

**Frederik Styns.**

**Research proposal: Annotation of musical Audio.**

Recent developments in the interdisciplinary research areas of Music Information Retrieval (MIR) and Musical Audio Mining (MAMi) do require the necessity to start investigating for a fundamental theoretical, methodological and practical music-analytical research of the annotation of musical audio. The annotation of musical audio concerns the detailed mapping, via specific methods and techniques, of musical Content.

This type of research is directly linked to a number of different projects currently in progress at the Institute of Psychoacustica and Electronic Music (IPEM) at the Ghent University in Belgium. In this research, I want to develop a new theoretical model based on these projects, supported by the “viewpoints” of evolutionary music science, as well as the mimetical theories of music representation.

This research project must lead to new annotation methods for musical audio, new techniques to bring musical contents in chart and create models of mimetical learning that are closely linked to the computerised approach in MIR and MAMi (for example Hidden Markov Models , Neural Networks).

The basic problem of music annotation concerns the definition of “musical content “. Musical content can generally be described as that type of sound that makes people become aware of certain information streams that are perceived as musical entities. However, the musical content situates itself on different abstraction levels, which are on itself carrier of specific concepts through which “musical content“ becomes an extremely complicated issue. Low-level concepts are referring to the acoustical and sensorial features of the musical signal. Medium-level concepts form the perception level in which time – space transformations do divide the musical signal into spatial terms (timbre, key, chords) and temporary terms (beat, metre, rhythmic patterns ...). High-level concepts are concerning the learning processes and categorisations which go further than the representation of the *now*.

The ultimate consideration prescribing “musical content“, is to find a link between the different abstraction levels, and by doing so, try to discover liaisons between, for example, emotions that are raised whilst listening to specific music on one side and the sound spectra of repeating frequencies on the other side. At the same time, considerations have to be made for the social-cultural, economical, ethical and normative context in applying music annotation. There is, in other words, a teleological (with a clear goal) and a cultural aspect connected to music annotation. It is therefore that these problematic issues also are incorporated in the discussion concerning the application of naturalistic and cultural opinions in music science.

In order to arrive at a representation of the musical content it is essential to also include the concept of imitative learning as well as the mimesis theory (1, 2). Recent discoveries in brain research are supporting the idea of a common representation for observation and synthesis (of music) by means of so called “mirror neurones“. This approach also allows to establish a link with computational models of imitative learning which are currently used in “voice recognition programs“ (Hidden Markov Models, Neural Networks ). All this will lead to what Leman (3) calls the “Instrumental Theory of Musical Content Processing”.

After having established an adequate theory regarding the meaning of the concept musical content, the development of different annotation methods and techniques has to follow. During this process the bottom-up methodology (nature – determined -from signal to concept) will be combined with the top-down methodology (culture –determined- specific targeted concepts to signal) to bring together the different abstraction levels of the musical content .

The targeted results concern a theoretical and methodological foundation of an instrumental theory of music content processing, a model that is based on the concept of imitative learning and a model that allows the automatic prescription of the musical content by the computational models – this in close corporation with applied sciences.

[1] Godoy, R. I. (2001). Imagined Action, Excitation, and Resonance. In R. I. Godoy & H. Jorgensen (Eds.), *Musical Imagery* (pp. 237-250). Lisse: Swetz & Zeitlinger B. V. [2] Inamura, T., Nakamura, Y., Toshima, I., Hideaki, E. (2003). An Integrated Model of Imitation Learning and Symbol Emergence based on Mimesis Theory. *Journal of the Robotics Society of Japan* 22, (4), (In Press). [3] Leman, M. (2003). Foundations of musicology as a content processing science. *Journal of Music and Meaning* 1, (1).