where processes build images of different kinds (e.g. local and global pitch images, Fig. 10). Inferences can be drawn from these images (such as the tension curve, Fig. 10). To validate the model, inferences should be compared with human behavioral responses. To make the model adequate of human information processing, modules should be compared with human physiology.

Figure 13. Snapshot of the EyesWeb system that captures body movement and extracts information related to the affective, emotive, expressive field. This information can be used to steer the synthesis of music.

Figure 14. Architecture of a music retrieval system, with target audio-database (left), and query interface (right). A similarity-matching engine with optional parts accounts for user profiling.

References

Barker, A. (Ed.) (1989). *Greek Musical Writings - Volume II: Harmonic and Acoustic Theory.* New York: Cambridge University Press.

Bell, D. (1973). *The Coming of Post-Industrial Society.* New York: Basic Books.

Bigand, E., Poulin, B., Tillmann, B., Madurell, F., & D'Adamo, D. (2003). Sensory Versus Cognitive Components in Harmonic Priming. *J Exp Psychol Hum Percept Perform*, *29(1)*:159-71.

Broeckx, J. (1981). *Muziek, Ratio en Affect: Over de Wisselwerking van Rationeel Denken en Affectief Beleven bij Voortbrengst en Ontvangst van Muziek.* Antwerpen: Metropolis.

Burns, E. & Ward, W. (1982). Intervals, Scales, and Tuning. In D. Deutsch (Ed.), *The Psychology of Music.* New York: Academic Press.

Butler, D. (1989). Describing the Perception of Tonality in Music: A Critique of the Tonal Hierarchy Theory and a Proposal for a Theory of Intervallic Rivalry. *Music Perception*, *6*, 219–241.

Camurri, A., De Poli, G., Leman, M., & Volpe, G. (2001). A Multi-layered Conceptual Framework for Expressive Gesture Applications. In *Proceedings of the Workshop on Current Research Directions in Computer Music*, Barcelona, nov 15-16-17. Pompea Fabra University, Barcelona: Audiovisual Institute.

Camurri, A. & Leman, M. (1997). Gestalt-based Composition and Performance in Multimodal Environments. In Leman, M. (Ed.), *Music, Gestalt, and Computing - Studies in Cognitive and Systematic Musicology* (p. 495-508). Berlin, Heidelberg: Springer-Verlag.

Cohen, H. (1984). *Quantifying Music: The Science of Music at the First Stage of the Scientific Revolution, 1580-1650*. Dordrecht: Reidel.

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

Desain, P., & Windsor, W. (2000). *Rhythm Perception and Production*. Lisse, The Netherlands: Swets & Zeitlinger.

Deutsch, D. (Ed.). (1982). *The Psychology of Music.* New York: Academic Press.

Eerola, T., Jarvinen, T., Louhivuori, J., & Toiviainen, P. (2001). Statistical Features and Perceived Similarity of Folk Melodies. *Music Perception, 18 (3),* 275-296.

Essens, P., & Povel, D.-J. (1985). Metrical and Nonmetrical Representations of Temporal

Patterns. *Perception and Psychophysics, 37 (1),* 1-7.

Faltin, P., & Reinecke, H. (Eds.). (1973). *Musik und Verstehen: Aufsätze zur semiotischen Theorie, Ästhetik und Soziologie der musikalischen Rezeption.* Cologne: A. Volk, H. Gerig.

Feldtkeller, R., & Zwicker, E. (1956). *Das Ohr als Nachrichtenempfänger*. Stuttgart: S. Hirzel.

Fodor, J. (1981). *Representations.* Cambridge, MA: The MIT Press.

Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The Psychology of Music.* New York, NY: Academic Press.

Gabrielsson, A., & Juslin, P. (2003). Emotional Expression in Music. In R. J. Davidson, H. Goldsmith, & K. Scherer (Eds.), *Handbook of Affective Sciences*. New York: Oxford University Press.

Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton-Mifflin.

Godøy, R., & Jørgensen, H. (Eds.). (2001). *Musical Imagery*. Lisse, The Netherlands: Swets & Zeitlinger.

Hamman, M. (1999). From Symbol to Semiotic: Representation, Signification, and the Composition of Music Interaction. *Journal of New Music Research, 28 (2)*, 90-104.

Hobsbawn, E. (1962). *The Age of Revolution 1789-1848*. London: Weidenfeld and Nicholson.

Houtsma, A. & Tholen, H. (1987). A Perceptual Evaluation. *Music Perception, 4*, 255-266.

Huron, D. & Parncutt, R. (1993). An Improved Model of Tonality Perception Incorporating Pitch Salience and Echoic Memory. *Psychomusicology, 12*, 152–169.

IFPI (2000). *The Recording Industry in Numbers 2000*. IFPI, London.

Janata, P., Birk, J., Van Horn, J., Leman, M., Tillmann, B., & Bharucha, J. (2002). The Cortical Topography of Cognitive Structures Underlying Western Tonal Music. *Science, 298*, 2167-2170.

Kameoka, A. & Kuriyagawa, M. (1969). Consonance theory Part I: Consonance of Dyads. *The Journal of the Acoustical Society of America, 45,* 1451-1459.

Karljalainen, M., Tolonen, T., Valimaki, V., Erkut, C., Laurson, M., & Hiipakka, J. (2001). An Overview of New Techniques and Effects in Model-based Sound Synthesis. *Journal of New Music Research, 30* (3). 203-212.

Koyré, A. (1957). *From the Closed World to the Infinite Universe*. Baltimore: John Hopkins Press.

Krajewski, W. (1977). *Correspondence Principle and Growth of Science*. Dordrecht, Holland: D.Reidel.

Krumhansl, C. (1990). *Cognitive Foundations of Musical Pitch.* New York, NY: Oxford University Press.

Krumhansl, C. & Kessler, E. (1982). Tracing the Dynamic Changes in Perceived TonalOorganization in a Spatial Representation of Musical Keys. *Psychological Review, 89,* 334-368.

Krumhansl, C. & Shepard, R. (1979). Quantification of the Hierarchy of Tonal Functions within a Diatonic Context. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 579-594.

Kunst, J. (1976). Making Sense in Music I: The Use of Mathematical Logic. *Interface – Journal of New Music Research, 5,* 3–68.

Laing, D. (1996). The Economic Importance of Music in the European Union. *EMO*, Brussels.

Langner, G. (1992). Periodicity Coding in the Auditory System. *Hearing Research, 60*, 115–142.

Langner, G. (1997). Temporal Processing of Pitch in the Auditory System. *Journal of New Music Research, 26*, 116-132.

Laske, O. (1977). *Music, Memory and Thought*. Ann Arbor, MI: University of Microfilms International.

Lehr, A. (1981). *Van Paardebel tot Speelklok – de Geschiedenis van de Klokgietkunst in de Lage Landen.* Zaltbommel, The Netherlands: Europese Bibliotheek.

Lehr, A. (2000). *Geschiedenis van Wetenschappelijk Onderzoek in de Campanologie.* Asten: Nationaal Beiaardmuseum.

Leman, M. (1990). Emergent Properties of Tonality Functions by Self-organization. *Interface – Journal of New Music Research, 19,* 85–106.

Leman, M. (1995). *Music and Schema theory: Cognitive Foundations of Systematic Musicology.* Berlin, Heidelberg: Springer-Verlag.

Leman, M. (Ed.). (1997). *Music, Gestalt, and Computing - Studies in Cognitive and*

Systematic Musicology. Berlin, Heidelberg: Springer-Verlag.

Leman, M. (1999a). Adequacy Criteria and Models of Musical Cognition. In J. Tabor (Ed.), *Otto Laske: Navigating New Musical Horizons* (p. 93-120). Westport, CT: Greenwood Publ. Comp.

Leman, M. (1999b). Naturalistic Approaches to Musical Semiotics and the Study of Causal Musical Signification. In I. Zannos (Ed.), *Music and Signs – Semiotic and Cognitive Studies in Music* (p.11-38). Bratislava: ASKO Art & Science.

Leman, M. (2000). An Auditory Model of the Role of Short-term Memory in Probe-tone Ratings. *Music Perception*, 17 (4), 481-509.

Leman, M. (2002a). The 18th-century Dumery-Carillon of Brugge: Manuscripts at the Central Library of Ghent University. In *Proceedings of the 16th Meeting of the FWO Research Society on Foundations of Music Research*. Ghent University, Belgium: IPEM.

Leman, M. (2002b). Musical Audio-Mining. In J.M. Meij (Ed.), *Dealing with the Data Flood: Mining Data, Text and Multimedia.* The Hague, Netherlands: STT Netherlands Study Centre for Technology Trends.

Leman, M. & Carreras, F. (1997). Schema and Gestalt: Testing the Hypothesis of Psychoneural Isomorphism by Computer Simulation. In Leman, M. (Ed.), *Music, Gestalt, and Computing - Studies in Cognitive and Systematic Musicology* (p. 144-168). Berlin, Heidelberg: Springer- Verlag.

Leman, M., Lesaffre, M., & Tanghe, K. (2001). A Toolbox for Perception-based Music Analysis (http://www.ipem.ugent.be/toolbox). Ghent: IPEM - Dept. of Musicology, Ghent University.

Leman, M. & Schneider, A. (1997). Origin and Nature of Cognitive and Systematic Musicology: An Introduction. In Leman, M. (Ed.), *Music, Gestalt, and Computing - Studies in Cognitive and Systematic Musicology* (p. 13-29). Berlin, Heidelberg: Springer-Verlag.

Leman, M., Vermeulen, V., De Voogdt, L., Taelman, J., Moelants, D., & Lesaffre, M. (2003). Correlation of Gestural Musical Audio Cues and Perceived Expressive Qualities. In A. Camurri & G. Volpe (Eds.) *Gesture-based Communication in Human-Computer Interaction,* Berlin, Heidelberg: Springer-Verlag.

Longuet-Higgins, C. (1987). Mental Processes: Studies in Cognitive Science. Cambridge, MA: The MIT Press.

Meddis, R. & Hewitt, M. (1991). Virtual Pitch and Phase Sensitivity of a Computer Model of the Auditory Periphery I: Pitch identification. The *Journal of the Acoustical Society of America, 89 (6),* 2866, 2894.

Meyer-Eppler, W. (1949). *Elektrische Klangerzeugung*. Bonn: Dümmler.

Moles, A. (1952). *Physique et Technique du Bruit*. Paris: Dunod.

Monelle, R. (1992). *Linguistics and Semiotics in Music*. Harwood Academic Publishers.

Moore, B. (1989). *An Introduction to the Psychology of Hearing*. London: Academic Press.

Panksepp, J. & Bernatzky, G. (2002). Emotional Sounds and the Brain: The Neuro-affective Foundations of Musical Appreciation. *Behavioral Processes, 60 (2),* 133-155.

Paradiso, J. (1999). The Brain Opera Technology: New Instruments and Gestural Sensors for Musical Interaction and Performance. *Journal of New Music Research, 28 (2),* 130-149.

Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms *Music Perception, 11,* 409-464.

Parncutt, R. (1997). A Model of the Perceptual Root(s) of a Chord Accounting for Voicing and Prevailing Tonality. In Leman, M. (Ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology.* Berlin, Heidelberg: Springer-Verlag.

Pichevin, A. (1997). *Le Disque à l' Heure d' Internet : l' Industrie de la Musique et les Nouvelles Technologies de Diffusion.* Paris: L'Harmattan.

Plomp, R. (1976). *Aspects of Tone Sensation*. London: Academic Press.

Plomp, R. & Levelt, W. (1966). Perception of Tonal Consonance. In M. Bouman (Ed.), *Studies in Perception*. Soesterberg: Institute for Perception.

Posner, R. & Reinecke, H. P. (Eds.). (1977). *Zeichenprozesse: Semiotische Forschung in der Einzelwissenschaften.* Wiesbaden: Akademische Verlagsgesellschaft Athenaion.

Pressnitzer, D. (1999). *Perception de Rugosité Psychoacoustique: d'un Attribut Élémentaire de l'Audition à l'Écoute Musicale.* Unpublished doctoral dissertation, Université Paris 6.

Rameau, J.-P. (1722/1965). *Traité de l' Harmonie*. New York: Broude Brothers.

Rayleigh, J. (1877/1976). *The Theory of Sound*. New York: Dover.

Risset, J. & Matthews, M. (1969). Analysis of musical-instrument tones. *Physics Today, 22 (2),* 23-30.

Roads, C., De Poli, G., & Pope, S. (1997). *Musical Signal Processing*. Lisse, The Netherlands: Swets & Zeitlinger.

Rösing, H. (1969). *Die Bedeutung der Klangfarbe in traditionaller und elektronischer Musik – Eine sonagraphische Untersuchung*. München: Musikverlag Emil Katzbichler.

Rowe, R. (2001). Machine Musicianship. Cambridge, MA: MIT Press.

Schaeffer, P. (1966). *Traité des Objets Musicaux: Essai Interdisciplines.* Paris: Editions du Seuil.

Sabbe, H. (1996). *"All that music!" – een Antropologie van de Westerse Muziekcultuur.* Leuven: AKKO.

Scherer, K. & Zentner, M. (2001). Emotional Effects of Music: Production Rules. In P. Juslin & J. Sloboda (Eds.), *Music and Emotion - Theory and Research* (p. 361-392). New York: Oxford University Press.

Schneider, A. (1997a). "Verschmelzung", Tonal Fusion, and Consonance: Carl Stumpf Revisited. In Leman, M. (Ed.), *Music, Gestalt, and Computing – Studies in Cognitive and Systematic Musicology*. Berlin, Heidelberg: Springer-Verlag.

Schneider, A. (Ed.). (1997b). *Tonhöhe, Skala, Klang – Akustische, Tonometrische und Psychoakustische Studien auf Vergleichender Grundlage*. Bonn: Verlag für Systematische Musikwissenschaft.

Serra, M.-H. (1997). Introducing the Phase Vocoder. In C. Roads, G. De Poli, S. Pope, & A. Piccialli (Eds.), *Musical Signal Processing*. Lisse, The Netherlands: Swets & Zeitlinger.

Sethares, W. (1997). *Tuning, Timbre, Spectrum, Scale*. Berlin, Heidelberg: Springer-Verlag.

Stefani, G. (Ed.) (1975). *Actes du 1er Congrès International de Sémiotique Musicale - Belgrado.* Pesaro: Centro di Iniziativa Culturale.

Stumpf, C. (1883/90). *Tonpsychologie.* Leipzig: Hirzel. (two volumes).

Terhardt, E. (1974). Pitch, Consonance, and Harmony. *The Journal of the Acoustical Society of America, 55*, 1061–1069.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Pitch of Complex Signals according to Virtual-pitch Theory: Tests, Examples, and Predictions. *The Journal of the Acoustical Society of America, 71 (3),* 671-628.

Tervaniemi, M. & Leman, M. (Eds.). (1999). *Cognitive Neuromusicology*. Lisse, The Netherlands: Swets & Zeitlinger. (Special issue of *Journal of New Music Research*)

Toga, A. & Mazziotta, J. (Eds.). (1996). *Brain Mapping*. London: Academic Press.

Ungvary, T. & Vertegaal, R. (2000). Designing Musical Cyberinstruments with Body and Soul in Mind. *Journal of New Music Research, 29 (3),* 245-255.

van Noorden, L. & Moelants, D. (1999). Resonance in the Perception of Musical Pulse. *Journal of New Music Research, 28 (1)*, 43-66.

Veitl, A. (1997). *Politiques de la Musique Contemporaine – le Compositeur, la "Recherche Musicale" et l' État en France de 1958 à 1991*. Paris: Editions L'Harmattan.

von Helmholtz, H. (1863/1968). *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik.* Hildesheim: Georg Olms Verlagsbuchhandlung.

Vos, P. & Leman, M. (Eds.). (2000). *Tonality Induction*. Berkeley, CA: University of California Press. (Special Issue of *Music Perception*)

Winckel, F. (1960). *Vues Nouvelles sur le Monde des Sons*. Paris: Dunod. (French translation)

Winograd, T. (1968). Linguistics and the Computer Analysis of Tonal Harmony. *Journal of Music Theory, 12 (1)*, 2-49.

Zannos, I. (Ed.). (1999). *Music and Signs – Semiotic and Cognitive Studies in Music.* Bratislava: ASKO Art & Science.

Zatorre, R. (1997). Cerebral Correlates of Human Auditory Processing: Perception of Speech and Musical Sounds. In J. Syka (Ed.), *Acoustical Signal Processing in the Central Auditory System (The Language of Science)* (p. 453-468). New York, N.Y.: Plenum Press.

Zatorre, R., Halpern, A., Perry, D., Meyer, E., & Evans, A. (1996). Hearing in the Minds's Ear: A PET Investigation of Musical Imagery and Perception. *Journal of Cognitive Neuroscience, 8,* 29-46.

Zatorre, R. & Krumhansl, C. (2002). Mental Models and Musical Minds. *Science, 298* (5601), 2138-2139.

Zatorre, R. & Peretz, I. (Eds.). (2001). *The Biological Foundations of Music*. New York: The New York Academy of Sciences.

Zölzer, U. (Ed.). (2002). *DAFx: Digital Audio Effects*. New York: John Wiley and Sons, Inc.

Zwicker, E. & Fastl, H. (1990). *Psychoacoustics: Facts and Models*. Berlin, Heidelberg: Springer-Verlag.