11 Expressing Coherence of Musical Perception in Formal Logic

Marc Leman

Abstract. Formal logic can be used for expressing certain aspects of musical coherence. In this paper, a framework is developed which aims at linking expressions in the formal language to an underlying interpretation in terms of musical images and image transformations. Such an interpretation characterizes truth within a framework of spatio-temporal representations and perception-based musical information processing. The framework provides a way for defining a semantics for the coherence of musical perception.

11.1 Introduction

Cognitive musicology aims at understanding the nature of musical information processing and imagination in composing, listening and performing. Of particular relevance is the choice of a proper description system.

Early cognitivist approaches of the seventies took formal (propositional and predicate) logic as the formalism for musical information processing [7,8]. Later on, and based on linguistic inspired ideas and models (see e.g. [14] for an overview), mathematical logics, extended formal logical systems (temporal, modal, non-monotonic, ...), and derived forms such as grammars, semantic networks, and Petri nets have been used as descriptive tools for musical conceptualization and subsequent reasoning [2]. What these approaches aim to achieve is the description of musical objects in connection to human information processing. In modeling perception and music analysis, assumptions often fit with the ontology of classical music theory (which is much *note* or *score* oriented).

Meanwhile, alternatives have been formulated on *ecological* and *naturalistic* grounds [9]. They ground human musical information processing on dynamic interactions with a musical sound environment, and model the physical and physiological processes using tools from the natural sciences. The so-called *functional equivalence models* give an account of the causal physical/physiological processes underlying musical behavior. (Examples of such equivalence models are dynamic systems, neural networks, cellular automata, and physical models.)

This paper is concerned with the question whether it is possible to connect the logical approach to ecological/naturalistic models, in particular, whether it is possible to express the coherence of musical perception in a formal

logic, by adopting a semantics grounded in functional equivalence models. Such a linking could provide possible new ways of dealing with relationships between musical concepts in computational systems.

A basic motivation for developing a logic of musical coherence pertains to its possible role in grasping in a concise way the inherent compelling forces of human musical perception in computational musical information processing and associated musical content extraction and data mining [1]. For example, may we deduce from the occurrence of a clarinet sound that its timbre is nasal? Given a sequence of chords, may we conclude that the addition of a new chord will create tension release? A proper logic of musical coherence may furthermore contribute to creative applications in composition, performance, and interactive computer music systems [4,3,5,15].

In what follows, we first discuss the role of formal logics in expressing musical coherence. Then we introduce a framework for describing musical coherence in terms of musical images, spatio-temporal memories, and transformation principles. A final section explores the idea of interpreting the truth of formal expressions of musical coherence in terms of a mapping onto spatio-temporal representations and transformations.

11.2 Formal Logical Accounts of Musical Coherence

An early representative example is the musical logic, developed by Jos Kunst (1978), intended to represent the *dynamics* of a musical conceptualization. We take it as a representative example of a logical approach. The logic takes modalities, such as *necessary* ($\Box$), and *possible* ($\Diamond$) as operations on musical propositions ($p, q, \dots$) in the temporal domain. The propositions stand for any kind of conceptualization of heard music. $\Box p$, for example, would mean that ' p is always with respect to past and present', and $\Diamond p$ would mean ' p is at some time in the past or present'. The modal (temporal) logic allows for the description of a changing conceptualization. For example, when some new sonoric object is encountered while 'listening' to music, the conceptualization of the music may change, and what formerly was conceptualized as $\Box p$ may change into $\Diamond \Box p$. The sudden change in conceptualization is part of the formal description. The so-called bi-valence function describes processes of conceptual *un-learning* and *learning* situation. For example, when a concept p is attributed to a sonoric object at a starting time t_0 , and this conceptualization is maintained until t_1 , then $\Box p$ will be instantiated. However, at a certain moment, owing to a new sonoric situation and related conceptualization, the expression $\Box p$ may no longer hold true due to $\neg p$, hence at a next time step, the expression $\neg \Box p \wedge \Diamond \Box p$ is derived. It means that $\Box p$ is not longer valid and that at some time in the past we had $\Box p$. The *learning* part, then, introduces a new concept, say q , and the derivation of a new law $\Box q$. The conceptualization and subsequent unlearning of p is conceived of in terms of a time line which is distinct from the time line in which q is conceptualized.

The (non-monotonic) bivalence function $BivF(p, q)$ provides a means to cross both time lines. It expresses the change of conceptualization as an unlearning and learning process.

The example illustrates some strong as well as weak points of a formal mathematical logical approach. One of the strong points is that the propositions may relate to any kind of musical concept, from general concepts such as *I hear a Sonata composed by Pierre Boulez, to Clarinet sound!*, or *Piece in C major*, up to *I like this rough percussion section*, to simply *Pitch C*. Musical content and relationships between content can thus be described at a *meta-level*. *Clarinet sound* is furthermore a very concise representation of an object in the external world whose properties as physical object and auditory image representation may be harder to define. The conceptualization and its associated reasoning based on general laws thus provides a very concise means of representation and reasoning.

The logical approach is appealing for the representation of musical imagery, a mental act in which a sonoric object or process is represented as a concept in a person's mind. Many composers acknowledge the use of metaphorical descriptions and general conceptualization before they may come to a more concrete auditory imagery of the imagined object of process. Hence, also for composition, the formal logical approach is relevant.

Yet any abstract formal description of music has a problematic character as well. The expression *Clarinet sound*, formally expressed as p , has nothing in common with an actual sounding clarinet, nor with an imagined sounding clarinet. p merely points to a content, which is furthermore not represented in the language. The classical formal logic furthermore does not specify a criterion for testing the truth of p , it merely specifies a semantics through a mapping on truth-values leaving open the way in which p is actually evaluated. In a similar way, the logical constants ($\vee, \wedge, \supset$, etc...) do not connect content but to symbols (= pointers to contents), and their mapping on truth-values is purely formally defined, in terms of the truth-values of the symbols.

11.3 Reasoning about Musical Coherence

Reasoning about musical coherence, in computational terms, involves knowledge about the meaning of the coherence. For example, if we say that "a clarinet sound has a nasal timbre", then we don't simply say something about the clarinet sound, and about the nasal timbre, but also about the intrinsic connection that a clarinet sound *perceptually* implies a nasal timbre. Moreover, if "clarinet sound" is given as a concept, we should be able to deduce from it "nasal timbre", provided that we can rely on some knowledge which says that "clarinet sound *perceptually* implies nasal timbre". Such a deduction may be realized through non-explicit, dispositional knowledge that is provided by a perceptual model which can test the truth of the connection

between "clarinet sound" and "nasal timbre". Provided that such a connection is found, then the deduction from "clarinet sound" to "nasal timbre" is accepted to be a valid one.

A possible way to proceed is to define the semantics of musical perceptual coherence in terms of interpretations within a mathematical model (a functional equivalence model of physical and physiological causality) that handles content and transformations between content. The truth of a statement expressing musical coherence then depends on its realization within that model.

11.4 Characterizing Coherence

Musical perceptual coherence is constrained by the *embedded physical environment*. It is the environment in which musical information processing takes place. It comprises the world of sounds and the constraints imposed by cultural conditions which allow us to characterize certain sounds (or sound sequences) as music.

The type of constraints that a logic of musical perceptual coherence should be able to grasp are those that define meaning when dealing with music perception and cognition. Meaning, however, is grounded in the way in which acoustical signals are dealt with through human sensory, perceptive and cognitive information processing principles and associated memories. A logic of musical coherence, therefore, has to deal with information processing. A particular example of coherence relates to aspects of tension creation and tension relaxation. Tension effects are typically based on the comparison of sound images that listeners build up in a memory. The coherence of tension effects then follows from the image transformation principles which model causal neuronal information processing.

It seems straightforward to define coherence as a logical concept for underlying causal interactions. In particular, transformation of sound into auditory images may be considered as a type of coherent transformation because *causal possible* transformation. As far as images are concerned, transformation of images into other images rely on auditory principles whose properties are grounded on causal principles (flow of neuronal *activation*). These transformations are called *coherent* when speaking about them.

11.5 Representations of Musical Content

In defining a semantics of musical perceptual coherence it is necessary to specify how musical content can be formally dealt with. Our strategy is indeed that an expression involving musical coherence is interpreted through a formal description of its content. This is achieved through the notion of *musical image*. Musical images are conceived of as spatio-temporal representational entities whose dynamics can be described in terms of image processing functions within different memory pools. In our approach, these entities provide

a link between sound and its conceptualization. Below are listed some main features and observations concerning musical images.

- Definition of Images

Images are conceived of as the result of feature extraction processes that ultimately rely upon sound waves. Images, therefore, reflect properties of the sound environment in a representational form carried by neurons. The activation of neurons is registered in terms of numerical values that account for the image information. The neural patterns that carry images may be described in different ways, occupying a defined space and time. Neural patterns occur as spike-trains, but in our modeling approach, a firing *rate-code* can be a useful encoding format for images. It allows the representation of images as vectors (see below). The rate-code refers to the amount of neuronal firing per time interval. Information can furthermore be time-coded as in the auditory periphery and then transformed into a space-encoded format. Transformation from time-code to space-code may be an important aspect of the musical logic.

- Image Processing

It is possible to describe image processing in the auditory system independently from the actual physical (neuronal) carriers in terms of signal processing such as filtering, integration, and correlation. The signal processing mathematics accounts for physiological plausible, coherent, and often causally justified image transformations, including an initial sound to image transformation. Higher-brain processing dynamics can be described by other functional equivalent models, such as cellular automata, self-organizing networks, or just by image processing principles.

- Flux of Musical Sensations

Low-level musical features, such as loudness, pitch, rhythm, and timbre are no fixed conceptualizations of sonoric objects. This is a problem for any formal logical approach to music. Instead, these musical features are to be understood as emerging primary features of underlying auditory processing. In the present model, musical images are based on the processing of signals in the time-domain, taking particular features of neuronal signal transmission in the auditory periphery such as the rate-code as a starting point. In this concept, the neural synchronization to temporal micro-structures in the neuronal patterns provides a foundation for dealing with the flux of musical sensations and hence of images as different as loudness, sensory dissonance, roughness, pitch, content-dependent pitch, density, volume, timbre, etc ...

- Different Time-Scales and Memories

Images and associated processing provide a domain for logical reasoning. But it should be noted, however, that images can be related to different time-scales of signal processing. Images at the periphery of the auditory

system are assumed to change faster than at the center, where we may conceive of images in terms of memory structures formed at long term. We may also note the use of different memory systems in which images are operating. Memories are containers of images. The echoic memory systems accumulates sensory information and uses it for local contextual information processing [10], while the long-term memory may store invariant information into a sort of statistical memory system [11]. Episodic memory may hold salient episodes of musical information.

- **Factualization**

Different time scales and memories suggests that a account of musical images should incorporate the notion of factualization, or the idea that images may have different degrees of abstraction from a concrete sound. Ultimately, images may connect to entities that can operate as abstract symbols in a formal logical reasoning process.

The many different types and appearances of spatio-temporal representational entities, and their particular emergent character suggest that a semantics of musical coherence has to rely on an integrated framework of image transformation operators.

To summarize thus far: A spatio-temporal representational system for musical images provides a formal model for dealing with the sub-symbolic description of musical content. Its specification as a functional equivalence model for causal processing guarantees coherence. In establishing a connection between the symbolic level and the sub-symbolic level we have to deal with the following objects:

- extra-logical and logical symbols
- images and image transformations
- vectors and operations

The symbols are objects at the formal logical level. The extra-logical symbols point to sounds, images, or image-transformation principles, while logical symbols specify logical connections between symbols. Images are objects that describe a state-of-the-art at the neural processing level. They reflect properties of sound objects that occur both in space and in time. A next level is the level at which images and image transformation are implemented. This level is vector-oriented. The vectors representing images are the objects of numerical operations that have a functional equivalence with causal processing.

11.6 Implementation

A spatio-temporal representational system for musical images has been implemented in a Matlab-based environment currently under development, which provides an analysis tool for musical content using auditory-based musical signal processing. The toolbox comprises a description at three different levels: functional-logical, signal processing, and implementation.

- The *functional-logical* level provides a description of image transformation concepts in terms of a functional formal logic. It allows a concise description which resembles the way in which the Matlab modules are to be used while programming in Matlab. Thus far, however, reasoning has not been implemented.
- The *signal processing* level provides a description of the Matlab functions in terms of the mathematics of signal processing where a trade-off exists between psychoacoustical accurateness and processing efficiency, hence the notion of functional equivalence model which assumes that different degrees of functional equivalence may exist, e.g. models of the inner ear can be based on filters, or on physical modeling of resonance.
- The implementation level concerns the way in which the function is actually implemented in Matlab. Signal processing functions of the Matlab signal processing toolbox have been used where possible.

11.7 Image-transformations

Having set the general framework, we now give a few examples of images and image-transformation principles and their expression in a functional logical language.

Fig. 11.1. Chart of image transformation modules, organized along the distinction between sensory, perceptual, and cognitive information processing (horizontal axis), and from textural to structural (vertical axis). The chart is not exhaustive

Figure 11.1 gives an overview of modules as conceived from the viewpoint of a distinction between sensory, perceptual, and cognitive processing. This chart, showing some processing modules as organized from sensory to cognitive (horizontal axis) and from textural to more structural (vertical axis), is not exhaustive. The modules are depicted as boxes, the associated indices as text, such as "roughness", "centroid", etc ... Important for the current purpose is that each module, which stands for a signal/image function or operator, has one or more associated image types from which indices are derived. In what follows, an example is given of a signal to image operator in the pitch domain. In the next section, its possible use within a formal logical reasoning system is discussed.

11.7.1 The "Primary" Auditory Images

The auditory peripheral module (APM) is a basic module in that it transforms sounds into a basic format on which all other images are based. The APM can be described as a function which transforms a musical sound signal $s(t)$ into a set of patterns $e_n(t)$ (with $n = 1 \dots 40$) that encode the responses of an array of auditory neuronal fibers spread along the basilar membrane (see (11.1)). The auditory filters, sometimes also called auditory *channels* thus perform an analysis of the sound in frequency sub-bands according to a critical band scale. They encode the signals as a spike-rate in time and frequency (time rate-code, and space rate-code). The rate-code amounts to the probability of neuronal firing during a short interval of time (in our implementation the interval is 0.4 ms). Space and time refer to the viewpoint: the *space code* considers a pattern in terms of an assembly of neurons, while the *time code* concerns the pattern of a single neuronal entity (such as a neuron or nerve fiber). The frequency range covered by the space code in terms of the center frequencies of the auditory filters is from 140 Hz to 8877 Hz but the highest frequency encoded in rate code is about 1250 Hz.

$$\text{APM} : s(t) \rightarrow \begin{bmatrix} e_1(t) \\ e_2(t) \\ \dots \\ e_{40}(t) \end{bmatrix} = \tilde{e}(t)$$

or, alternatively, we can write

$$\text{APM} : s(t) \rightarrow \tilde{e}(t) \quad (11.1)$$

or

$$\tilde{e}(t) = \text{APM}(s(t)) \quad (11.2)$$

The whole pattern, denoted $\tilde{e}(t)$, is called the *primary image* of $s(t)$. It should be considered the auditory counterpart of D. Marr's (1982) *primal*

sketch in the domain of visual perception. The tilde character denotes a vector and each vector-component corresponds to one auditory filter. Hence the tile character points to space code. A *running* vector means that the values of the vector-components change over time, where (t) denotes time. The pattern $\tilde{e}(t)$ thus represents the activation of a group of neurons at the level of the auditory nerve over subsequent time steps. All further information processing is based on these images, hence the connotation *primary image*.

The APM operator is effectively implemented in terms of signal processing functions that aim at simulating an auditory system at the periphery. We don't go deeper into the description of this module at the signal processing level, in order to have a more detailed focus at a next level, which relates to pitch.

11.7.2 Pitch Processing

Pitch processing can be captured by a formal logical description starting from the primary images. Starting from a neural pattern $e_{cf}(t)$, where cf represents an auditory channel, we may describe the particular transformation as:

$$e_{cf}(t) \rightarrow \tilde{p}_{cf}(t)$$

where $\tilde{p}_{cf}(t)$ is the result of a periodicity analysis on a section of that pattern. The detected periods in that section are registered along a spatial array of periods. Hence, the transformation involves a transition from images in the temporal domain (time-code) to images in the spatial domain (space-code). At each time step, one thus obtains a periodicity analysis of the primary pattern $e_{cf}(t)$ that fits the window.

A periodicity analysis is performed in all $e_{cf}(t)$ patterns, so that we can write the periodicity analysis (PA) as:

$$\text{PA} : \begin{bmatrix} e_1(t) \\ e_2(t) \\ \dots \\ e_{40}(t) \end{bmatrix} \rightarrow \begin{bmatrix} \tilde{p}_1(t) \\ \tilde{p}_2(t) \\ \dots \\ \tilde{p}_{40}(t) \end{bmatrix}$$

Secondly, a *coincidence mechanism* (CM) sums up the $\tilde{p}_{cf}(t)$ patterns and stores the result in a *summary auto-correlation* pattern called $\tilde{p}(t)$.

$$\text{CM} : \begin{bmatrix} \tilde{p}_1(t) \\ \tilde{p}_2(t) \\ \dots \\ \tilde{p}_{40}(t) \end{bmatrix} \rightarrow \tilde{p}(t)$$

The resulting pattern $\tilde{p}$ at time t is also called the *completion image* at time t because the pattern will focus on common periodicity and in a sense

complete a spectrum in function of its common sub-frequencies. The completion image is related to the notion of *virtual pitch pattern*. It gives an account of the common periodicity along the auditory neurons. Given an incomplete harmonic tone complex, its analysis in terms of a common periodicity of the auditory resolved harmonics will point to the fundamental of the harmonic complex, in a way which resembles a model proposed by Van Noorden (1982). The completion images are of interest in musical analysis, hence the function may be summarized as:

$$\text{PM} : \tilde{e}(t) \rightarrow \tilde{p}(t)$$

From the point of view of logical reasoning, care should be taken that the primary images $e(t)$ have a different dimensionality and meaning than the completion images $\tilde{p}(t)$. The space-code (as well as time-code) has a different nature and meaning.

Similar to the auditory peripheral module, a next level of detail allows the description in terms of a coherent signal processing mathematics. We do not enter into the details of signal processing here, suffice to say that the notion of periodicity can be implemented in different ways.

11.8 Perceptually Constrained Logical Reasoning

A semantics of formal logical reasoning can be set up which relates certain formal expressions to image transformations, hence providing a model theory in terms of mappings to a domain of image transformations or interpretations that characterize truth.

In that sense, the truth of a particular formal expression can be given by an interpretation in which coherence can be checked. Reasoning about musical concepts in the formal logical sense thus relies on a model which characterizes the derivations in terms of perceptually constrained transformation principles. These principles guarantee logical coherence in terms of causal transformation principles. The implementation of this idea in terms of a computational approach requires a linking between logical inference and the mathematics of signal processing.

In the logical language, the expression $p \rightarrow q$ is read as " p coherent-implies q ". It defines an operation on concepts whose truth-value is not grounded on the truth of p and q , but on an instantiation of the relation as a whole. The expression is validated (or falsified) by creating an interpretation of the expression in the domain of image transformation operators. This can be expressed by saying that in order to check the validity of $p \rightarrow q$ we must check the validity of $\exists p', q', \Phi : p' \rightarrow_{\Phi} q'$, where p' and q' are image instantiations (or interpretations) of p and q and where Φ refers to $a(n)$ (chain of) image transformation operations such as APM, PM, etc... The interpretation thus maps the expression $p \rightarrow q$ onto a statement which has a truth value.

For example, the expression "clarinet sound coherent-implies nasal timbre" is valid if nasal timbre can be obtained from an auditory-based signal analysis of the clarinet timbre. It involves a set of distinguished image transformation principles that start with the auditory peripheral module and subsequent analysis of the formant positions in derived spectral images. The truth of $p \rightarrow q$, obviously, is contingent, which means that there are possible instantiations where it is false.

Inference rules can be based on the coherence-implication, provided that on all instantiations, it cannot be false. In dealing with such an approach, we may also involve the classical logical operators (such as conjunction ($\wedge$), disjunction ($\vee$), negation ($\sim$), and implication ($\supset$), and their truth semantics in the classical way, and specify new rules of consequence for the coherent-implication operator. Such rules can be established if the expression allows any interpretation of the coherent-implication.

For example, we may add the following deduction rule:

$$p, p \rightarrow q \vdash q$$

which says that if p is true, and $p \rightarrow q$ is true, then we can conclude that q is true. For example, if we have a 75 Hz amplitude modulated sound, and a 75 Hz amplitude modulated sound coherent-implies high roughness, then we may conclude that we have high roughness. $p \rightarrow q$ can be seen as a disposition. The idea is indeed that q may be derived from p , provided that $p \rightarrow q$ holds.

In a similar way, we may say, for example, that if a 75 Hz amplitude modulated sound coherent-implies high roughness, and we don't have high roughness, then we may conclude that we don't have a 75 Hz amplitude modulated sound. Hence, the following rule could be added to the inference system:

$$p \rightarrow q, \sim q \vdash \sim p.$$

In a similar way, we may add $p \rightarrow \sim q, q \vdash \sim p$.

If a particular sound excerpt coherent-implies to be not in the key of C, and we are in the key of C, then we can conclude that we don't have this particular sound excerpt.

On the other hand, the following derivation can not be accepted:

$$\sim (p \rightarrow q) \vdash q \rightarrow p.$$

For example, if it is not true that a particular performance of a musical piece coherent-implies the expression of aggression, then we cannot derive that the expression of aggression coherent-implies that particular performance of a musical piece.

However, if $p \rightarrow q$ is true then it follows that p implies q :

$$p \rightarrow q \vdash p \supset q.$$

In other words, from the coherence between p and q we may conclude that if p is true, q cannot be false. The reverse is not accepted.

From p to deduce that $q \rightarrow p$ is not acceptable either, because $q \rightarrow p$ is checked as a whole, not solely on the true-value of p . The occurrence of a concept in the logics does not imply that this concept is coherent-implied.

Although the coherent implication is rather restricted, it can be combined as in the following general inference rules of propositional nature:

$$\begin{aligned} p \rightarrow (q \wedge r) &\vdash (p \rightarrow q) \wedge (p \rightarrow r) \\ (p \vee q) \rightarrow r &\vdash (p \rightarrow r) \vee (q \rightarrow r) \\ p \rightarrow q, p \rightarrow r &\vdash p \rightarrow (q \wedge r) \\ p \rightarrow q, r \rightarrow q &\vdash (p \vee r) \rightarrow r \\ p \rightarrow q, q \rightarrow r &\vdash p \rightarrow r \end{aligned}$$

The propositional system may furthermore be extended with temporal operators: X (= next time), G (= always), F (= at some time), U (= until), B (= before). Extensions towards modal and temporal logic may involve the following rules:

$$F(p \rightarrow q) \vdash F(p) \supset F(q).$$

If at some time p coherent-implies q then we may conclude that if p some time, then q some time.

$$G(p \rightarrow q) \vdash G(p) \supset G(q)$$

If always p coherent-implies q , then we may conclude that if always p , then always q .

$$(p \rightarrow q)Br \vdash (pBr) \supset (qBr).$$

If p coherent-implies q happens before r , we may conclude that if p happens before r , then q happens before r .

$$pB(q \rightarrow r) \vdash (pBq) \supset (pBr).$$

If p happens before q coherent-implies r , then p happens before q implies that p happens before r .

$$(p \rightarrow q)Ur \vdash (pUr) \supset (qUr).$$

If p coherent-implies q happens until r , then p happens until r implies that q happens until r .

$$pU(q \rightarrow r) \vdash (pUq) \supset (pUr).$$

If p happens until q coherent-implies r , then p happens until q implies that p happens until r .

The above examples show that the coherent-implication operator can be incorporated in a logical inference system. Its semantics can be defined in terms of a mathematical theory of perceptually constrained signal processing and inference rules can be specified which allows reasoning in a conceptual domain.

11.9 Discussion

The concern for the development of a concrete framework has forced us, thus far, to focus exclusively on *perceptually* coherence. Similar ideas, inspired by the idea that musical perception and musical action are strongly interconnected, could indeed be developed in view of *motoric* coherence (see Goday, 1997 who argues along these lines). Yet, in the current framework, such an approach should involve equivalence models of bodily movement (e.g. physical models of movement) which we don't have at our disposal at this moment. In addition, we are well aware that the present account leaves open a number of problems and difficulties. Of particular interest also is the connection with modal logic and temporal logic. The coherence-implication discussed in the present paper seems to have a lot of properties in common with the implication in a modal logics called S4. Connections with temporal logics can be clarified in view of Marsden (2000).

11.10 Conclusion

Expressing musical coherence in a formal logic can be based on the representational entities of an underlying coherence system of spatio-representational entities. The framework consists in a link between lower-level perception-driven spatio-temporal representational structures on the one hand, and high-level conceptualizations on the other hand (See Fig. 11.2). A major advantage of this approach is that it ensures coherence between conceptualization in terms of inherent compelling forces (relating to sound and sound processing), rather than imposed meta-level relationships (established by an interpreter providing meaning to symbols).

Fig. 11.2. Connection between concept, images and sonic objects

Acknowledgments

Thanks to Francesco Carreras and Jean Paul Vanbendegem for reading a first draft.

References

1. Aigrain, P. (ed.): *Musical Content Feature Extraction*. Swets & Zeitlinger 1999. (Special issue of *Journal of New Music Research*)
2. Balaban, M., Ebcioglu, K., Laske, O.: *Understanding Music with AI: Perspectives on Music Cognition*. Cambridge, MA: MIT Press 1992
3. Camurri, A.: Music Content Processing and Multimedia: Case Studies and Emerging Applications of Intelligent Interactive Systems. *Journal of New Music Research* 28(4), 351-363 (1999)
4. Camurri, A., Ferrentino, P.: Interactive Environments for Music and Multimedia. *Multimedia Systems* 7, 32-47 (1999)
5. Camurri, A., Leman, M.: AI-Based Music Signal Applications. In: Roads, C., De Poli, G., Pope, S. (eds.): *Musical Signal Processing*. Lisse, The Netherlands: Swets & Zeitlinger 1997
6. Godoy, R.: Knowledge in Music Theory by Shapes of Musical Objects and Sound-Producing Actions. In: Leman, M. (ed.): *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology*. New York: Springer 1997
7. Kunst, J.: *Making sense in music: an enquiry into the formal pragmatics of art*. Ghent: Communication and Cognition 1978
8. Laske, O.: *Music, memory and thought*. Ann Arbor, MI: University of Microfilms International 1977
9. Leman, M. (ed.): *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology*. New York: Springer 1997
10. Leman, M.: An Auditory Model of the Role of Short-Term Memory in Probe-Tone Ratings. *Music Perception* 17(4), 481-509 (2000)
11. Leman, M., Carreras, F.: 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*, pp. 144-168. Heidelberg, New York: Springer 1997
12. Marr, D.: *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman and Company (1982)
13. Marsden, A.: *Representing Musical Time: A Temporal-Logic Approach*. Lisse, The Netherlands: Swets & Zeitlinger 2000
14. Monelle, R.: *Linguistics and Semiotics in Music*. Chur: Harwood Academic Publishers 1992
15. Paradiso, J.: The Brain Opera Technology: New Instruments and Gestural Sensors for Musical Interaction and Performance. *Journal of New Music Research* 28(2), pp. 130-149 (1999)
16. van Noorden, L.: Two channel pitch perception. In: Clynes, M.: *Music, mind and brain: the neuropsychology of music*. London: Plenum Press 1982