Cognitive science and music analaysis - the case of musical pitch

Ian Cross
CENTRE FOR MUSIC & SCIENCE
FACULTY OF MUSIC
UNIVERSITY OF CAMBRIDGE
WEST ROAD
CAMBRIDGE CB3 9DP
UK
E-MAIL: ic108@cam.ac.uk

Introduction

This paper is intended as an attempt to answer the question "how are the findings of cognitive science relevant to music ?". It will do so by focusing on the ways in which cognitive science has explored musical pitch organisation in the context of three different types of relations between cognitive science and music.

At the outset, one might ask why cognitive science should be employed in attempting to understand music. A general answer to that question is that the application of cognitive science to music can be thought of as being intended to bridge the gap between what music feels like - its experiential texture - and the language that is used to describe it and to teach it. To be more specific, the development of a cognitive science of music can help to span the disjunction that exists between the ways that music is experienced by listeners and by practising musicians and the rational frameworks of discourse that conventionally constitute music theory, i.e. that are used to describe and to define music. This development proceeds by seeking to provide accounts of music that are consonant with the concepts of computational logic - "computability" - and with empirically-derived evidence about musical perception, performance and creation.

A cognitive science of music should describe explicitly the manner in which processes of memory, of perception, of consciousness and of action come together within the experience of music. It should enable the empirical evaluation of theories about musical experience whether these exist at the level of intuitions about music or as explicit music theories. It may even prove capable of providing explanations - in terms of the capacities of cognising systems - of phenomena such as the consonance or dissonance of sounds or chords: the experience of sequences of pitches as lines or melodies: or the feeling that some notes in a melody are "stable" while others are "unstable".

How are the findings of cognitive science relevant to music ?

Rosner (1988) suggests that cognitive science and theories of music can be related in three distinct ways. Cognitive science can clarify the ontology of theories about music by experimentally testing their applicability to our musical perceptions and productions, in effect evaluating the musical theories. Alternatively, cognitive science can explain music theories by providing us with accounts of cognitive processes and psychophysical and neurophysiological constraints that explain our musical experience, in effect functioning to replace music theory by explaining it. On the other hand, theories of and about music may provide ways to expand the theoretical bases of cognitive science; music is, after all, a unique activity, and it would not be surprising if the cognitive processes that were revealed by the study of its perception and production advanced our understanding of our broad cognitive capacities. However, most cognitive-scientific studies of music exemplify more than one of these types of relation between music theory and cognitive science, while some employ all three types of relation (seeking to evaluate and explain music theory and to contribute to general theories of cognition from a specifically musical perspective)

Limits on explanatory power

Cognitive science has undoubted explanatory power in respect of musical experience. However, when addressing issues in music that are predicated on cultural determinants of mind and are not directed towards contemporary perceptions and productions (e.g., when cognitive science is used to account for the historical origins and development of tonal musical pitch organisation), cognitive science cannot provide "complete" scientific accounts as its empirical content is rooted in contemporary perceptions and productions. It can contribute to historical accounts but cannot supplant them. Despite this, and despite the arguments advanced by Treitler (1980) and by Cook (1989) that science should not play a role in the elucidation of aesthetic domains, cognitive science must be deemed fully capable of providing wholly scientific accounts of music as it exists now in our current perceptions and actions.

Musical pitch organisation

Studies of musical pitch organisation have both evaluated and explained music theoretic accounts of pitch organisation. They have done so by testing music-theoretic accounts of musical pitch organisation and by investigating general psychophysical or cognitive principles that underlie the ways in which pitch is organised in musical perceptions and productions. However, the degree to which different groups of studies constitute evaluations or explanations differs considerably. I shall begin by outlining a body of research which may have started by evaluating music theory and appears to have provided a new level of explanation for the phenomena that gave rise to that theory (for a more detailed account see Cross, West and Howell, 1991; Cross, 1997).

Music theory - evaluation and explanation I: sensory and psychoacoustical process as determinant of musical pitch organisation

Most conventional Western theories of musical pitch refer to one single specific "scientific fact" in describing the foundations of musical pitch organisation: the harmonic series. This "scientific fact" is usually cited (often in a very perfunctory way) as determining which pitches should "sound well" together and thus as providing a foundation for theories of harmony. However, while the harmonic series is undoubtedly a "scientific fact" there is no reason to accept that it should determine how pitch should be employed in music unless it can be shown that listeners and composers are directly and involuntarily sensitive to the relationships embodied in the physical reality of the harmonic series. In other words, if we are to accept that a scientific fact - the harmonic series of low-integer ratios between the constituent frequencies of a complex tone - determines the ways in which pitch is used in music, we must be able to demonstrate that the "facts" of musical pitch usage can be explained in terms of this scientific fact. In the nineteenth century Helmholtz explored this issue; while he found some evidence that we are sensitive to low-integer relations between frequencies, he came to the conclusion that the harmonic series could not wholly account for the ways in which pitch was employed in music.

Much recent research has come to the same conclusions in showing how it may be our sensitivities to the physical facts of musical sound rather than the physical facts themselves that determine our musical practices. This research has focused on the processes employed by our auditory systems in analysing out components of complex waveforms and integrating these into unitary percepts in order to explain the way in which chords and chordal successions are experienced and employed in music. The most theoretically sophisticated instance of this research is that conducted by Parncutt (1989), which derives from that of Terhardt (see Terhardt, Stoll and Seewan, 1982). Parncutt proposes that chords are heard as having unitary identities or roots because listeners have learned to hear complex periodic waveforms as having unitary pitches (generally the fundamental frequency of the complex waveform). When a complex periodic sound is encountered, it is usually experienced as one pitch; this occurs through a process of analysis of the waveform and its re-integration as a single percept by the auditory system. Parncutt suggests that a similar process governs the perception of chords in music. Thus, a sounding triad - major or minor - can be heard as being one thing rather than three notes sounding together. The likelihood of a chord being heard as a unitary percept - as having an unambiguous "root" - is largely dependent on a pattern-recognition process based on a "best fit" to the harmonic series. Major or minor triads would probably be heard as having unambiguous "roots", would be experienced as "stable", and would be likely to be employed in some referential way, while a configuration such as the "Tristan" chord would be "unstable" (and hence non-referential) as it could be experienced as having several different possible "roots". Parncutt has conducted a series of experiments that on the whole support his theory's predictions.

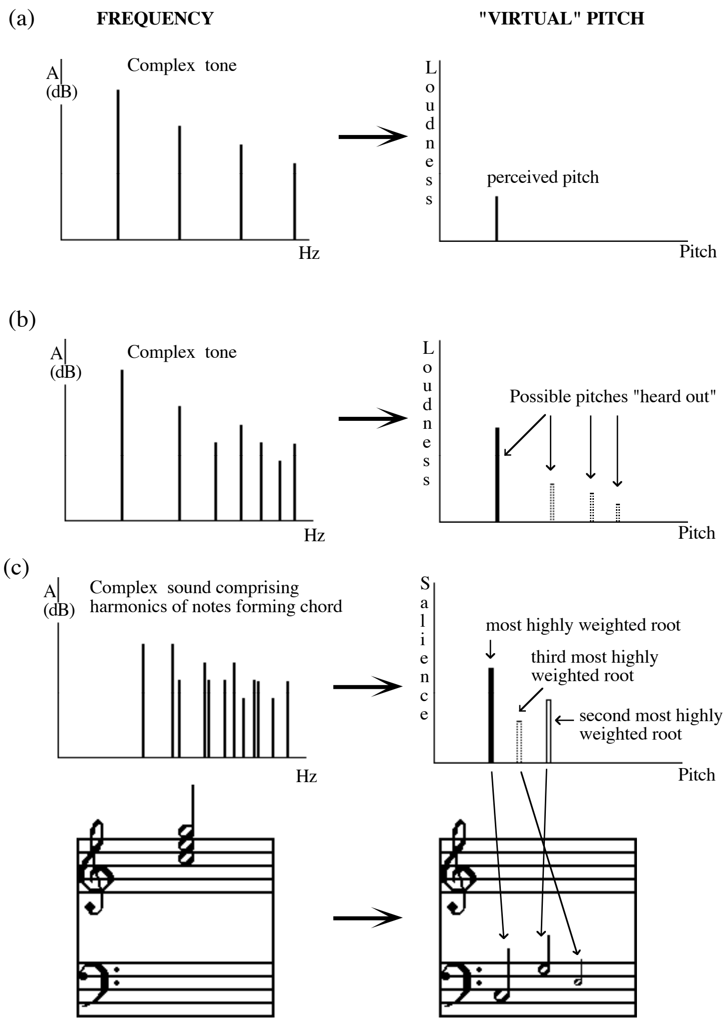

Figure 1

(a) A complex tone, comprising four harmonics, is heard synthetically as a sound with one definite pitch.

(b) A complex tone, comprising seven harmonics, is heard analytically as a sound with at least four component pitches.

(c) The complex sound-wave produced by a three-note triadic chord (comprising three complex tones) is heard as having an identity that corresponds to one of three possible "chord-roots".

A theory focusing on psychoacoustical functioning such as Parncutt's can offer explanations of how it is that we hear chords as more - or less - "stable" and closural, or final. However, such theories do not seem capable of explaining other fundamental phenomena of musical pitch such as why we should experience sequences of successive pitches as co-ordinated entities or melodies, nor why we should employ particular sets of pitches to compose such melodies.

Music theory - evaluation and explanation II: cognitive processes as determinants of musical pitch organisation - surface structure and auditory scene analysis

The question of how we come to experience a series of pitches as a melody has troubled philosophers such as Scruton (1983), who resorts to the idea that such "musical motion" is perceived "metaphorically", without ever making clear what such "metaphorical listening" might involve. Music theory is scarcely more explicit on this point, although it does prescribe that pitch intervals between successive notes of a melody should generally be small. The issue has been investigated extensively by Bregman, who developed the concept of Auditory Scene Analysis (see Bregman, 1990) in the course of conducting a wide range of empirical studies of the factors that determine our perception of connectedness between pitches as they unfold sequentially in time.

His research started by focusing on the Gestalt "laws" proposed by German psychologists of the twenties and thirties; these laws take the form of general principles that are supposed to determine how objects will be grouped in our perceptions of them. Hence the "laws" of similarity and proximity suggest that objects that are similar, or are spatially or temporally close together, will be perceived as grouped together, while the "law" of good continuation dictates that a sequence of separate objects aligned according to a common trajectory will be perceived as outlining the course of that trajectory and hence as grouped.

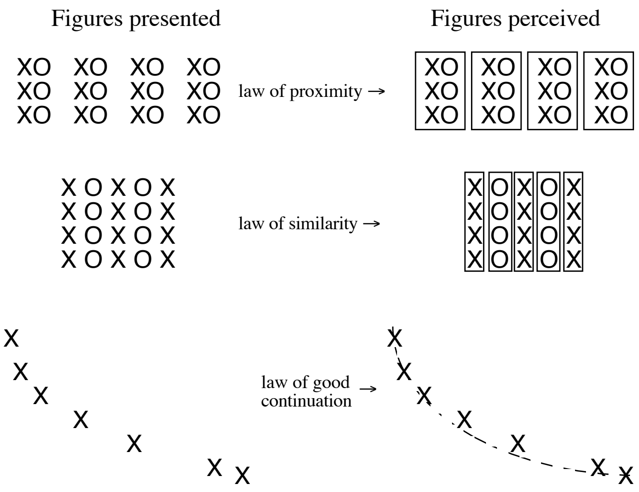

Figure 2
Examples of the visual groupings produced by the application of the laws of proximity, similarity and good continuation.

Bregman suggests instead that these "laws" represent "best guesses" about what is likely to be happening in the world around us and that they enable us to organise the information that our senses receive. He terms this process "auditory scene analysis", and states that this "is the process whereby all the auditory evidence that comes, over time, from a single environmental source is put together as a perceptual unit" (Bregman, 1993). The "evidence" from our auditory environment reaches our ears as sound waves (i.e., patterns of small changes over time of air pressure) and is transduced into patterns of neural activity by processes in the auditory periphery, these being interpreted in the light of our expectancies about what "evidence" should be available. Bregman sees the Gestalt "laws" as deriving from regularities at each of these three levels, the acoustical, the psychoacoustical and the cognitive, enabling us to infer the existence of sound-producing objects or sound-sources in our environment and to group successive sound events by assigning them to one or other of these inferred sound sources. It is easy to see how the Gestalt principles appear to be at work in our use of this evidence; the operation of the laws of similarity or proximity are implied in our grouping together of sounds from one (or neighbouring) locations and of the same or similar waveforms, and that of good continuation can be referred to our expectancies that changes in the pitch produced by a single sound source over time will be small and gradual rather than large and sudden. In the Auditory Scene Analysis model, sound sources may be "real", as in the case of two instruments in different spatial locations simultaneously producing sequences of pitches, or "virtual", as in the case of compound melody, where one instrument produces alternate pitches in different pitch ranges but gives the impression that it is producing two more-or-less simultaneous and discrete melodic lines.

Music theory - evaluation and explanation III: cognitive processes as determinants of musical pitch organisation - tonality

While Auditory Scene Analysis provides the framework for an explanation of our experience of sequences of pitches as being melodies, it cannot provide a rationale for the use of certain sets of notes and of chords in much of the music of our culture, nor for the ways in which certain notes or chords within these sets seem more important or more "stable" in our perceptions than others. The "cognitive-structuralist" research programme conducted by Krumhansl and others and described in Krumhansl (1990) has examined just these issues, taking as its starting point the ways in which these phenomena are described within conventional music theory, and resulting in a justification and explanation of that theory in cognitive terms.

The cognitive-structuralist programme is intended to examine the judgments that listeners make about pitches and relations between pitches in musical contexts, and employed simple music-theoretic concepts such as scales and keys to organise its experimental material and to define these contexts. In effect, in using concepts from music theory it is evaluating the applicability of those concepts to the description of musical experience while developing cognitive-scientific explanations for the phenomena to which those concepts had been applied.

A fundamental postulate of the cognitive-structuralist approach is that underlying our perceptions or judgments of pitches and pitch relations is some form of schema, or mental structure in long-term memory that organises the information received from our senses and is itself altered by that information (see Neisser, 1976), shaping our interpretations of what we encounter, and determining the nature of our experiences. In examining listeners' judgments about musical pitch in context, the cognitive-structuralists aim to elucidate the structure and functioning of our schemata for musical pitch and to determine the factors and processes that give rise to such schemata. The initial experiments that Krumhansl and others carried out in pursuit of these aims required listeners to rate the degree to which pitches completed or "fitted with" short melodic phrases (the "contexts"); further experiments examined the "fit" of single chords that succeeded short sequences of chords, the perceived similarity between pitches or chords following short contexts, or the time taken to discriminate between pitches or chords following contexts. The melodic or chordal contexts were intended to ensure that listeners made their judgments on the basis of the "tonal schema" that they possessed by providing appropriate contexts for the musical perception of pitches and pitch relations. In most of the cognitive-structural experiments these contexts consisted of an ascending or descending major or minor scale traversing an octave from tonic to tonic, of a major or minor triad, or of a short "key-defining" sequence of chords such as I-IV-V-I, although in a few studies more extensive contexts have been employed.

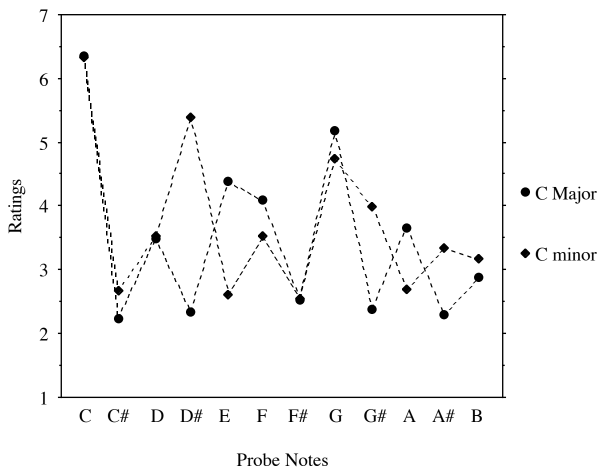

Figure 3
Typical ratings derived from the responses of a number of listeners asked to rate how well each chromatic note "fits" following a context of an ascending or descending major or minor scale.

Initial experiments asked listeners to rate how well each chromatic note within an octave range completed either seven-note ascending or descending major or minor scales or major or minor triads. It was found that notes which were diatonic in respect of the context were rated more highly than non-diatonic notes, and that, within the diatonic notes, notes were rated more highly on the basis of their relative strength or stability within a conventional tonal hierarchy (with the tonic most highly rated, followed by the dominant and the mediant, then the other diatonic notes). Further studies on the perception of chords and their inter-relations in context found that chords which could be interpreted as being on the tonic, dominant, and subdominant of a key tended to be clustered together in a multi-dimensional representation of all the listeners' responses, other diatonic chords being detached from the cluster and, chromatic chords being even more distant. These results were taken to indicate that the cognitive representation of chords and chord relations embodied some hierarchical structure, distance within the multi-dimensional representation being taken as analogous to the functional differentiation of chords within the listeners' musical pitch schemata.

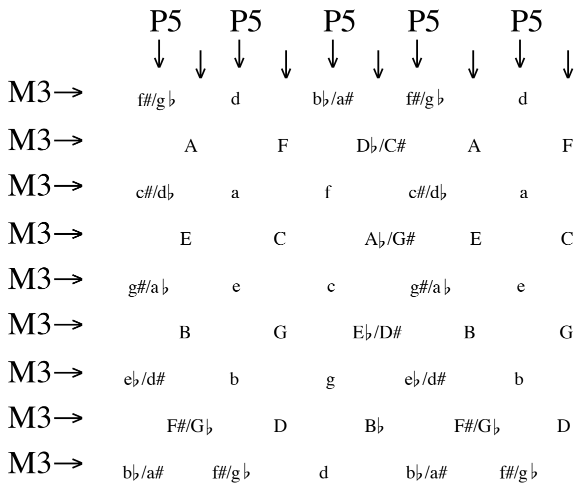

Figure 4
"Harmony space" (after Krumhansl, 1990), showing fifth-related chords adjacent along one diagonal and "minor-related" chords adjacent along the opposite diagonal (upper-case=major, lower-case=minor).

Krumhansl suggests that a principal factor in giving rise to the tonal hierarchy is the statistical distribution of pitches in tonal music. Statistical analyses of pitch distribution in various corpora of tonal music indicate that pitches that are highly rated by listeners in the cognitive-structuralist experiments - such as the tonic, dominant and mediant - tend to occur more frequently and persist for longer proportions of the total duration of tonal pieces than do other pitches. Moreover, such notes also tend to occur in salient positions in the musical texture, e.g. at the ends of phrases and sections. The high degree of correlation between tonal hierarchy rating profiles and the frequency distribution of pitches in tonal music, implies that learning - whether formal, or by acculturation - is a primary factor in enabling listeners to form schematic representations of musical pitch organisation that embody the tonal hierarchy. Empirical support for this hypothesis came in a study by Castellano, Bharucha and Krumhansl (1984), who conducted a cross-cultural experiment using North Indian music in which two groups of listeners, one having had extensive experience of that music while the other had not, were presented with short pieces of music as contexts and judged pitch “fittingness” in respect of the contexts. Analysis of the experimental data showed that all listeners had rated pitches according to the total sounding duration of pitches, but also demonstrated that only listeners with extensive prior exposure to Indian music showed a sensitivity to the hierarchy of pitches described in Indian music theory. In other words, the inexperienced listeners had employed a strategy of basing their judgments on the abstracted regularities of the pitch structure - the event hierarchy - of the music that was presented as context, while the Indian listeners had been capable of employing the representation of an appropriate tonal hierarchy that they had built up through years of exposure to or training in Indian music.

Thus tonal-hierarchical schemata are built up through exposure to bodies of pieces of music that exhibit regular and consistent event hierarchies, and are brought to bear, generally non-consciously, by listeners as they experience pieces of music unfolding in time in order to "sense" which of the pitches in a piece is more or less important or stable, stable pitches coming to serve a referential function and other pitches being perceived as being organised around these reference elements. In this way the schematic representation of the tonal hierarchy assists in the perception (and creation) of complex and multi-layered pattern in the domain of musical pitch; its actuality in cognition can be modelled computationally in the form of a neural net or connectionist model as suggested by Bharucha (1987, 1991) and more recently by Leman (1995).

An alternative account of the dynamics of pitch organisation in cognition is found in the work of Brown and Butler, who suggest (Butler 1989) that cognitive-structuralist findings might not reflect the operation of schematic representations of pitch relations in long-term memory but instead can be accounted for by local pattern implications or by short-term memory processes (with highly-rated notes being those that start, end, or simply occur within, the context sequences). Their work is based on recent theories of pitch organisation in music that treat pitches as analogous to numbers (as "pitch-classes") and focus on the set-theoretic properties exhibited by the sets of pitch-classes conventionally employed in Western music.

Within the set-theoretic representation of pitch (see Forte, 1973; Balzano, 1980), the diatonic set (the pitch-class set analogous to the diatonic scale) has several unique attributes. The property upon which Brown and Butler have concentrated is the fact that the tritone is a "rare" interval within the diatonic set, two pitches separated by a tritone plus any other pitch being sufficient to define a single diatonic set; hence of all the possible musical intervals, the tritone has the greatest capacity to signal to a listener the function of the pitches in a tonal piece. They showed (Brown and Butler, 1981) that listeners exhibited a high degree of accuracy in producing a feasible tonic when presented with a three-note context consisting of two notes separated by a tritone plus one other note (a "minimal cue-cell"). Further studies (e.g., Brown, 1988) indicate that certain orderings of the minimal cue-cell are most efficacious in enabling listeners to infer the position in pitch space of a tonal centre. On the basis of this result and of further studies, Butler (1989) proposed an alternative to the tonal hierarchy theory of the cognitive-structuralists, the intervallic rivalry theory. This holds that listeners do not simply apply their pre-existing cognitive schematic representations of musical pitch to a piece as it unfolds, but engage with musical pitch organisation in the act of perception in a quite directed and dynamic way that is partly determined by the structural properties of those schematic representations[1]. They will tend to assume that pitches occurring early in a piece are reference pitches, and will rely on rare intervals such as the tritone in orienting themselves within "tonal space". Thus Brown and Butler's account of musical pitch in cognition appears more dynamic than that of the

---

1 It should be noted that many of the processes involved in listening under this account are still likely to be non-conscious.

cognitive-structuralists, focusing on the ways in which particular strategies are employed in the perception of pitch pattern and on the fluidity of the inferences enabled by the structural properties of the schematic representation of musical pitch.

The cognitive structuralists and the intervallic rivalry theorists thus propose two different mechanisms to account for musical pitch in cognition. The cognitive-structuralist research suggests that the primary processes involved in the perception of musical pitch are largely based on the pickup of information about the frequency distributions (or aggregate sounding durations) of pitches in the music of the perceiver's culture, both in terms of the "event hierarchies" of individual pieces and the "averaged event hierarchies" of the totality of the music to which the perceiver has been exposed, or tonal hierarchy. Pickup of frequency distribution information thus leads to the formation of schemata within which pitches are functionally differentiated within keys in terms of their relative stabilities. By contrast, the intervallic rivalry theory suggests that processes of musical pitch perception are likely to be based on both order-information and on sensitivities to structural relations between pitches, emphasising the active, discovery-oriented nature of the processes involved. Such processes rely on the application of schemata embodying information about the functional characteristics of intervals and sets of pitches - in particular, their diatonic multiplicity and their harmonic implications - in the listening process.

The differences between the two theories can be interpreted as a difference in focus rather than in fact. Both theories acknowledge the role of order information in determining the "stability" of pitches in perception, and both rely on some representation of diatonicism. However, while the cognitive-structuralist theory focuses on the differential frequency distributions of pitches within music as an operational characteristic in perception, the intervallic rivalry theorists emphasise the role of diatonic structure. This divergence between the two models can be interpreted as deriving from their focus on different perceptual stages and strategies in the perception of musical pitch. Brown, Butler and Jones (1994) themselves suggest as much in stating that the two models accentuate different aspects of tacit knowledge about tonality; the intervallic rivalry model centres on process of key discovery, the cognitive-structuralist account on reinforcement of tonal function, but both processes are necessary for a listener to follow tonal music in real time.

This suggestion, that the two theories reflect different stages or strategies in listening, is given force by the findings of both Castellano, Bharucha and Krumhansl (1984) (in which Western listeners responded to the event structures of the unfamiliar, North Indian, contexts while the responses of Indian listeners reflected "tonal hierarchical" relations appropriate to the music) and a recent study by Cuddy (1994). She conducted probe-tone experiments where diatonic contexts were specially composed so as to have frequency distributions of pitches (a component of the event structure) that differed from the tonal "norm". She found that subjects' responses to this unfamiliar music directly reflected the frequency-distribution aspect of the contexts' event-structures. This implies that when the detail of the flow of events in a piece is unfamiliar it is likely that listeners will respond to global statistical characteristics such as the frequency distribution of events, while in listening to pieces that conform more to cultural norms, prior knowledge concerning functional characteristics of intervals and pitch sets will be employed in interpreting the music's structure. A recent study by one of my final-year undergraduates, Naomi Gregory, reinforces these findings. She employed two brief atonal pieces as contexts in respect of which listeners judged the "fittingness" of subsequent single pitches, Schoenberg's Opus 19/2 and Webern's Opus 27/2; within the Schoenberg, pitch-distribution is markedly unequal (with the pitches G and B sounding for markedly longer than other pitches), while in the Webern, pitches are distributed more or less equally. She found that listeners' responses correlated significantly with pitch distribution in the Schenberg, while other factors determined the pattern of responses subsequent to the Webern piece; these other factors were related both to superficial characteristics of the piece (such as which pitches were most salient at the start and end of the piece) and to a complex structural characteristic of the piece, its symmetry around the pitch E@. A full report of the experiment can be found on the World-Wide Web at: http://www.mus.cam.ac.uk/PandP/PandP96.html

In other words, when presented with music in an unfamiliar but "followable" idiom in which structural relations between notes are used in unexpected ways, expectations based on abstractions of structure (prior tonal knowledge) will be violated and attention will tend to be directed towards the holistic properties of the music rather than to the detailed "argument" of its unfolding. In contrast, the experience of music in a "familiar" idiom will be determined more by a listener's capacity to engage pre-established cognitive schemata in deciphering the complexities of pitch structure that the music embodies.

Reinforcement of musical ideology versus evaluation of music theory

The accounts of musical pitch in cognition described above are tied in general to the experience of tonal music, the most widespread form of musical experience in Western culture; most of the music that surrounds us (in the form of radio broadcasts, recordings, or as the environmental pollutant, *Muzak*) subscribes to the principles of tonality. However, in what ways are these accounts relevant for contemporary composers? For instance, do these accounts of musical pitch in cognition have much to say about the music that has been composed in a self-consciously post-tonal idiom? We can answer this question in two very different ways; either we can assume that the processes involved in listening to tonal music determine our experiences of <u>all</u> music and should help us understand our experience of post-tonal music, or we can hypothesise that quite different cognitive processes are employed when we listen to or produce such music. Adopting the first strategy has led some researchers to employ the results of cognitive scientific findings and theories in order to support value-judgments about the "musicality" of post-tonal music; in other words, ideological claims about what <u>should</u> be regarded as music have been made on the basis of cognitive-scientific theories. The second approach has involved empirical testing of musical theories that bear specifically on post-tonal music and evaluation of the ontological claims of those theories in an attempt to discover whether the pitch organisation of post-tonal music affords listeners ways of experiencing patterns in musical pitch that are different from those that might be exploited within tonal music.

<u>Music theory - explanation and evaluation: Lerdahl and the cognitive inefficacy of post tonal structures</u>

The research that this paper has discussed so far appears to indicate that several fairly general cognitive principles are facilitated by the structure of tonal pitch organisation. Tonal structure affords a basis for judgments of similarity in pattern; it facilitates the formation in cognition of internally-structured categories (scales, keys, modes) whose members - pitches differentiated in terms of tonal function - are related to one another to different degrees; it enables pitch patterns to be organised in perception around pitches that are more central (prototypical) within categories; it enables musical structure to be perceived over long time-spans. Does this then mean that there is an immutable relation between cognitive processes and musical materials ? Are the functional principles of tonality a cognitive - and hence, compositional - "given", in that those functional principles "fit" uniquely well with our cognitive propensities and capacities ?

This is effectively the position adopted by Lerdahl (1988), who asserts that post-tonal music within which pitch is organised according to serial principles is "cognitively opaque". Serial organisation involves structuring pitch relations in a piece so that series of pitches or vertical combinations of pitches can be related to a basic ordering of the twelve available chromatic pitches in which no pitch is repeated, the original row form by operations of transposition, inversion and retrogradation. Hence pitch structure in a serial piece can be described in terms of the ways in which pitch successions and verticalities can be derived from - or can be thought of as more or less similar to - the serial pitch structure of the original row. According to Lerdahl, a serial piece's pitch structure is not "penetrable" by a listener, because the way in which pitch is organised in the music does not facilitate the operation of those cognitive processes implicated in judgments of pattern similarity, category formation and internal differentiation, etc. The implication here is that music that does not conform to tonal principles is of less value than that which does. Is such an inference justified ? Does science offer the basis for making value-judgments about cultural artefacts and the contexts from which these emerge ?

Lerdahl starts by suggesting that Boulez's *Le marteau sans maître* is "impenetrable" to a listener because "the serial procedures [employed in its composition] profoundly influenced the stimulus structure, leading to a situation in which the listener cannot form a detailed mental representation of the music. The result is a piece that sounds partly patterned and partly stochastic". He then examines the relationship between composition and listening by means of the concepts of generative grammar[2]; he suggests that there exist composition

[2] A generative grammar is a logical theory that enables the formal description of the structure of instances of classes of phenomena such as linguistic utterances or perhaps pieces of music; it also enables the generation of instances of such classes - such as sentences, or pieces of music, from abstract structural descriptions.

grammars (means of generating pieces) and listening grammars (means of generating abstract structural descriptions of pieces that correspond to their representations in the cognitions of listeners), and that in the case of serial music there is little or no correspondence between the composition grammar and the listening grammar. He draws a distinction between "natural" and "artificial" compositional grammars (the former depending on "the listening grammar as a source") and suggests that "a [natural] compositional grammar must be based on the listening grammar".

Lerdahl offers a detailed rationale for this assertion that is largely based on the specifics of the theory that he and Jackendoff outlined in their comprehensive 1983 book *A generative theory of tonal music* (*GTTM*). He states that underpinning the development of *GTTM* was the aim of producing a theory that could act as "the basis for artificial compositional grammars that could be intellectually complex yet spontaneously accessible to mental representation". *GTTM* provides the basis for his description of "psychologically plausible" constraints on compositional grammars, all of which contribute to the "projection" of "event hierarchies" to a listener (i.e., contribute to a listener's experience of detailed and articulated musical structure at different time-scales). These constraints include the propositions that the musical surface must be capable of being parsed into sequence of discrete events on the basis of "salient distinctive transitions", that the formation of groups of events at higher levels in a piece will depend on symmetry and parallelism, and that a range of factors will contribute to the formation of hierarchies (of importance, and of stability) of events as the piece unfolds.

The constraints - particularly those concerning stability - are related to a further set of constraints on "underlying materials": as he puts it, "stability conditions are contingent on the basic materials and properties out of which the event structure of a piece is made". Lerdahl suggests that the pitch materials employed in tonal music (discrete pitches organised within scales, modes and keys) and their particular properties (such as those outlined by Balzano, 1980 and experimentally examined by Butler and Brown) are uniquely well-suited to the delineation within perception of complex hierarchical structure within which different events are differently stable.

Having demonstrated that tonal pitch organisation is well-adapted to facilitate the cognition of complex musical pattern, Lerdahl then contrasts this organisation with that afforded by serialism, arguing that serialism is "cognitively opaque", because, *inter alia*, it relies on pitch permutations and "permutations are hard to learn and to remember", and hence is far less efficacious as a means of controlling and shaping structure in musical pitch than are the principles of tonality. He concludes by proposing that "the best music utilizes the full potential of our cognitive resources" and "the best music arises from an alliance of a compositional grammar with the listening grammar"; on this view, it follows that composers should adhere to the principles of tonal pitch organisation and should shun serialism (and all other "music-generating algorithms"), as in this way they will best be able to communicate their musical ideas to listeners by exploiting those musical attributes to which listeners are most sensitive

Can this argument - which seems to propose that composers should adopt a traditionalist approach that would have seemed reactionary a hundred years ago - really be taken seriously ? Is this seemingly backward-looking view in reality supported by the scientific evidence ? In fact, most of the claims that Lerdahl makes for the perception of music and of pattern in musical pitch are derived not from the experimental evidence of empirical studies but from the implications and details of his own theory, the *GTTM*; if his claims deserve serious attention it is because of the explicitness and comprehensiveness of *GTTM*, but they remain hypothetical in the absence of empirical research. Fortunately, a number of recent studies explicitly investigate some of the issues that he raises, in particular, those of the perceptibility of serial pitch structure, and of permutational pitch structure as described within set-theoretic approaches.

Music theory - evaluation and explanation IV: serialism

A fundamental tenet of post-tonal music is that all pitches are generally regarded as functionally equivalent; in particular, the use of serial principles to organise pitches so as to avoid pitch repetition reduces the likelihood that any single pitch will fulfil the referential role that the tonic plays in tonal music. A point that is thus of interest is the question of whether the use of serial pitch structure does in fact minimise the use in perception of single pitches as reference points, or whether such a strategy is so "ingrained" in our cognitive processes that it is employed despite the serial nature of the music that we are perceiving. In other words, is music which does not conform to the usages of tonality in its patterning of pitch still necessarily perceived as though it conforms to tonal usages because of our general tendency to employ reference elements in cognition (as Lerdahl appears to be claiming), or can we make use of alternative cognitive processes when we listen to such music ?

This question was explored by Krumhansl, Sandell and Sergeant (1987) in a series of probe-tone experiments. Listeners heard a row from Schoenberg's op 26 *Wind Quintet* or from his op 37 *String Quartet No 4*. and judged how well a subsequent probe note fitted with the preceding row within the "sense of the atonal idiom". It was found that those listeners who had had extensive experience gave low ratings to pitches that could be construed as possible tonics, dominants, etc. because of local key implications, and high ratings to pitches that did not fit in with these implications. Inexperienced listeners gave high ratings to pitches that could be interpreted as tonics, etc. within the contexts. In other words, listeners with little experience of post-tonal music appeared to employ very much the same cognitive processes as they would have employed in listening to tonal music, deriving tonal allusions from the rows that they heard, and thus treating certain pitches as referential if they could be interpreted as important on the basis of the tonal allusions. On the other hand, experienced listeners appeared to be doing something quite different. However, their responses can also be accounted for in terms of the perceptual processes that are operational in listening to tonal music, in that they also appeared to infer a key structure from the rows but gave pitches ratings that were inversely related to the pitches' "tonal importance" because they felt that this would reflect the "atonal" nature of the rows.

It seems that the results of this study imply that music which does not conform to the usages of tonality in its patterning of pitch is still perceived as though it conforms to tonal usages; post-tonal music seems to be experienced through a set of "tonal filters", rather in accordance with the claims of Lerdahl. However, this may be the case only for music that conforms to serial principles (as, for instance, the pieces employed in Naomi Gregory's experiment); what of the relations and structures that set theory describes ? Are these accessible to our musical perceptions ? The set-theoretic approach involves conceiving of relations between pitches in a piece as being governed by the relations between sets of pitches; these sets can include different numbers of pitches, and each set can be defined by the intervals that can be formed between all of the pitches of the set irrespective of the actual order of occurrence of the pitches in the set. Similarity between sets of pitches in a piece is determined by the degree to which different sets can be reduced (by operations of re-ordering, transposition, inversion) to the same pitch-class set, or by the similarity between the sets of intervals that can be formed between pitches of different sets. Hence the set-theoretic pitch structure of a piece is dependent on global properties of sets of pitches that is largely independent of the order in which pitches occur in different sets.

Music theory - evaluation and explanation V: the imperceptibility" of pitch-class set theory

There is, of course, evidence in the work of Brown and Butler (referred to earlier) that aspects of the set-theoretic description of musical pitch relations are operational in cognition, although the research that has produced that evidence is focused exclusively on the special properties of diatonic structures when viewed from a set-theoretic perspective. Much less empirical research has been conducted on the perceptibility of the relations between pitch-class sets in terms of the global characteristics described by set theory. In the 1980's three separate studies were carried out in the 1980's in an attempt to shed light on this issue.

A study by Millar (1984) examined the perception of equivalence relations - i.e., of similarity - between pitch-class sets, requiring listeners to compare a "standard" three-note set with subsequent three-note sets which could be related to the standard set by operations of transposition, inversion, re-ordering, or could be unrelated according to set-theoretic principles but might share some common pitches with, or share the contour of, the standard set. The results of this study were inconclusive; while some evidence for a sensitivity to pitch-class set relatedness could be inferred from subjects' responses, the pattern of results was also explicable in terms of a sensitivity to simple pitch contour (patterns of successive ups and downs in the experimental pitch sets).

Two other studies (Bruner, 1983; Gibson, 1986) tested a different aspect of set-theoretic similarity in perception. Both studies employed small sets presented as chords to test the perception of similarity between pairs of sets, specified in terms of the sets of intervals that can be formed between the pitches within each set (similarity measures devised by Morris, 1979, and by Lord, 1981). Two sets would be maximally similar if the set of intervals that could be formed between all the pitches within one set differed in only one respect from the intervals within the other set[3]. Both studies concluded that the theoretical similarity measures applied to sets do not bear any relation to the representation of similarity between groups of tones in cognition. Gibson's results demonstrated no correlation with similarity predicted by theoretical measures. In fact, Bruner (1983) found that one of the most potent factors in determining subjects' judgments of similarity between two sets was the degree to which the relations between the sets could be construed as in terms of tonal structures (e.g., whether or not two sets to be rated could be heard as "consonant" or dissonant").

The results of these studies appear to support Lerdahl's claim that we employ certain cognitive processes in listening to music that are best exploited by musics that conform to the principles of tonality, and that post-tonal musics make use of methods of pitch organisation that are "cognitively impenetrable": <u>we may see and understand them in the score, but we cannot hear them</u>. However, a series of studies by my research student, Diana Stammers (Cross and Stammers, 1992; Stammers, 1995) indicates that this may not be the case; one study in particular indicates that it may be possible that we <u>can</u> "hear" relations between sets of pitches in very much the ways prescribed within pitch-class set theory. This particular study has profound implications for our understanding of our capacities to experience structure in musical pitch.

[3] If the two sets of intervals were to be the same, then the pitch-class sets would not be similar but **identical** under this measure (which does not differentiate between identical and Z-related sets).

Music theory - evaluation and explanation VI: the "perceptibility" of pitch-class set theory

Stammers conducted two initial experiments (Cross and Stammers, 1992) which examined perception of similarity between sets under basic set-theoretic transforms of transposition, transposition-inversion and reordering. Her results hinted at a sensitivity to set-theoretic relations; they showed that, over all conditions, subjects (with and without musical training) perceived equivalent pairs as more similar than non-equivalent pairs. However, pairs of sets that were non-equivalent in set-theoretic terms but that shared common tones or contour or were both diatonic were also perceived as similar.

A second experiment was conducted in order to assess the effect of set-theoretic equivalence on perceived similarity using pairs of sets that were non-diatonic, did not share common tones and did not have similar melodic shapes or contours. The only difference between equivalent and non-equivalent pairs was match or mismatch of total interval content. The results of this experiment showed that subjects made no distinction between equivalent and non-equivalent sets, suggesting that the responses given by subjects in the first experiment were not made on the basis of total interval content but on the basis of the extent of matching other features - such as contour, common tones and diatonicism - across the two sets of a pair.

These two experiments demonstrated that equivalence between sets in terms of total interval content is unlikely to provide such sets with any greater perceptible similarity than that which occurs between non-equivalent sets. But they tested only the relations explored by Millar (1984), if rather more rigorously; they did not consider the possible effect of the similarity relations tested by Bruner and by Gibson. It is possible that because the relation between the total interval contents of the non-equivalent set pairs was not controlled in Stammers' first two experiments, the difference between the sets of intervals that could be formed between the pitches in each of the non-equivalent sets used was not sufficiently marked to make them stand out as different in perception.

Stammers' third experiment (Stammers, 1995, Ch 4) aimed to examine the possible correlation between theoretical similarity measures and the perception of similarity between sets under conditions which encouraged subjects to use interval-content similarity in making judgements about similarity between sets by minimising the possible effects of the factors that her earlier studies had found to be operational in perception. This required Stammers to adopt extraordinarily complex methods in order to control her experimental materials, but resulted in a situation wherein she could be certain that any consistency in listeners' judgments of similarity could be traced back to set-theoretic factors. She asked subjects to listen to two pairs of four-note sets and indicate which of the two pairs was more similar; each pair of sets was either unrelated (in set-theoretic terms) or was maximally similar (in terms of interval content) or maximally different (in terms of interval content). She used two groups of subjects in the experiment, one of whom had had extensive musical training while the members of the other group had experienced very little formal musical learning.

Her results showed an unequivocal effect of set-theoretic similarity between sets; those pairs of sets that were theoretically similar were indeed selected as being similar, but only by those of her subjects who had had extensive musical training. Subjects without that training did not appear to differentiate between sets on the basis of set-theoretic similarity. It should be noted that the musicians who participated in Stammers' experiment had had no formal training in set theory, and had indeed very little experience of post-tonal music (despite having undergone, on average, some 13.2 years of formal musical training) !

This experiment suggests that listeners only employ total interval content in the perception of pitch structures if all other possible means of relating groups of pitches are removed. The fact that subjects *can* use total interval content if forced suggests that such a concept is not altogether alien to their perceptual processes. People tend to use associations which have previously been used in another situation to encode a new set of properties, and the ability of musically trained subjects to draw on total interval content similarity to make their judgements in this experiment may suggest that this type of association has previously been tried and tested. It is possible that subjects may have been forced to apply a categorisation rule usually only applied to diatonic contexts (or to a small range of special pitch-event types) to a new context; alternatively, listeners may in fact employ total interval content similarities and differences in the cognitive processing of atonal music if necessary, but the design of previous experiments has masked this ability.

This series of experiments indicates that musically-trained listeners, at the least, can employ set-theoretic relations in making judgments about sets of pitches; Lerdahl's assertion that music that is not organised according to the principles of common-practice tonality is thus "cognitively opaque" is just plain wrong. However, Stammers' results do suggest that music organised according to set-theoretic principles may not be completely "cognitively transparent"; after all, set-theoretic relations only seem operational in the cognitions of musically-trained listeners, who might be expected to be able to devote more cognitive resources - and apply a wider repertoire of cognitive strategies - to the experience of pattern in musical pitch than untrained listeners.

Music theory as input to cognitive scientific theory

Stammers' results indicate that the fact that listeners tend to employ "tonal filters" in listening to music that is not tonal is not because the "principles of tonality" are immutably embedded in our cognitions; they suggest that we have learned to employ these "tonal filters" because these are best adapted to to help us to make sense of the majority of the music that surrounds us. The implication of these findings is that we cannot use an understanding of the cognitive processes employed in listening to make claims about the types of music to which we should listen; while cognitive science can explain musical experience, it should not be used to prescribe it.

Cognitive science can evaluate music theory; it can explain music theory, and can account for the nature of musical experience. I wish to consider briefly the third type of relation between cognitive science and music suggested by Rosner, that in which concepts from music theory and from the dynamics of musical experience contribute to cognitive science.

There are at least four ways in which an understanding of music and of our musical cognitions can contribute to the development of cognitive science. Firstly, music is a highly-structured non-verbal auditory domain. It does not appear to refer to anything beyond itself in the way that language does, yet music theory suggests - and empirical research demonstrates - that we are capable of perceiving extremely complex patterns in musical sound. The cognitive processes that subserve this capacity are beginning to be illuminated by research of the type that I have outlined in this paper. The processes that enable us to abstract the regularities of our auditory - our musical - environment, to organise these regularities in our perceptual and memory systems and to employ the resulting schemata to articulate and to shape our experience of complex pattern in sound, surely has profound implications for ourselves as cognising systems. It indicates clearly that our experience of intentionality, of meaning, is not tied solely to our use of language or even to our capacity to comprehend and manipulate information in formal terms, as in the domains of mathematics or formal logic.

Secondly, music conveys affect; it is a domain of experience that is intimately bound up with our emotional life. An understanding of the relation between our experience of music and the patterns of emotion to which it gives rise provides us with some insight about the function of emotion in cognition. The recent development of cognitive theories of emotion, and the increasing amount of research that indicates that our experience of emotion in music is tied to our experience of musical structure are converging to produce an ever more clear account of the roles and functions of emotion in our everyday cognitions.

Thirdly, much recent research (see, e.g., Clarke, 1993) indicates that our experience of music is consequent upon our embodiment. In other words, the ways in which we experience music and we produce music are at least in part determined by the possibilities afforded by the human body. The structure and dynamics of the human body determines much about how we can produce music and thus determines a large part of our experience of music. The fact that human beings, by-and-large, possess fairly similar bodies with fairly similar motoric capacities provides a framework within which we can share aspects of musical experience, a framework that provides us with a very basic, implicit and common understanding that enables musical communication. The need to "situate" cognition, to study cognition as a phenomenon that is conditional on its embodiment in the world, is becoming increasingly recognised within cognitive science (see, e.g., Hendriks-Jansen, 1996), and music provides us with a rich example of "situated cognition".

Finally, music is a product of culture. It arises through the interactions of members of a culture as they identify and select regularities of the world of sound and of the human sensitivity to sound in time (or inherit and modify such a selection) by conferring value on, codifying and transmitting principles of organisation that enable members of the culture to share more-or-less abstract representations of musical phenomena. The dynamics of processes of cultural interaction, and the impact of those processes on our cognitions, are complex and still obscure, but are more and more recognised as integral to our understanding of ourselves as cognising beings. Bruner (1990) suggests that "It is culture…that shapes human life and the human mind, that gives meaning to action by situating its underlying intentional states in an interpretive system". Music, and the interactions between composers, performers and listeners, provides us with a rich field in which to explore the shaping action of culture on cognition. As we have seen, an understanding of musical experience in terms of cognitive science does not enable us to prescribe the forms that music may take; however, it may be that an understanding of musical experience as an aesthetic process situated firmly in the domain of culture might be capable of enlightening cognitive science as to the qualities, the function, and the importance of artistic apprehension.

References

Balzano, G. (1980). The group-theoretic representation of 12-fold and microtonal pitch systems. Computer Music Journal, 4, 66-84.

Bharucha, J. J. (1987). Music Cognition and Perceptual Facilitation: a Connectionist Framework. Music Perception, 5, 1-30.

Bharucha, 1991. Pitch, harmony and neural nets: a psychological perspective. In P. Todd and D. G. Loy (Eds.), Music and connectionism. Cambridge, Mass: M.I.T. Press.

Bregman, A.S. (1990). Auditory scene analysis: the perceptual organisation of sound. Cambridge, Mass: M.I.T. Press.

Bregman, A.S. (1993). Auditory scene analysis: hearing in complex environments. In Macadams, S. & Bigand, E. (Eds), Thinking in sound: the cognitive psychology of human audition. Oxford: O.U.P.

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. Music Perception, 5, 219-250.

Brown, H. and Butler, D. (1981). Diatonic trichords as minimal cue-cells. In Theory Only, 5, 39-55.

Brown,H. Butler, D. and Jones, M. R. (1994). Musical and temporal influences on key discovery. Music Perception, 11 (4), 371-407.

Bruner, C (1984) The perception of contemporary pitch structures Music Perception 2(1): 25-39

Butler, D. (1989). Describing the perception of tonality in music: a critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. Music Perception, 6, 219-242.

Castellano, M. A., Bharucha J. J. and Krumhansl C. L. (1984). Tonal hierarchies in the music of North India. Journal of Experimental Psychology: General, 113, 394-412.

Clarke, E. F. (1993). Generativity, mimesis and the human body in music performance. Contemporary Music Review, 9, 207-220.

Cook, N. (1989). Music, imagination and culture. Oxford: O.U.P.

Cross, I. (1997). Pitch schemata. In J Sloboda and I Deliège (Eds.) Perception and cognition of music. Hove, Sussex: Psychology Press.

Cross, I. & Stammers, D. (1992). Music analysis and the perception of set-theoretic relations. In Proceedings of the Secondo Convegno Europeo di Analisi Musicale, R. Dalmonte & M. Baroni (Eds), Università degli Studi di Trento: Trento.

Cross, I., West, R & Howell, P. (1991). Cognitive correlates of tonality. In P. Howell, R. West & I. Cross, (Eds), Representing musical structure. London: Academic Press.

Cuddy, L. L. (1994). Tone distributions in melody: influence on judgments of salience and similarity. In I. Deliège (Ed.), Proceedings of the 3rd International Conference on Music Perception and Cognition, Liège, July 1994.

Forte A (1973). The structure of atonal music. New Haven: Yale University Press.

Gibson, D. B. (1986) The aural perception of non-traditional chords in selected theoretical relationships: a computer-generated experiment. Journal of research in music education 34(1): 5-23

Hendriks-Jansen, H. (1996). Catching ourselves in the act. Cambridge, Mass: MIT/Bradford Books.

Krumhansl, C. L. (1990). The cognitive foundations of musical pitch. Oxford: Oxford University Press.

Krumhansl, C. L., Sandell, G. J. and Sergeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. Music Perception, 5, 31-78.

Leman, M. (1995). A model of retroactive tone-centre perception.. Music Perception, 12 (4), 439-471.

Lerdahl, F. (1988). Cognitive constraints on compositional systems. In J. Sloboda (Ed), Generative processes in music. Oxford: O.U.P.

Lerdahl, F. and Jackendoff, R. (1983) A generative theory of tonal music. M.I.T. Press, Cambridge, Mass.

Millar, J. (1984) The aural perception of pitch-class set relations: a computer assisted investigation. Unpublished PhD thesis, N.Texas State University.

Neisser, U. (1976). Cognition and Reality. New York: WH Freeman & Co.

Parncutt, R. (1989). Harmony: a psychoacoustical approach. London: Springer-Verlag.

Rosner, B. (1988) Music Perception, Music Theory, & Psychology. In E. Narmour, E & R. Solie (Eds), Explorations in Music, the Arts and Ideas. New York: Pendragon Press.

Scruton, R. (1983). The aesthetic understanding. London: Methuen.

Stammers, D. M. (1995). Set theory in the perception of atonal pitch relations. Unpublished Ph.D thesis, University of Cambridge, UK.

Terhardt, E., Stoll, G. and Seewan, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. Journal of the Acoustical Society of America, 71, 679-688.

Treitler, L. (1980). History, Criticism and Beethoven's Ninth Symphony, Nineteenth Century Music, 3, 193-210.