Remarks on the Cognitive Reality of Musical Representation:
A Response to Lerdahl's "On Tonal Motion and Force"

John Halle
Deptarment of Music
Yale University
john.halle@yale.edu

Talk presented at:
MIND & MUSIC ROUNDTABLE
Columbia University
March 4-5, 2006

I should confess to a small degree of discomfort in responding to Fred Lerdahl's paper in the context of a colloquium mostly attended by philosophers. The reason is that the issues and concerns which I will be addressing might better be described as foundational: that is, they should be seen as addressing matters internal to the field of cognitive music theory rather than necessarily externally relevant to philosophy. That said, when a field begins to examine its underlying premises closely, those within it often find themselves encroaching on disciplinary ground well trod by philosophers. Best known among these are foundational issues in physical sciences such as the ontological status of atomic particles, the directionality of time, the probabilistic nature of quantum reality. These are topics which have for many generations engaged the attention of both philosophers and physicists and while most physicists have agreed with Stephen Weinberg's assessment that "*no one* who has participated actively in the advance of physics in the postwar period. . . has been significantly helped by the work of philosophers (Weinberg's italics)" (Weinberg 1994) certain proposed revisions to the generally accepted historical picture (Howard 2005) have suggested that a more charitable view may be in order.

More recently, a new interdisciplinary common ground shared by philosophers and researchers in other fields has emerged with the development of the cognitive sciences. It is easy to see why. Cognitive science is, after all, a scientific investigation into our capacity to develop, exercise and represent knowledge in a variety of psychological and perceptual domains: language, vision, spatial awareness, pragmatics, problem solving, logical paradoxes, the formation of belief etc.[1] Given that it is concerned with investigating these particular forms of knowledge, cognitive science is a virtual synonym for epistemology. The only difference is arguably a superficial terminological one, residing in the fact that the latter is a theory of knowledge approached from the philosophical perspective.

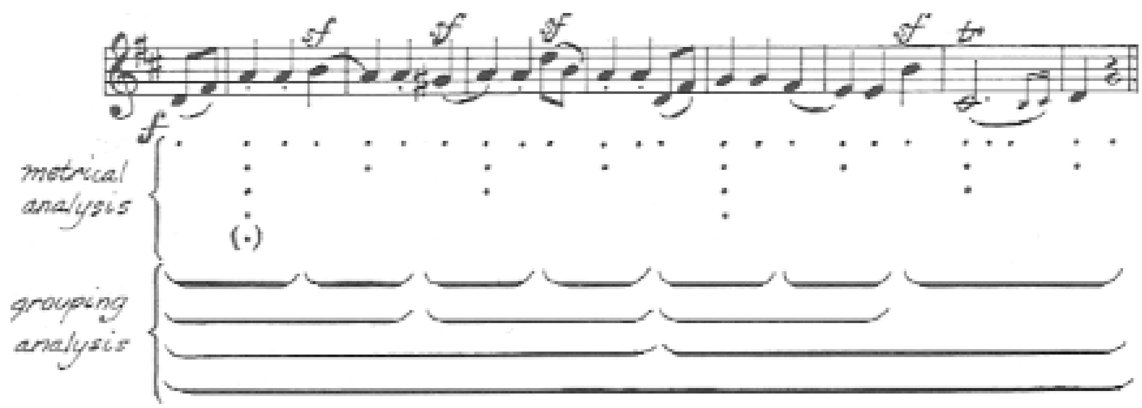
Lerdahl's presence at this conference is indicative of a recognition of this fact, that is, of the awareness that his work is directed to answering a fundamental epistemological question: namely, what is the form in which one type of knowledge is represented in our minds and ultimately in our brains. Among the questions which Lerdahl's work has addressed are what sorts of formalisms must be adopted, what types of computational machinery must be posited, what sorts of idealizations are required such that we can account for the particular class of cognitive facts which constitute musical knowledge. The category of facts implicated here are the facts of musical intuition: more or less obvious and uncontroversial shared responses of experienced listeners with respect to various aspects of musical structure. Among these are the fact that certain pieces are heard in ternary rather than binary meters, that certain harmonic sequences are heard/felt as tensing or relaxing, certain sequences of notes require the hearing of constituent (or phrase) boundaries at some locations as opposed to others. These intuitions are the basic

[1] For an overview of these and other studies see, for example, Osherson (1995)

data which a theory of musical knowledge must account for and all of Lerdahl's work and that of those working in the field constitutes an attempt to account for them.

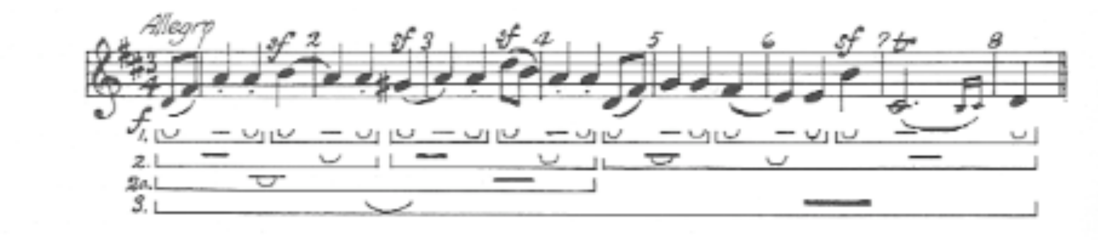
To give some idea of how Lerdahl and Jackendoff's generative theory and its extensions within Lerdahl's *Tonal Pitch Space* and elsewhere (I will refer to all of this work hereafter as GTTM) succeed in providing answers but perhaps more importantly, allow us to ask questions which we have been unable to coherently formulate previously, I will focus for a moment on a subcomponent of Lerdahl's theory of musical knowledge, namely his description of the inferred rhythmic structure of music and the formalism which must be posited in order to account for it.   After having been first advanced two decades ago, GTTM's descriptive apparatus for representing the intuitions implicated in metrical structure, namely, the grid representation shown immediately below the Haydn minuet in example 1 has gained general acceptance.

Example 1



The acceptance of this formalism, like that of any other formalism, resides crucially in its capacity to capture important characteristics of the facts under consideration thereby allowing us to view the essential mental reality of metrical intuitions obscured within other representations of inferred metrical structure. Thus, to take one such example, the prosodic representation of metrical structure proposed by Cooper and Meyer shown in Example 2 is a reasonable attempt at characterizing metrical intuitions.

Example 2



The essential shortcoming inherent in the prosodic representation, as noted in GTTM, is in its attribution of metrical prominence to portions of phrase segments. This can be seen by observing the markings within the horizontal brackets in example 2. It will be noticed that the two eighth note upbeat and the entire first measure of the passage is assigned a macron on level 2 by Cooper and Meyer. This is to indicate that this half of the bracketed time span is relatively strong compared to the second half of the bracket which, as indicated by the breve, is heard as weak. Metrical prominence is, however, understood by listeners as an inherent characteristic not of continuous segments but of discrete points in time, that is, of the temporal locations marked by what we refer to as the beat. Thus, it is the first beat of the first measure, and not the time span extending from the upbeat to the third beat of the measure which is heard as strong. The grid representation identifies the former point in time as strong by means of the five dots assigned to this precise location while the latter is identified as relatively weak by virtue of its being assigned only a single dot. Since the grid correctly represents these and other facts[2] which are essential components of our understanding of meter, it and not the prosodic representation is now broadly adopted, including among those who are not necessarily sympathetic with the cognitivist approach in explaining the metrical structure of music.

Another contribution of GTTM to our understanding of the mental representation of rhythmic structure is the recognition that the "accent" associated with the awareness of meter is of a special perceptual type, namely the metrical accent, a perception of relative prominence as an inherent characteristic of specifiable temporal positions. The metrical accent is categorically distinct from what is informally understood as an "accent" namely, the phenomenal accent, which is inherent in the physical signal itself resulting from a local extreme in amplitude, pitch, or sudden alteration in timbre. While it is natural for listeners, even experienced ones, to conflate all types of accent, they are distinct cognitive objects: not all metrically strong positions correspond with a phenomenal accents, nor are phenomenal accents enjoined from occurring on relatively weak positions of the metrical hierarchy. The phenomenal accent can function as an important cue to metrical accent however, it is only one factor within a complex computation and is not determinative of the result. While they are easily confused, any reasonable account of musical structure must distinguish between these fundamentally distinct types of accent and GTTM, while not the first to do so, was the first to provide a comprehensive and coherent account of their distinct character.

[2] For example, the grid will also be seen to be preferable in that it is able to represent the downbeat of the second measure as relatively strong compared to its neighbors but relatively weak compared to the first beat of the first measure.

A significant achievement of GTTM's approach to meter was the recognition that the informal term accent is in fact a kind of homonym, referring to two fundamentally distinct albeit interactive psychological entities. In order to make sense of metrical phenomena, GTTM made use of a gambit familiar within the natural sciences namely to devise a specialized terminology which overlaps with but which is not coextensive with informal usage. Just as the terms, work, energy, momentum, etc. have a precise, technical definition within physical sciences while evincing only a metaphorical connection to their use in such expressions as "I don't have the energy to read that book," "Joe Lieberman's campaign is picking up momentum" and "Women's work is never done," the operative definition of "accent" within the GTTM system, as has just been discussed, is one which has only an oblique relationship to how we informally understand the term. While the folk science understanding may provide us with certain types of evidence with respect to our perception of meter in some instances, there is no reason why a theoretical account of the underlying basis of metrical structure and its computation should be expected to conform to our informal sense of the term. Indeed to do so would guarantee its failure.

Insofar as GTTM is correct in the specific case of its account of metrical structure and its account of musical structure generally, knowing a piece of music means having access to the sorts of representations which are posited in the system, maybe not, indeed certainly not, exactly like them but having the broad outlines proposed therein: metrical structure designated by the grid generated by the beat induction computation outline therein or perhaps making use of proposals elsewhere[3], constituency structure according

---

[3] See, for example, Eck (2001), Desain and Honing (1999), Povel and Essens (1985).

to the computations designated by the grouping well formedness and preference rules[4], vertical/harmonic structure within the outlines designated by prolongational and time span hierarchies. To the extent to which a listener has succeeded in constructing these representations, it is in this precise sense that one can be said to "know" a piece. More pointedly, insofar as a piece resists the construction of these representations, it can be said to be fundamentally unknowable, and perhaps, hardly a piece of music at all, a claim which has generated some controversy in the two decades since it was first proposed.[5]

****

To reiterate, these and many other insights with respect to the nature of musical form as it is psychologically represented constitute a contribution to our understanding of one aspect of what human knowledge is. But if one grants this to be the case, certain contradictions and confusions along the lines just mentioned immediately come to the fore not just with respect to musical knowledge but to knowledge generally. As we saw in the case of "accent" just as the informal usage of the term obscures what is required in order to devise a descriptively adequate account of the phenomena, so too does the common usage of the term "knowledge" and the verb "to know" in many cases come into conflict with the technical sense of "knowledge" which the cognitive sciences require us to adopt.

It follows from this that it is likely that no generalized definition of the term will come close to offering an adequate account for all our uses of it. There is no such thing as knowledge but rather a menagerie of cognitive states of distinct types which we refer to

[4] Revisions to the grouping preference rules are proposed in Frankland and Cohen (2004).

[5] See Lerdahl (1988) and its discussion in volume 5 of Taruskin (2005).

by the same word but which are by no means the same thing. That this is the case seems fairly obvious. Thus, while we might say that

(a) "John knows the Eroica symphony"

and

(b) "John knows that the moon is made of cheese"

it is clear that the knowledge implicated in (a) relies on the sorts of computations and formalisms which form the basis of cognitive musical theory. As for (b), while this aspect of John's knowledge must have an ultimate cognitive basis, there is no reason to believe a priori that the sorts of primitives, formalisms and computations which are appropriate to it are at all comparable to those which are appealed to in forming the sort of a musical representation which constitutes the musical knowledge implicated in the former statement. Although we use the term knowledge to refer to John's relationship to both (a) and (b), there is no reason to assume that the term designates anything close to the same the same entity.

Examining the two sentences more carefully, whereas (b) necessarily implicates propositional knowledge on the part of John, his awareness of things in the world, their real world characteristics and their logically necessary and logically possible relationships, (a) assumes encoded psychological representations of the general variety proposed in GTTM. That there is an inherent ambiguity to the term we are employing is, of course, a common linguistic circumstance and does not necessarily pose any serious impediment to effective communication. The problem for us as investigators comes when we assume that the particular form of knowledge implicated in (b), namely propositional knowledge, is in some sense theoretically primary and on this ground must dictate the form of what we posit to make sense of our psychological capacities.

Why this matters for GTTM or, for that matter, any minimally plausible cognitive theory, is that it appears that the sorts of representations which one must posit within a cognitive theory of music are, it would seem, notable precisely for their absence of propositional content. A strong statement of this position can be found in Ray Jackendoff's essay the "Problem of Reality" (1992). According to Jackendoff:

> "it hardly makes sense to say that the representations one constructs in response to hearing a performance of the Eroica (Symphony by Beethoven) are true or false. Nor does it make sense to claim one has propositional attitudes towards musical representations, which aren't, as far as I can tell, propositions. The whole story becomes even more absurd if we think of the process of composition. Presumably, Beethoven wrote the symphony by virtue of creating mental representations (imagining music) that he wanted the written notes to evoke—in the absence of any overt musical signal. Were his representations false until he wrote the symphony down? Did they suddenly acquire a truth-value when the score was written or the first performance took place? I can't make any sense of such questions—they entirely misconstrue the character of musical experience.

Musical representations provide a specific application of a general characteristic of cognitive representations which Jackendoff describes as follows:

> "A representation is not *about* anything. It does not strictly speaking represent anything, hence my hesitation about using it except as a rigid designator of what cognitive scientists believe the mind has in it."

If mental representations generally and musical representations specifically do not embody propositional attitudes, intentional states, or intentional content what is a representation and what are its salient characteristics? According to Jackendoff for the purposes of cognitive inquiry a mental representation is "a space of possibilities having an inherent organization." Accordingly, a particular mental representation R

> can be thought of as a location or region in the total space of possibilities. What makes R what it is is simply its relationship to other possibilities in

> the space--how it is distinct from other possibilities, what possibilities are close to it in what dimensions, and so forth. When we as theorists use symbols to state a theory of mental representation, it is not the symbols themselves that are significant but rather the range of distinctions possible in the system of symbols we adopt: these distinctions are claimed to be homologous to the organization of the relevant subsystem of brain states.

It should be apparent to those who are familiar with it that Lerdahl's attempt to describe our perceptual relationship to the vertical/harmonic structure of tonal works provides a parade example of what Jackendoff has in mind in this passage and it is worth making a brief digression here to focus on certain specifics of Lerdahl's model for this point to become clear. In particular, it will be noticed that the geometrical or more specifically topological organization proposed by Lerdahl in his paper (the framework is given its complete exposition in *Tonal Pitch Space*) has an "inherent organization" of the sort Jackendoff refers to which has the potential to express a wide range of logically possible perceptual relationships to harmonic structure. But only a small subset of these organizations are reasonable to entertain as cognitive models consistent with the facts. For this reason, it is appropriate to characterize Lerdahl's investigation in Jackendoff's terms as attempting to determine which of "the possibilities in the space" are optimal as a model for how we construct tonal relationships in pieces of music.

Among these possibilities is the pitch space model which Lerdahl applies to the Chopin E major Prelude in his figure 2 which we repeat here as example 3 with, as will be seen, certain alterations:

Example 3

↑	g#	B(V)	b	D	d			
	c#(vi)	**E(I)**	e(i)	G	g			
(fifth axis)	f#	A(IV)	a	C	c	Eb		
	b	D	d	F	f	Ab	g#	
↓	e	G	g	Bb	bb	Db	c#	E

← (minor third axis) →

Example 3 fills out slightly more of the partial space represented in Lerdahl's figure 2 by identifying a few additional vertical and horizontal regions. The purpose in doing so is to show one aspect of the "inherent organization" of the space more clearly, namely, while the horizontal axis is derived from alternating minor and major key regions a minor third apart, the vertical axis maps out major or minor key regions along the circle of fifths. While this spatial orientation offers an attractive representation of the modulatory schemes of certain works, including the Chopin Prelude under consideration, the purpose of structuring the space in this way is not merely aesthetic, it is empirical. That is, it makes correct predictions with respect to the relevant empirical domain, which is, to reiterate, the listener's unconscious understanding of one component of musical structure, namely the relative distance of tonal regions.

The central intuition which must be represented within a pitch space model is the perception of relative tonal distance of key areas or tonal regions, as we will refer to them: certain regions, namely, the dominant (V), subdominant (IV), parallel minor (i) and relative minor (vi), are understood as[6] closest to the central tonic I. Others, for example, major keys a tritone or a minor second away from the tonic (Bb and F) are highly distant. As for the major third cycle of modulations characteristic of later 19th century harmonic practice and found in the Chopin prelude discussed in Lerdahl's paper, these are somewhere in between: more distant and more surprising than the closely related keys though not maximally distant and unexpected. It will be seen that Lerdahl's pitch space model succeeds in representing all of these regional distances accurately: the four most closely related keys are located immediately adjacent to the tonic I. Somewhat further away on the space are the key areas reached in the Chopin prelude. More distant still- requiring a traversal of two vertical and two horizontal steps is the bII region F major, and maximally distant within this model is the region built a tritone above the tonic Bb major which requires three vertical and two horizontal steps.

It should be obvious that not all models of pitch space which we might construct represent these and other regional distances accurately. This is what Jackendoff has in mind in referring to cognitive representations as defined by their "relationship to other possibilities in the space." The relative success of Lerdahl's pitch space model can be verified by a comparison of example 3 with the pitch space represented in example 4.

---

[6] The evidence for adjudicating the validity of the model is, on the one hand, external: as is well known, modulations were for a period of nearly a century limited to these key areas and these would remain the most common modulatory destinations within tonal music for a century thereafter. More significant, however, is internal evidence: these key areas are heard by listeners as being subjectively "closer" to the tonic than other more distantly related keys. Much of this data is introspectively obvious to anyone familiar with the tonal idiom. The experimental work of Krumhansl and others referred to in Lerdahl's paper confirms what most experienced listeners take for granted along these lines. We will return later to a discussion of the important distinction between the external and internal empirical domains.

Example 4

↑	A(IV)	Bb	B(V)	C	C#	D	Eb	E	F	F#	G	Ab
minor 2nd	bb	b	c	c#	d	d#	e	f	f#	g	g#	a
axis	B(V)	C	C#	D	Eb	**E(I)**	F	F#	G	Ab	A(IV)	Bb
	c	c#	d	d#	e	f	f#	g	g#	a	bb	b
↓	C#	D	Eb	E	F	F#	G	Ab	A(IV)	Bb	B(V)	C

← minor 2nd axis →

Rather than being constructed from horizontal third and vertical fifth axes, both axes map out key areas immediately adjacent along the chromatic scale, with the vertical axis alternating, as does Lerdahl's minor third horizontal axis, major and minor regions.

There is no a priori reason why the arrangement of tonal regions in example 4 should not describe the relevant cognitive facts; indeed, there is at least some basis for believing that a fully chromatic space, as example 4 might be called, is not altogether unreasonable as a model of how we hear tonal regions. For example, musical instruments tend to situate chromatic pitches adjacent to each other on fingerboards or keyboards. So does conventional musical notation which represents half steps as occurring on vertically adjacent staff locations (with the addition of the appropriate accidentals).

But while there is no obvious a priori reason why the space in example 4 shouldn't succeed in characterizing perceptual reality, the a posteriori fact of the matter is that it does not: we simply don't happen to hear music in that way the model indicates. To return to the previous discussion, the closely related tonic and the dominant key areas, (B and A) are located far away from the tonic E major on the space, requiring a traversal of six vertical or horizontal steps to the left or right respectively. In contrast, regions which are experienced as distant to the tonic, F and Eb major, are represented as immediately adjacent on both axes. Of the closely related keys, only the parallel minor i is represented as close to the tonic, and this is only diagonally adjacent. As for the Chopin major third modulations indicated on Lerdahl's path these appear as, on the one hand, closer in space than the dominant and subdominant regions I and IV and also more distant than the immediately adjacent chromatic regions 9 (c.f. example 3). Thus, this model of our competence clearly fails-not only due to its own empirical shortcomings, but, more significantly, in comparison "other possibilities within the space" namely Lerdahl's pitch space in example 2 which succeeds in describing the data more satisfactorily.

Now it may be the case that the broad class of spatial representations of which Lerdahl's model constitutes one possible instantiation will turn out to be, for some reason, incapable of accurately representing certain important aspects of our internalized understanding of tonal hierarchies. Indeed, other organizations have been proposed, most notably the neo-Riemanian model which has become probably the dominant analytic framework for the description of late Romantic repertoire within music theory proper.[7] This is not the place to evaluate these competing models except to remark that which representation is chosen is entirely an a posteriori empirical question. And insofar as a representation succeeds in explaining the structure which listeners actually assign to it, it is reasonable to assume that, according to Jackendoff's description, the model successfully "designates what cognitive scientists believe the mind has in it-a symbolic

[7] See, for example, Cohn (1999).

description which is homologous to the mental structures according to which (the musical experience) is ultimately instantiated."

****

The discussion above allows us to clarify two aspects of the cited remarks by Jackendoff. First, it should be clear that whatever form these representations ultimately must take they are, as Jackendoff insists, by their nature, "not *about* anything". That is, they do not refer to "things" in the external world independent of our perceptual relation to it but rather designate salient aspects of the construction of particular things-namely pieces of music. As such, the representation is not of an external musical object, but of the internal cognitive structure which the external object gives rise to in the mind of the listener.[8]

It follows that, while indeed "not *about* anything" a mental representation is very much *a thing* in any reasonable sense of the term. Specifically, it is a "natural object", in precisely the sense discussed in Chomsky's important essay (Chomsky 2000). Mental representations are natural objects of a particular sort, that is, they are theoretical constructs, like gravitational fields, atomic particles, or the null subject parameter (in linguistics), whose existence while not directly apparent to our senses, has been shown to be necessary in order to account for a range of observable empirical facts. The "occult" nature of these objects has, as Chomsky recounts, been repeatedly brought into question over the centuries. By this point in the history of science their existential status is taken as entirely unproblematic and barely discussed. Quarks, gravity, and twenty seven dimensional hyperspace (if it exists) are for the purposes of science, in the words of

---

[8] For a discussion of internalist conceptions of musical structure as compared to the the dominant structuralist perspective in music theory, see Halle (forthcoming).

Steven Weinberg, "real in pretty much the same sense (whatever that is) as the rocks in the fields" (Weinberg 1994). If the mental representations posited by Lerdahl withstand scientific scrutiny they must be accepted as just as unproblematically real.

With this perspective in mind, a second claim made by Jackendoff becomes eminently reasonable: insofar as a musical representation is a natural object (in the Chomskyan sense) "it hardly makes sense to say that the representations one constructs in response to hearing the Eroica Symphony are true or false." For while propositions such as "John knows the Eroica Symphony" may be true or false, to say that any natural object is true or false, is a semantic absurdity. One doesn't say "a chair is true", "a noble gas is false" or "a mountain range is true": one says that these objects exist (or not) depending on the evidence provided for them. The same applies to the mental structures which a listener constructs in response to hearing a piece of music. It is not a question of the truth or falsity of the propositional content of these representations, but rather of the existence (positive or negative) of the things themselves. By their nature, they are mute-incapable of expressing anything about the world aside from the fact of their own existence.

*****

To conclude, I will suggest that the confusion surrounding the status of mental representations above is part of a larger picture. The tendency to ascribe content to mental representations of cognitive objects which only grudgingly and unnaturally accept such ascriptions is one characteristic of what Jackendoff describes as the "philosophical" relationship to how our mind construes the world around us and the things in it. According to Jackendoff, the philosophical approach is concerned with answering the following question:

> "What is the relationship of the mind to the world, such that we can have knowledge of reality, such that we can have beliefs and desires about things in the world, and such that our sentences can be true and false?"

In contrast, the psychological approach while superficially related asks a fundamentally distinct question:

> "How does the brain function as a physical device such that the world seems to us the way that it does, and such that we can behave effectively in the world."

Jackendoff argues that the philosophical approach "in case after case . . . leads to uncomfortable metaphysical problems (most but not all of them well known) whereas the 'psychological' approach permits-at least in principle-a revealing account of the phenomena." While I will not develop this aspect of Jackendoff's discussion, as we begin to discover more about music and our cognitive relationship to it and how we as a species construct our knowledge of the world around us in a variety of experiential domains, it would appear that Jackendoff's assessment is becoming increasingly uncontroversial.

References

Chomsky, N. (2000) "Language as a Natural Object" in *New Horizons in the Study of Language and Mind*. Cambridge, England: Cambridge University Press.

Cohen, A. J. and B. Frankland, (2004) Parsing of Melody: Quantification and Testing of the Local Grouping Rules of Lerdahl and Jackendoff's "A Generative Theory of Tonal Music" in Music Perception 21:4 [Summer 2004] p. 499-543

Cohn, R (1997) "Neo-Riemannian Operations, Parsimonious Trichords, and their Tonnetz Representations," Journal of Music Theory 41.1, pp. 1-66.

Desain, P. & Honing, H. (2003) "The formation of rhythmic categories and metric priming." *Perception*. 32(3), 341-365.

Eck, D. "Finding downbeats with a relaxation oscillator." *Psychological Research*, 66(1):18-25, 2002.

Essens, P.J. and D.J. Povel. (1985) Metrical and nonmetrical representations of temporal patterns. *Perception and Psychophysics*, 37(1):1-7.

Halle, J. (forthcoming) "Whose post-Structuralism: A Response to Pat McCreless"

Howard, D (2005) "Albert Einstein as a Philosopher of Science," Physics Today, December 2005

Jackendoff, R., (1992) "The Problem of Reality" in *Languages of the Mind*. Cambridge: MIT Press.

Lerdahl, F. (1988) "Cognitive Constraints on Compositional Systems " , Contemporary Music Review 6:2 pp. 97-122.

Lerdahl, F. (2001) *Tonal Pitch Space*. New York: Oxford University Press, 2001.

Osherson, D The Encyclopedia of Cognitive Science, MIT Press, Cambridge, 1994.

Taruskin, R (2005) The Oxford History of Western Music, New York: Oxford University Press.