OBJECTIVISM AND STRUCTURE-BUILDING IN MUSICAL COMPREHENSION

Erkki Huovinen
University of Turku, Finland

ABSTRACT

Background. Theories of music perception may be characterized according to the relationship that they hypothesize between some perceiver-independent description of musical events and the mental representations of the listener. Pitch-time objectivism is the view that accepts a strict correspondence relation between these two types of representations.

Aims. The aim of this paper is to show how the correspondence principle of pitch-time objectivism may be refined in order to take into account listeners' perceptual interpretations.

Main Contribution. Both pitch perception and musical time perception seem to rely on centric arrangements that the listener imposes on the perceived stimuli. Such arrangements account for a variety of phenomena in which different perceptual interpretations of a stimulus seem incompatible with each other. The mental representation of a musical event arises through a process of mental structure-building in which the centric arrangement plays an important role as a core around which the mental representation progressively develops.

Implications. Pitch-time objectivism with its set-theoretical picture of mental representations of music is judged to be inadequate. The structure-building account solves the problems of objectivism, but presents some new challenges for research methodology.

1. PITCH-TIME OBJECTIVISM

Objectivism (Lakoff, 1987), also commonly known as metaphysical realism (Putnam, 1983) is a philosophical position which subscribes to the following two claims:

(i) There is an external reality that is independent of human understanding;

(ii) Truth is a correspondence relation between linguistic or mental items and the entities of the mind-independent world.

Together these two premises yield the conclusion that there should be a single correct description of the way that the world is. If the world is mind-independently dissected into objects with their associated properties and relations, then we may say that a description of the world is correct only to the extent that it mirrors the ready-made structure of the world. I do not want to enter into the philosophical dispute concerning such objectivistic claims; my aim is to discuss a similar issue that concerns music cognition and music theory. It seems that some researchers in this area hold an implicit view that is somewhat analogous to the general philosophical position of objectivism. This music-theoretical view may be called *pitch-time objectivism*.

Pitch-time objectivism takes as its starting point the hypothesis that music (or at least some specific musical objects of interest), may be sufficiently described in terms of a two-dimensional musical space that involves a pitch dimension and a time dimension. This seems sensible, considering that traditional music notation is essentially a stylized two-dimensional graph that describes changes in pitch as a function of time. A more modern representation of musical events in this two-dimensional space is found in the "piano-roll" representations which are used by some researchers of music cognition (*e.g.*, Temperley, 2001). Following Lakoff (1987, p. 159), we may now realize that a *pitch-time representation* of this sort can be understood in a set-theoretical fashion: each note is an object that has certain properties (its locations with respect to the pitch and time dimensions) as well as certain relations with respect to the other notes. Of course, we are now talking about a theoretical description of a musical event and not about the event itself as it unfolds in the real world. Nevertheless, the success of musical praxis suggests that we may, for the purposes of argument, accept pitch-time descriptions as sufficient representations of the actual musical events. Given this, we have achieved something like a musical analogue to the first premise in the general objectivist thesis. In some sense, musical events (understood in the simplified two-dimensional way) certainly are the way they are independently of any human observers.

The second step follows when we move from the pitch-time representation of a musical object or event to the *mental representation* of the same object or event. The former is studied by the music theorist, the latter by the music psychologist. What is the relationship between them? Is there a correspondence between the mental representations of music and music as it is "objectively" represented in musical notation or in some more elaborate forms of pitch-time description? It seems that many music analysts, at least, implicitly subscribe to the following correspondence principle, which hypothesizes some sort of one-to-one mapping between the two representations:

(C1) The mental representation of a musical work corresponds to the pitch-time representation of the work.

We may call this view *naive pitch-time objectivism*. Its most obvious mistake is that it seems to transform all relationships between pitch elements in a piece of music into *heard* relationships without paying attention to the limitations of human memory. The view has, nevertheless, been rather well entrenched in the minds of music researchers, judging by the fact that recently several studies have been published with the sole purpose of demonstrating that listeners in general do *not* have a firm grasp over all music-analytically meaningful aspects of large-scale musical form (Cook, 1987; Edlund, 1997; Marvin & Brinkman, 1999). These insights lead to a restricted version of the naive thesis — one that restricts its temporal scope to shorter musical events. "Event" is here taken to mean anything that falls within the capacity of short-term memory:

ISBN 3-931852-67-9
ISSN 1617-6847

(C2) The mental representation of a musical event corresponds to the pitch-time representation of the event.

This version of the correspondence principle is related to the view that Levinson (1997) has called *concatenationism*, which states that understanding music centrally involves apprehending immediate progressions and not necessarily aural or intellectual grasp of large-scale connections in musical works. From a cognitive point of view, the new version of the principle sounds much more reasonable than the previous one. One way to develop this view is to think of the mental representation of an "atemporal" pitch configuration (such as a note or a chord) as a pattern in a conceptual space (Gärdenfors, 2000), recalling that many geometrical models of pitch are in fact multi-dimensional spaces in themselves (Shepard, 1982). A pitch-time representation then corresponds to the temporally changing patterns (of activation) within such a mental structure.

The principle (C2) differs from concatenationism in two respects. First, it still leaves open the very real possibility that large-scale relationships may *occasionally* be grasped by the listener. More importantly, however, it claims direct correspondence between what is represented by the listener and all that there is within the musical event in the pitch-time sense. This is still a rather strong claim concerning the perceptual abilities of human listeners. Common experience and studies of attentional mechanisms in music listening leave no doubt about the fact that some pitch elements may be more easily heard than others and that music listening involves a certain amount of necessary but effortless selection. This calls for the following modification to the earlier principle:

(C3) The mental representation of a musical event corresponds to a subset of the event's pitch-time representation.

This already seems to be a very useful principle. It is compatible with observations concerning figure/ground relationships in situations where musical events tend to split into several perceived streams (*e.g.*, Dowling, 1973). It is also compatible with the general music-analytical idea of pitch reduction, in which some pitch events are considered structurally more important than others. The central feature of this refined principle is that it nevertheless accepts the set-theoretical nature of the mental representation. Given that we can point out the subset of the pitch-time representation that is going to be highlighted in perception, the principle (C3) allows the music analyst a complete freedom in explaining his musical experience of the event with reference to any pitch relationships that occur within that specified subset. In terms of a connectionist model this might mean that all nodes in a neural network that represent the pitch elements of the specified subset should be connected to all other such nodes.

2. CENTRIC ARRANGEMENTS

We have progressively refined the correspondence principle by restricting it first to temporally shorter segments of music, and then to selected subsets of pitches that occur within such time spans. However, the version (C3) still has the flavor of objectivism: there is a strict correspondence between how the musical event is heard and how some subset of it really exists in a "neutral" pitch-time representation. Now, recall that the pitch-time representation is understood in a set-theoretical way: it consists of entities with properties and relations. The correspondence relation has the effect that the attended subset of pitch elements is reflected in a similar set of elements in the mental representation.

The weakness of the refined version of pitch-time objectivism lies in this set-theoretical premise. The problem is that similar sets of elements in a pitch-time representation may be given not only different but altogether incompatible perceptual *interpretations*. Such perceptual interpretations, I submit, arise through *centric arrangements* that are imposed by the listener on the mental items that correspond to the pitch elements in the attended subset of the pitch-time representation. The phenomenon of centric arrangement is so familiar that it almost seems trivial. Consider a simple ascending melodic interval. Subjective perceptual centers may enter the perception of such an interval in both of the two dimensions of the pitch-time space. This is illustrated in figure 1, where the dashed lines in horizontal and vertical orientations stand for perceptual centers, mental points of accent, in the pitch dimension and the time dimension, respectively. The two-dimensional subjective grid of perceptual centerpoints allows for four possible placements with respect to the simple ascending interval. Within the pitch dimension, the melodic movement may be experienced either as ascending *to* the perceptual center or ascending *from* it. Similarly, the movement either *begins* on a temporal center or *ends* on one (cf. trochees and iambs in lyric meters).

Our common terminology for dealing with centric arrangements derives largely from the more elaborate situation in which there is a *periodicity* of perceptual centers with respect to both dimensions. Within the temporal dimension, periodic repetition of the centerpoints is, of course, called *meter*, and the centers themselves acquire the role of *downbeats*. Within the pitch dimension, periodic repetition of the center is called *octave equivalence*, and the phenomenon of centric arrangement is discussed under the topic of tonality, the centers being called *tonal centers*. The terms "meter" and "tonality" should be understood here in a simple perceptual sense, disregarding the specific connotations that these terms are given in discussions of historical developments of musical styles and compositional techniques.

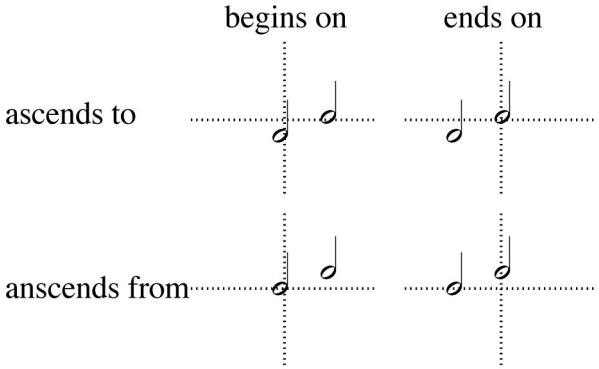

Figure 1: Four centric arrangements imposed on an ascending melodic interval.

ISBN 3-931852-67-9
ISSN 1617-6847

Elsewhere, I have shown by way of extensive experiments that the concept of tonal centricity is well applicable to the description of strategies of perceptual "grasping" even when the musical stimuli are highly unconventional or "non-tonal" (Huovinen, 2002). In rhythm perception, we find a similar tendency to interpret novel stimuli through centric arrangements: only a few clicks or beats are often enough to give us an impression of where the imaginary downbeat should be located. Indeed, it seems a reasonable assumption that centric perception is what the listeners generally strive for in order to comprehend musical events. This urges us to modify the correspondence principle accordingly:

(C4) The mental representation of a musical event corresponds to a centric arrangement of a subset of the event's pitch-time representation.

The notion of centric arrangement has the effect that the mental representation no longer exhibits a neutral, set-theoretical structure. Centricity means interpretation, and perceptual interpretations lead us away from objectivism.

The centric arrangement of a set seems to involve two things: (i) one element of the set is chosen as a perceptual center, and (ii) the relations between the center and the other elements of the set somehow become highlighted in the representation. The important point about this is that *all* relations between all elements do *not* have to be processed, which saves the listener from a combinatorial explosion when more elements come into play. It is noteworthy that some objectivistic theorists do accept the role of centricity but attempt to derive it from calculations concerning the relations between *all* elements (consider, for instance, the derivations of metrical weight by Lewin, 1981). Others attempt to reduce the idea of centricity to nothing more than a hierarchy of "stability" concerning scale-steps within an octave or temporal locations within a measure. This leads to a search for optimal fittingess between the hierarchy and the musical stimulus by taking each element in the periodic cycle in turn as the nominal center of perception (see *e.g.*, the Krumhansl-Schmuckler key-finding algorithm, see Krumhansl, 1990). In these approaches, all relations within the stimulus are given an equal status.

What is not admitted in such approaches is that some of the relations in the pitch-time representation may be *suppressed* through centric arrangements. In other words, the act of taking some relations as the direct relations of a perceptual center may render other equally prefarable relations unaccessible. Consider a simple example from tonal perception. You cannot hear the chord C–E–G–A simultaneously in C major and in A minor, and depending on your decision the perceived quality of the chord will change, because each time only one of the thirds above the tonal center (or one of the sixths below it) that determine the perceived mode will have to be processed as a direct relationship of the center (for a more elaborate discussion, see Huovinen, 2002, pp. 35-45). Similar kinds of subjective "creation" of perceptual objects through the selection of some relationships and the suppression of others also happens in the perception of rhythmic patterns. Consider, for example, situations in which the listener is unable to place the metrical downbeat in the "right" location. However hard they try, listeners with a Western musical education may sometimes not be able to adjust their perception of African rhythms to match that of the African musicians themselves (Kubik, 1983). As Bamberger and Brody (1984) suggest, even intellectual understanding of one metrical interpretation cannot make it seem compatible with another interpretation in terms of which the listener *hears* the stimulus. Centric arrangements create incompatible interpretations with respect to the attended subset of the musical stimulus.

3. STRUCTURE-BUILDING

It should be noted what the principle (C4) does not say: it does not say that there first exists a correspondence between the mental representation and some subset of the pitch-time representation, and that the centric arrangement only develops after that. Such a view would still count as pitch-time objectivism, because it allows that there is a level of mental processing at which strict correspondencies are created between the stimulus and the mind. In (C4), the correspondence relation is no more a one-to-one mapping between some fragment of the mind-independent musical event and the mental representation of the listener. The listener now has a more active role.

Elsewhere, I have suggested (Huovinen, 2002, pp. 223-235) that the phenomenon of centricity can be integrated within a more general theory that is analogous to Gernsbacher's (1990) structure-building account of language comprehension. In hearing music, listeners construct mental representations by integrating incoming information to mental structures that have been initiated during the time of hearing the previous stimuli. The leading idea in the structure-building framework is that such integration of new information to previously initiated mental structures is done until the new information becomes (subjectively) incongruent with the developing mental structure, and the listener has to "shift" to a new substructure. Shifting requires cognitive effort and it is thus avoided as much as possible by suppressing irrelevant aspects of the incoming information.

Within a musical context, the developing mental structure may now be described in terms of a centric arrangement of a subset of the event's pitch-time representation. The listener may begin the process of mental structure-building by picking up one or two elements from the stimulus (that we describe as a pitch-time representation), and imposing a centric arrangement on them. The process may continue through a progressive exploration of the stimulus, in which single elements are given representational counterparts in the developing mental structure. Thus, the centric arrangement is *not* imposed on a ready-made set-theoretical representation that has been transferred from the music to the listener's mind. Rather, centric arrangement provides the initial core of the mental structure that may subsequently continue to develop when new information is added to it.

The concept of centric arrangement reflects many of our basic musical intuitions: scale degrees and metrical positions are identified in terms of their relationships to the perceptual center — other relationships are suppressed. Also the structure-building account of musical understanding looks appealing, because it takes into account the manner in which the listener's own previous actions are partly responsible for her perceptual interpretations at any given time. However, these benefits have a high price for research methodology. Along with the set-theoretical treatment of elements in the pitch-time representation, we have to throw out a

ISBN 3-931852-67-9
ISSN 1617-6847

host of formal approaches to the study of mental representations of music. As seen above, the proposed account emphasizes the role that centric interpretations have during an early phase of perceptual processing. According to this view, then, any algorithm that does *not* severely restrict the number of alternative centric interpretations that it considers would not seem to qualify as an appropriate model of human musical processing. Furthermore, statistical results concerning the manipulation of pitch-time data in human perception should not be elevated into a theory of musical processing in "the mind", if it turns out that even simple stimuli may receive seemingly incompatible interpretations in the minds of different listeners. Consequently, an algorithm that *does* restrict the alternatives in the way that is typical of human listeners cannot claim generality for its specific results: it only mimics the actions of a single perceiver.

In short, these problems seem to encourage the divorce of music psychology from music theory: psychological theory may only concern the general form of the relationships between pitch-time representations and mental representations, but the latter cannot be reduced to the former. This is a provocative conclusion — not because it would contradict any good musician's intuitions, but because it challenges us to ask, to what extent our views concerning musical perception have been driven by our research methodology.

4. REFERENCES:

1. Bamberger, J. & Brody, M. (1984). Perceptual problem solving in music: Some proposals for future research. *Psychomusicology*, *4(1-2)*, 33-58.

2. Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception*, *5(2)*, 197-205.

3. Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, *5*, 322-337

4. Edlund, B. (1997). Tonal closure—fact and/or fiction. In A. Gabrielsson (Ed.), *Third triennial ESCOM conference. Proceedings* (pp. 140-144). Uppsala University.

5. Gärdenfors, P. (2000). *Conceptual spaces. The geometry of thought*. Cambridge, MA, and London: The MIT Press.

6. Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

7. Huovinen, E. (2002). *Pitch-class constellations. Studies in the perception of tonal centricity*. (*Acta musicologica Fennica 23*.) Turku: Finnish Musicological Society.

8. Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

9. Kubik, G. (1983). Kognitive Grundlagen afrikanischer Musik. In A. Simon (Ed.), *Musik in Afrika* (pp. 327-400). Berlin: Museum für Völkerkunde.

10. Lakoff, G. (1987). *Women, fire, and dandgerous things. What categories reveal about the mind*. Chicago and London: The University of Chicago Press.

11. Lewin, D. (1981). Some investigations into foreground rhythmic and metric patterning. In R. Browne (Ed.), *Music theory: Special topics* (pp. 101-137). New York: Academic Press.

12. Levinson, J. (1997). *Music in the moment*. Ithaca and London: Cornell University Press.

13. Marvin, E. W. & Brinkman, A. (1999). The effect of modulation and formal manipulation on perception of tonic closure by expert listeners. *Music Perception*, *16(4)*, 389-408.

14. Putnam, H. (1981). *Reason, truth, and history*. Cambridge: Cambridge University Press.

15. Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review*, *89(4)*, 305-333.

16. Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, MA, and London: The MIT Press.

ISBN 3-931852-67-9
ISSN 1617-6847