Musical Schemata in real-time listening to a piece of music

Irène Deliège*, Marc Mélen*, Diana Stammers**, Ian Cross**

* Unité de Recherche en Psychologie de la Musique
Centre de Recherches Musicales de Wallonie
Université de Liège

** Faculty of Music
University of Cambridge

Requests for reprints may be addressed to I. Deliège, URPM, Laboratory of Cognitive Psychology, 5 boulevard du Rectorat, B 32, B - 4000 LIEGE (Belgium)
e-mail: URPM@VM1.Ulg.ac.be

ABSTRACT

A series of experiments investigated cognitive processes involved in listening to a piece of music, focusing in particular on the abstraction of surface features (here referred to as cues). Subjects listened to an unfamiliar piece in a familiar musical idiom and their sensitivities to aspects of the just-heard piece were employed to elucidate the nature of their representations of the piece in recent memory. The study also sought to assess the capacities of subjects to employ any declarative knowledge of aspects of tonal structure that they possessed in organising musical material. Three experiments made use of different procedures to address these issues, employing either a single short tonal piece - Schubert's Valse Sentimentale, D 779, Op 50, n° 6 - or a variant of this. The first two experiments employed non-musician subjects and examined, respectively, the cues abstracted in listening to the piece, and subjects' post-listening ability to identify the temporal location of segments of the piece. The third experiment explored the constructional abilities of musician and non-musician subjects, requiring them to create a coherent piece by ordering the segments that made up the original piece. The results of these experiments indicated that while the abilities of musicians differed from those of non-musicians, both groups of subjects exhibited a weaker sensitivity to features of musical structure than to cues abstracted from the musical surface.

Introduction

One of the most complex and seemingly intractable problems in the study of music cognition is that of just how a listener may experience a complete tonal musical work as it unfolds in time. While the issue has long been the focus of much diverse and stimulating theoretical work, it is largely within the last decade that empirical methods have been brought to bear on this question. A wide range of approaches has been applied in experimental studies, and much light has now been shed on processes that may underlie the real-time perception of whole musical works. However, to date, certain facets of the problem remain unexplored. This is in some ways unsurprising; after all, listening is a private process, rarely involving any overt behaviour, and the application of experimental method that relies on observation of behaviour to elucidate the listening process is immediately problematic. It is thus extremely difficult to monitor the cognitive processes involved in listening without running the risk of interfering with these processes to some degree by requiring subjects to exhibit overt and recordable behaviours. Moreover, the complexities of "real" pieces of music are difficult to square with the requirements of experimental method stringently to control the stimulus material employed.

These difficulties can be seen to have conditioned the ways in which the problems of elucidating the listening process have been tackled. Much of the research upon which current hypotheses about how we experience whole musical works is based addresses itself to the ways in which various musical parameters (rather than complete pieces) are represented in cognition. This research has generally employed short and somewhat artificial sequences in which a given dimension could be isolated and manipulated in order to study its particular effect and interactions with other parameters. This approach has provided valuable advances in our understanding of aspects of the cognition of musical pitch organisation (see Krumhansl, 1990 for a review), of melody (see, e.g., Tighe & Dowling, 1993), of rhythm (see, e.g., Handel, 1989) and of timbre (see McAdams, 1993). Although this research has provided information on the cognitive resources which may be employed by musicians and non-musicians in the perception of music, they have left largely untouched issues concerned with representation of larger-scale - or even "real-life" - musical structures in cognition.

Over the last decade there has been an increasing number of studies directed towards the elucidation of just such "real-life" listening processes. An increasing number of these studies require listeners to experience complete pieces of "real" music and to make judgements about their experience in ways that are intended to interfere minimally with that experience. However, while much valuable insight is being gained from these studies, the majority are directed at the elucidation of processes involved in the perception of local rather than global features of musical structure (for a general overview of these studies see Krumhansl, 1991a; McAdams & Bigand, 1993), and the question of how complete pieces are experienced remains relatively unexplored.

A large body of work does exist that can be claimed to have a direct bearing on that question. While this has been conducted from the viewpoint of production rather than perception, much has been learned about the forms that the cognitive representations of complete pieces might take (see Gabrielsson, 1987; Clarke, 1988, 1993; Palmer & Vandesande, 1993). Nevertheless, this research of necessity employs subjects who are highly musically-trained; the ways in which such subjects might grasp the totality of a piece in cognition might be expected to be very different from those that less experienced listeners might employ.

It seems important to address the perceptions of listeners of different degrees and types of musical experience, and to do so by employing "real" music in a variety of experimental situations so as to enable broad and robust inferences to be made about the processes involved in listening to music. Several recent studies meet just these criteria, and appear to fall into four main groups. The first group can be thought of as those studies that have focused on demonstrating how a piece may be segmented in cognition (e.g. Krumhansl and Jusczyck, 1990; Jusczyck and Krumhansl, 1993; Clarke & Krumhansl, 1990; Stoffer, 1985, Deliège, 1989). Using a range of different methodologies (although those employed by Deliège, 1989 and Clarke & Krumhansl, 1990, were similar), these studies have demonstrated high degree of convergence in the segmentations perceived by subjects (musicians, non-musicians or infants). Whatever the musical background (in terms of degree and type of musical training), and whatever the musical material employed (pieces from Mozart to Messiaen, Berio, Boulez or Stockhausen), subjects were shown to experience segmentations at points which coincided with the main structural articulations of the piece. These studies contribute significantly to the understanding of the cognitive representation of complete pieces of music, but on the whole have not addressed the issues of the nature or of the complexity of relations between musical elements within pieces. This question has been the focus of the second type of study (e.g. Cook, 1987; Deliège I., 1989; Deliège & El-Ahmadi, 1990; Krumhansl, 1991b; Pollard-Gott, 1983), which shed light on subjects' ability to establish relationships between different musical structures drawn from a same piece. Questions such as which segments are perceived as similar, or what determines the perceived similarity, were addressed by these approaches. Nevertheless, these studies again did not focus on the representation of a piece as a whole. This was the explicit aim of a third type of study, of which some have concentrated on what might be thought of as artificial or restricted subsets of tonal music (e.g. Serafine, Glassman, & Overbeeke 1989; Sundberg & Lindblom, 1976, 1991). More recently, Dibben's (1994) study investigated listeners' experience of complete atonal pieces and of extracts from tonal works, and demonstrated that the cognitive representations that were formed incorporated abstract and complex structural relations between elements of the pieces when these were tonal. However, her subject-group was highly musically-trained, and the same limitations that apply to the results of production studies might be thought to apply here: that is, the question of whether her results can be generalised to listeners whose competence must be assumed to have arisen through processes of acculturation (see Sloboda, 1985) remains open.

The present research can be considered as representing a fourth type of study. We report the results of a series of experiments employing the same short piece within different experimental procedures, which addresses not only the question of the relationships that listeners could establish between the segments, but also the issue of the ways in which they could assemble these segments to form a complete piece, employing both musician and non-musician subjects. This research was undertaken with the aim of expanding existing perspectives on the processes and constraints that bear on the formation, structure, stability and accessibility of the representation of a piece in cognition.

The empirical investigation of how a cognitive representation of an entire piece of music arises in the course of listening is necessarily complex, requiring consideration of many diverse factors. The "real-time" processes involved in listening are influenced by, on the one hand, a listener's prior musical experience and competence and on the other by the specific idiosyncrasies and the cultural and historical provenance of the musical material employed. The piece used was the Valse sentimentale D 779, Op. 50 No. 6, by Franz Schubert (see Figure 1). This piece was chosen as it was deemed unlikely to be familiar to any of the subjects and conformed to a structure which can be said to be archetypal[i] in

[i] Ratner (1980, p. 209) suggests that "when we look at classic forms according to the harmonic plan, by far the large majority of movements…carry out a two-phase action" which has its origins in the binary forms of dance music. He cites Koch's treatise of 1793 in support of this contention, suggesting that harmonic layout usually conforms to one of the simple strings *I-I-x-I* or *I-V-x-I*, stating that Koch himself claimed that the

respect of tonal music of the classical period, and hence, highly representative of works in a tonal idiom.

Figure 1: The figure shows Schubert's <u>Valse sentimentale,</u> D 779, Op. 50 No 6, as well as the partitioning into eight segments that was used in the ensuing experiments employing Version A.

Three different experimental procedures were used. The first employed a <u>cue localisation</u> task (adapted from I. Deliège, 1991, 1992a) in which subjects, after hearing the piece, were required to locate temporally the auditory events that they judged to be most salient. The second and third procedures made use of a partitioning of the piece into eight segments (as shown in Figure 1), and required the subjects to organise the segments and to establish the relationships between them. The second procedure was based on the idea of the <u>mental line</u> (adapted from I. Deliège, 1991, 1993); in this procedure, following a complete hearing of the piece, subjects were asked to position segments on a line that they were informed symbolised the temporal progression of the piece. The third procedure examined subjects' constructional capacities, employing a <u>puzzle</u> task (see Deliège, Mélen, Stammers & Cross, 1994) in which segments were presented to subjects as visually undifferentiated but sounding objects via computer with no prior hearing of the piece. The subjects' task was to create a coherent ordering of the segments with no direction given by the experimenters. All these experiments were relatively "free-form" in that subjects' responses and activities during the experiments were minimally constrained.

<u>The perception of tonal works - theory</u>

The most comprehensive account to date of the perception of entire musical works is the well-known <u>Generative Theory of Tonal Music</u> (<u>GTTM</u>) of Lerdahl and Jackendoff (1983). Their theory seeks to describe the ways in which a listener relates musical surface structure

---

simplest piece would be 16 bars in length, each of the symbols in the above symbol-strings standing for a four-bar phrase.

to underlying structure. The impact of this theory has reinforced the general acceptance of the idea that harmonic relationships can predict the way in which the listener organises his or her perception of a piece of tonal music. Harmonic relations play a significant role in their theory, both in their contribution to the segmentation of the flow of events in a work and in their articulation of patterns of stability and instability, or tension and relaxation. While Lerdahl and Jackendoff did not explicitly propose a model of the time-course of musical listening processes (although see Jackendoff, 1991), their conception of underlying structure (perhaps primarily comprising harmonic events and their relations) could constitute a parsimonious way of storing tonal material in memory. The same underlying structure could be related by a listener to many different surface structures. The capacity to abstract an appropriate underlying structure in listening to a given piece could thus be held to represent the most highly developed mode of musical listening.

An alternative account is explored in the work of I. Deliège, which focuses more on characteristics of the musical surface in cognition than on the cognitive consequences of harmonic structures in an attempt to provide a theory that is generalisable beyond the tonal repertoire. Deliège's approach takes as its starting point the grouping structure aspect of the GTTM, grouping preference rules having been claimed by Lerdahl and Jackendoff to be empirically verifiable. Their theory of elementary rhythmic groupings appears indeed to be the most powerful part of their model, as it is claimed to be applicable to any music irrespective of historic or cultural origin. I. Deliège's experimental work has appeared to validate Lerdahl and Jackendoff's grouping theory, at least at local levels (Deliège, 1987b); however, as she points out "the accumulation in memory of small rhythmic units in a long series of groups is not in itself sufficient to constitute a cognitive representation" (Deliège, 1991, 1992a, p. 8).

Deliège (1987a, 1989) hypothesised the operation of a "mécanisme d'extraction d'indices", here translated as <u>cue-abstraction mechanism</u> . The cue would be a "very restricted entity... often shorter than the group itself, but always embodying striking attributes" (I. Deliège, 1989, p. 307) which suffice to signify longer entities or groups. The cue abstraction mechanism may, in this sense, be compared to the well-known figure/ground phenomenon. Information is reduced by focusing attention on small features that can be distinguished from a more diffuse or global environment. The cues play the figural role and become abstractions used to lighten the load on memory storage. It should be noted that, as in the figure/ground phenomenon, the perceptual qualities of the cue are differentiable from the environment from which it is abstracted. As is the case for a figure, a cue would be more salient and consequently more precisely defined than elements of its environment (or "ground") which would be less structured (Deliège, 1992 b).

In this way, longer structures (motives, phrases, etc.) may be labelled and encoded in memory by means of cues which alone are stored in immediate memory and enable access to the entire structure. The cues can then be used as sign-posts; they embody invariant characteristics of the musical material, and take on the function of referential entities enabling constant evaluation of new material[ii]. Thus the cue provides the basis for grouping and for the chaining together (concatenation) of groups at different hierarchical levels of listening; concatenation of groups will continue for as long as a particular cue continues to be recognised. When a new and contrasting cue is perceived, it establishes the boundaries of the higher-level grouping - periods or sections of the piece - and can initiate a new grouping. In other words, grouping can be seen to make use of two principles; the principle of "sameness", determining the length of the concatenated groups and that of "difference" which establishes their boundaries (Deliège, 1989, p. 308).

[ii] Krumhansl (1995)'s notion of "new musical idea" seems synonymous to the notion of cue.

In determining what might constitute a cue it is necessary to take into account the cultural and historical provenance of a given piece. As pointed out by C. Deliège (1989), those elements of music that we might now consider as "motivic" (constituting local structures of which the role as an invariant basic cell is reinforced by forms of repetition, imitation and variation of the original cell) make their first appearance in Western music only in the fifteenth century, their role becoming progressively integral to common musical practice during the course of the eighteenth century. Thus from this period throughout the period of common tonal practice, cues are likely to be provided principally by motivic components. In subsequent periods - and most likely, in other cultures - other types of musical elements may operate to fulfil this role (I. Deliège, 1991, 1992a).

Cues become cumulatively more numerous in the course of listening but only those that are more resilient will be retained in memory. I. Deliège accounts for this in the context of a further hypothesis; either as a work proceeds, providing variations of cue structures or after several listenings, the traces left by these cues in memory will come to constitute imprints, i.e. types of schematic representations, exhibiting prototypical characteristics wherein the particularities of the different cues are unlikely to be preserved (I. Deliège, 1989, 1991, 1992a & b). An imprint is not static, but self-adjusts in accordance with the various presentations and manifestations of the cues; thus the time-course of the perception of a piece of music must be regarded as extremely significant in any study of music perception.

The temporal structuring of a musical work by both listener and performer has been the focus of the work of Imberty since 1975. His theory outlines processes that permit the structure of the musical piece to be grasped by means of the repetition of the organisation of musical segments at different hierarchical levels (Imberty, 1981). On a first hearing of a piece, perceptual segmentation relies on the perception of qualitative changes in the musical flow**.** In respect of perception of these changes, Imberty defines two types of hierarchical organisation (weak and strong) directly linked to the perceptual salience and to the number of changes encountered in listening; a strong hierarchy arises when changes are few but very clear, while a weak hierarchy will exist when segmentations are numerous but are all of similar salience. In addition to these two types of organisation, Imberty has developed the ideas of schemata of order (concerned with simple successions and juxtapositions of material) and schemata of order-relation (where materials can be related by processes of reduction or elaboration)[iii]. These concepts can be linked together (see also Deliège, 1991, p. 184) in order to differentiate between characteristics of global musical structure as perceived. In Imberty's view, pieces from the tonal repertoire are likely to elicit the perception of weak changes ("appearing more as links that function coordinatively rather than as contrasts or breaks in the temporal unfolding": Imberty, 1981, p. 91-92), leading to schemata of order-relations and the formation of strong hierarchies. Cases where strong changes are elicited, leading to schemata of order and the formation of weak hierarchies, are more likely to be found in post-tonal works.

However, it does not seem feasible to argue that music that exhibits strong hierarchical organisation necessarily invokes only schemata of order-relations, or that a weak hierarchical organisation predicates solely schemata of order. What may be observed in "real-world" music, on the contrary, is the dominance of one organisational type or the other, that changes according to historical period (Imberty, 1985, p. 113). This consideration makes apparent the elements that confer directionality on the piece in perception, called by Imberty the dynamic vectors. These are set up and reinforced during listening by the successive changes which define the alternation of tension and relaxation; this idea provides the basis of Imberty's concept of macrostructure, a concept that he borrowed

[iii] Cf the differentiation between temporal structure and alphabetic structure outlined in West, Cross and Howell, 1991.

from Kintsch and van Dijk (1975) who advanced this idea in the context of text-comprehension.

Imberty has recently (Imberty, 1991; 1993) taken over the idea of cue put forward by I. Deliège, and has proposed an extension of this idea in suggesting that a listener is likely to abstract two types of cues which would underlie the articulation of the macrostructure: "objective" cues (themes, registral usages, etc.) which are structural, and are not only perceivable in listening but also identifiable in the musical score; and "subjective" cues which have psycho-dynamic functions (impressions, for example, of development, or of commencement) which may be experienced differently from one listener to another and are not necessarily identifiable in the score. The first type of cue would refer to time "in the abstract" and would express schemata of order-relations. The second type of cue would relate to time as experienced wherein events are ordered in a specific and irreversible order, representing the expression of schemata of order. In tonal music, dynamic aspects are evidently perceived but their organisation will tend to coincide with thematic organisation. The incidence of dynamic functions predominates, however, in atonal music, as the sign-posts pre-established by the tonal system disappear.

<u>The perception of tonal works - experiment</u>

In the main, studies that are intended to address the issue of how complete tonal pieces may be perceived have tended to employ fragmentary musical materials, and have extrapolated from their results in proposing theories applicable at the level of complete pieces. A number of recent studies have explored the implications of Lerdahl and Jackendoff’s theories (in particular, the idea that harmonic structure is likely to play a significant role in listeners’ perceptions of a whole piece) by using such materials and methods. Bigand, for example, in exploring the idea that underlying harmonic structure can be more predictive of accurate recognition of a melody than can elements present at the surface structure (Bigand, 1990), required subjects to recognise familiar melodies either after the melodies had been subjected to a transformation that altered their rhythmic grouping structure or after they had been "reduced" on the basis of their harmonic framework. The first transformation maintained all the pitches of the original melody but made their rhythmic grouping incompatible with its implied harmonic organisation. The second transformation preserved the harmonic structure but eliminated some pitches. This study showed that the original melodies were better recognised after reduction than after transformation of the rhythmic groups. However, recent experiments (Mélen & Deliège, 1995) have indicated that Bigand’s findings may be explained in terms of limitations of short-term memory; recognition of melodies appeared to depend crucially on the degree to which cues at the musical surface had "degraded" in memory rather than on ostensible adherence to underlying harmonic structure.

Another instance of empirical testing of implications of Lerdahl and Jackendoff’s theory is provided by Oura and Hatano (1988, 1991). They proposed a model for long-term memory storage of familiar melodies, which would be stored in a reduced form similar to that obtained by application of Lerdahl and Jackendoff’s prolongational reduction rules. They recognised, however, that elements of the surface structure might intervene in the formation of a long-term representation; when the theoretical reductions of the melodies are presented to musicians, the rate of correct identification is very poor unless some information from the surface structure is provided (e.g., the pitches and the proportional durations of notes from the original melodies). This gloss on Lerdahl and Jackendoff’s model appears highly relevant to further experimental studies; whatever the subject’s competence in respect of the tonal system - notably aspects of it concerned with tension and relaxation (as studied by Bigand, 1993a & b; Nielsen (1983) - it seems very likely that the cognitive organisation underlying this competence arise through processes that rely on elements from the musical surface as suggested by I. Deliège (1987a, 1989).

Further studies by Bigand (see Bigand, 1993a and 1993b) appear to demonstrate that listeners, both musicians and non-musicians, are sensitive to the "tension-relaxation" schemata in terms of which tonal pieces can be characterised. Bigand employed short self-composed melodies which had ostensibly similar "surface features" of rhythm and contour, but different underlying "tension-relaxation" schemata (harmonic structures). Listeners were presented with fragments of the melodies that ended on notes that exhibited higher or lower degrees of "tension" in the context of the melody as a whole, and were required to rate the "completeness" of each fragment on a seven-point scale. Ratings were integrated so as to produce graphs of tension and relaxation, and it was found that the graphs produced by musician subjects showed features that were predictable from the prolongational analyses of the melodies (i.e., analyses of tension-relaxation relations), and that graphs for pieces with different structures were significantly different from one another. Evidence was found that even non-musician subjects could perform the task, although their performance exhibited more variability than did that of the musicians.

While Bigand's results appear to add weight to the contention that underlying (harmonic) structure of a piece is a principal component of its representation in cognition, his findings are limited in the degree to which his experimental procedures can be thought of as tapping "into the moment-by-moment history of mental involvement with the music" (Sloboda, 1985, p.120). Moreover, the fact that the musical materials employed in the experiment were solely melodic (relying on implicit harmonic organisation to differentiate the functions of different notes) again limits the generalisability of his results to the perception of complete tonal works.

A recent study by Krumhansl (1995, see also Fredrickson, 1995) explored the perception of tension and relaxation and the identification of new musical ideas in the course of listening, employing subjects who varied considerably in their level of musical training. Analyses showed a great degree of convergence between judgments and were strongly correlated with the structural gestures within the music. At first glance, this study's finding that this pattern of results was found using a complete piece extends Bigand's results. A close look at the structural features suggests, however, that factors other than harmonic relationships might underlie the pattern of subjects' judgments. Release in tension coincided with the structure of the piece derived from Lerdahl and Jackendoff's theory. As Krumhansl (1995) herself acknowledges, most of the features that were responsible for subject judgments were available at the musical surface. As she also stresses, the ends of sections coincided in the piece with changes in surface characteristics. For this reason, the possibility cannot be ruled out that subjects' judgments were determined in the first instance by surface features of the music that in turn correspond to the underlying structure of the piece, a hypothesis that will be studied further in the present paper.

These experimental studies follow Lerdahl and Jackendoff in holding that the cognitive representation of music is hierarchically organised. I. Deliège has conducted a number of experimental studies within a theoretical framework that does not make this assumption. The differentiation between processes involved in the perception of tonal and of non-tonal works that is explicit in the theories of Lerdahl and Jackendoff is hence not present in Deliège's theory. Her approach is based on a set of experimental procedures that are intended to be adaptable to any musical genre and provides multiple perspectives on the listening process. These procedures relate to the theory of cue abstraction and imprint formation outlined above. The segmentation procedure requires subjects to listen to a piece as they would listen to a story. They are asked to indicate segmentations of an unfolding piece by pressing a button when a "change" is perceived, analogous to a full stop, paragraph- or chapter-ending in a piece of text (Deliège, 1989, 1990). These studies, employing both musically trained and non-musicians subjects, show that major segmentations intervene where important contrasts in the perceived cues appear. Clarke and Krumhansl (1990) provided also evidence of this tendency. However, in their experiments on Stockhausen's Klavierstück IX, they focused primarily on the segmentation responses of the subjects rather than on the musical content of the delimited groups. The segmentation points were sorted into four categories that corroborated Lerdahl and Jackendoff's grouping preference rules: long rests; contrasts in dynamics, texture, register and rhythm; changes in pitch content, melodic contour; repetition of previously heard material. Their experiment on Mozart's Fantasie in C minor K. 475 supported the previous findings (except for segmentation for long rests, as this piece did not provide such instances). Krumhansl (1995) reported further empirical evidence of the incidence of surface characteristics on the segmentation of a piece into several regions.

The mental line approach (Deliège, 1989, 1991, 1993) makes use of different procedures, and is aimed at investigating the traces in memory of a piece's unfolding in time. Again the reference points that are provided by the abstracted cues appear to play a leading role. Her 1989 study, employing a piece by Berio (Sequenza VI for viola solo, with a duration of ca 10 minutes) was first segmented by the subjects into six main sections, each based around distinct cues, using the segmentation procedure. Non-musician and musician subjects were then asked to listen to about forty segments and to indicate from which section the presented segment was derived. The abilities of both groups of subjects to accurately locate the segments within the large-scale framework of the piece was closely tied to the nature of the cues embedded in each segment. Clarke and Krumhansl (1990) use a similar procedure (experiment 2 on Stockhausen, and experiment 5 on Mozart). In a second study, employing Boulez's *Eclat*, Deliège (1991, 1993) employed a new version of the mental line procedure; 15 segments had to be localised by the subjects on to a horizontal line divided in 15 boxes which symbolized the complete duration of the piece. The second experiment of the present research is based on the latter version of the procedure.

The framework is further elucidated by categorisation approaches, examining the effects of frequency of occurrence, family resemblance, and discriminability of cues. Deliège (1991, in press) employed a solo violin work by J. S. Bach, requiring non-musician and musician subjects to estimate frequency of occurrence of motifs and to classify and discriminate between more and less related musical fragments. A clear advantage for the musician subjects appeared in this procedure, their responses more accurately reflecting motivic characteristics of the music across all levels of relatedness between motifs; non-musicians were able to undertake the task, but were less likely to perceive distantly-related motifs as being at all related. In general, these experiments appear to indicate that the listening processes of musicians and of non-musicians are dependent on similar sensitivities to characteristics of the musical surface, but that musicians' cognitive representations are more accessible and more detailed than are those of non-musicians.

In the present study, subjects' sensitivity to the musical surface and prior knowledge of the hierarchies implied by the tonal system are investigated through comparison of results of experiments which involve or preclude previous exposure to the piece as well as comparison of the performance of musicians and non-musicians. All the experiments required subjects to establish the relationships between constituent elements of the piece, irrespective of the specific methods employed in the different experiments. A cue-localisation procedure was employed to illuminate the types of structures most immediately salient for the subjects, and, as it involved two consecutive listenings, also provided information about their stability in perception. A mental line procedure, involving prior listening to the piece, was intended to enable an evaluation of the role played by episodic information - cue information - in the make-up of the piece in cognition. A puzzle procedure made straightforward use of subjects' declarative memory, i.e. of their musical knowledge prior to hearing the piece, in exploring their capacity to call on that knowledge in making judgments about musical coherence as they constructed a piece from pre-existing musical segments. Analysis of the placement of segments within the pieces that were constructed by subjects indicated the degree to which relationships between segments (which can be thought of in terms of tension and relaxation) were cognitively accessible. This last experimental procedure also provided a valuable way to compare expert and non-expert performances.

EXPERIMENT 1 - CUE-LOCALISATION PROCEDURE

It has been shown that non-musicians, after listening to a piece, can provide reports of those characteristics of the piece they judge to be most significant; these characteristics can be described as acting as cues that are abstracted in the course of forming a cognitive representation of a piece (see, e.g., I. Deliège, 1992a). The cues that are potentially abstractable by the listener are in general specific to the particular piece heard. The aim of this first experiment was to ascertain which features of the piece in question might constitute actual cues that could be abstracted by the non-musicians.

Method

Subjects: Seven non-musicians (between the ages of 18 and 24, mean 21.9) took part in the experiment. They were students of different faculties at University of Liège (Philosophy, Psychology and History). None of them had received music tuition.

Material: A sixteen-bar piece for piano by Schubert, Valse sentimentale D 779, Op. 50 No. 6, comprising four main four-bar phrases (see Figure 1), was used in this experiment. This piece was selected as it was short (30 seconds), and apart from one phrase, the third, included no direct repetition of musical material. In addition (see Note 1 above) it can be thought of as conforming to a simple "archetypal" structure, and hence as being highly representative of tonal music from the classical period. The repeat marks in the original score were not observed so as to ensure that the order in which events occurred in the piece was as unambiguous as possible.

Equipment: The piece was recorded onto a Macintosh IIci computer using Performer 3.61 software, was temporally quantized and had intensities equalised; it thus represented the musical score but embodied no irregularities of timing and intensity that could have acted to segment or differentiate elements of the piece in perception (Clarke, 1985; Sundberg, Friberg & Fryden, 1991). In the experiment the piece was output via MIDI to a Yamaha Disklavier computerised piano so that subjects experienced the music on a "real" piano. The intensity of the piece as performed corresponded to a moderate mf marking. The piece was performed at a tempo of two crotchets per second. A Yamaha RX5 rhythm box was used to record the subjects' responses.

Procedure: Subjects were informed that they were going to hear a short piece for piano. Their task was to indicate those elements of the piece playing the role of cue. This term was not used explicitly in the instructions; subjects were asked instead to select the events they found most salient, those which might be considered as points of reference or "landmarks", a metaphor more-or-less comparable to that suggested by Dowling (1989, p. 325). In order to accomplish this task, they were asked to imagine that they were taking a walk in an unfamiliar forest and were picking out "landmarks" in order to help them plan a future walk. They could consider any aspect of the piece to be a "landmark", or point of reference. They indicated their choices by pressing on a button on the rhythm box, as demonstrated by the experimenter. They were asked to respond as soon as they felt that they had identified a reference point. They were not given a prior hearing of the piece, but gave their response on first hearing. The procedure was then repeated. After the second hearing, the sections containing the landmarks pointed out by the subjects were played again. The sections began two seconds before the location of the landmark. The subjects were required to delineate, as precisely as possible, the reference point that the extract contained, i.e. to indicate the exact duration (beginning and end) of the selected cues by holding down a button of the rhythm box for the complete duration of the extract and to explain verbally why they found it salient. Subjects were allowed to describe the features as precisely as possible in their own words; although they were not required to use musical terminology they responded with sufficient specificity to enable the experimenters to relate their responses to precise elements of the piece[iv].

Finally, for each subject, the extracts containing the reference points that had been identified were presented aurally a second time and the subject was asked to evaluate their importance on a five-point scale ranging from "highly important" to "of little importance", each reference point being identified by a different box on a response-form. The reference points were presented in a different random order for each subject. A reference point was played and subjects assigned it to the response-category (box) of their choice, using letters of the alphabet: "A" for the first reference point played, "B" for the next, etc. They were allowed to assign as many reference points as they wished to any one category [v].

Results and discussion

Table 1 gives the reported frequency of occurrence, the mean and standard deviation of the evaluations of cue-value and the number of the confirmations (i.e., the number of times the same cue was abstracted in both trials) of the cues abstracted by the subjects.

	First Hearing			Second Hearing			
Cue	Frequency	Mean	SD	Frequency	Mean	SD	Confirmations
I	6	4.67	0.52	7	3.90	0.88	6
II	5	4.20	0.84	4	3.75	1.50	3
III	4	4.50	1.00	4	3.25	1.26	4
IV	4	4.75	0.50	2	4.00	1.41	2
V	1	-	-	5	3.00	1.58	-
VI	1	-	-	2	4.50	0.71	1
VII	1	-	-	2	2.00	-	-
VIII	-	-	-	2	1.50	0.71	-
IX	1	-	-	-	-	-	-
			-				

Cue labels and descriptions
I = Shift to a higher register, measure 4
II = Anacrusis prior to measure 9
III = Introduction of accompaniment, measure 1
IV = Caesura tied to longest inter-event time-interval of piece, measures 10 and 12
V = Occurrence on downbeat of densely-textured triads, measures 10 and 12
VI = Repeated motif, measures 8 to 12
VII = Final cadence, measures 15/16
VIII = Ascendant melodic movement , measures 12 to 16
IX = Central cadence on dominant and inferred rallentando, measures 7/8

Table 1: Observed frequency, the mean and standard deviation of the evaluations of cue-importance given in the first and second hearings, and the number of confirmations on second hearing of the

[iv] For example, cue I has been described as contraste grave/aigu; cue II as des sons annonciateurs; cue III as entrée de sons graves; cue IV as temps vide; temps creux; cue V as accords pesants; plus de poids, plus dense; etc. thus terminologies clear enough to be summarized as proposed in Table 1.

[v] This procedure represents a development of an exploratory procedure used by I. Deliège (1991) in which four pieces or excerpts from larger pieces for solo instrument (by Wagner, Stravinsky, Debussy or Schubert) were presented to the subjects.

cues abstracted by the non-musicians in Experiment 1 (Version A of the Schubert's Valse). Subjects could give more than one response.

It appears that four cues were judged most important (see Figure 2): the contrast of register at bar 4 (cue I); the anacrusis prior to bar 9 (cue II); the introduction of the accompanimental part (cue III); and the caesura associated with longest inter-event time-interval of the piece at bars 10 and 12 (cue IV). The structural accent on the chord of the dominant at bars 10 and 12 (cue V) were noted only rarely on the first trial though more frequently on the second.

Figure 2: The figure indicates the nine different cues (labelled I to IX) that were abstracted by subjects in Experiment 1 over two hearings of Version A of Schubert's Valse sentimentale, D 779, Op. 50 No 6.

It has been shown that change of register is a powerful means of segmentation in music (I. Deliège, 1987b; see also Krumhansl & Jusczyk, 1990); in the present piece the change of register at bar 4 was particularly salient. Although there are anacruses at bars 0, 4, 8, 10, 12, subjects appeared sensitive only to that leading into bar 9. There appear to be several possible bases for this. Subjects may have required more time than was available to become aware of the anacruses that permeate the entire piece (which might explain why the anacrusis at bar 0 was not noted) while the salience of the anacrusis in bar 4 can be held to be reduced by the simultaneous change in register, thinning of the texture of the upper part and integration of the anacrusis into the metrical framework of the piece by its co-occurrence with the accompaniment. Again, at bar 12, the texture is less dense. The perceptual functionality of this change in textural density is prefigured in the results of I. Deliège (1987b) which indicate that it is one of the most powerful cues in the segmentation of short musical excerpts.

The fact that the melodic ascent functions only weakly as a cue through bars 12 to 16 (being noted only at the second trial as cue VIII) is consistent with previous results. Deliège (1992c) showed in an experiment on Wagnerian leitmotivs that motifs reliant on melodic contour, like the Vertrags-Motiv from Rheingold were more difficult to memorise and to recognise than motives incorporating other cues, such as the Riesen-motiv. Mélen and Deliège (1995) present further evidence showing that "global" melodic contour is a weaker basis for recognition of melodies than are more local cues.

The subjects indicated the existence of 23 cues in the first trial and 28 in the second. A series of binomial tests indicated that there were no significant differences between the frequency distributions of any cue on first and second trial. A series of Wilcoxon tests were undertaken to test for differences between the mean evaluations of the cues at first and second trial. These tests were applied only if a cue had been chosen by more than one subject, but all were non-significant ($p>.05$).

These results indicate that subjects were generally consistent from one trial to the other. The four main cues in the first trial remained equally important in the second trial, even if the cue based on the caesurae in bars 10 and 12 showed a slight decrease in efficacy. It appears that when the frequency of identification changes from one trial to the other it tends to increase. The mean evaluations of cue importance remained relatively stable from one trial to the other.

Although a very few of the cues that were abstracted - like the final cadence or the central cadence to the dominant - are primarily harmonic in origin, most of the cues appear at the level of the musical surface: change of register, change of density, and anacrusis (cf Pollard-Gott, 1983). Thus the musical surface appears to have governed subjects' responses (for similar conclusions, see Krumhansl, 1995). In order to clarify the issue of this sensitivity to the cues of the musical surface, an experiment using the mental line procedure was conducted here, again on non-musicians. The mental line procedure was applied using the same piece by Schubert (Version A, see Figure 1).

EXPERIMENT 2 - MENTAL LINE PROCEDURE

Method

Subjects: Eight non-musicians (between the ages of 18 and 24, mean 20.9, otherwise similar to those who had participated in Experiment 1) took part in the experiment. None of them had taken part in the previous experiment.

Material: The same piece as was used in Experiment 1 (Version A) was used in this experiment. Again, the repeat marks in the original score were not observed so as to ensure that the order in which events occurred in the piece was as unambiguous as possible. This experiment (in common with the puzzle experiment) required that subjects be presented with the material in segmented form. A number of different segmentations were considered; the structure of the piece allows several distinct possibilities for segmentation, depending on whether more importance is given to the melodic, rhythmic, harmonic and metric organisation of the piece. A segmentation into four four-bar units was rejected as subjects' error-rates would have been likely to be too low for any adequate inferences to be made about the bases for their performance.

It was decided that partitioning the piece into eight segments on the basis of metrical and harmonic structure would be most appropriate. Examination of the rhythmic and melodic structure of the piece indicated that a division into two-bar anacrustic segments would be possible throughout; however, this possibility was rejected in favour of the segmentations shown in Figure 1. The piece is explicitly composed as a Valse, within which the coincidence of the metrical accents on the downbeat of each bar with the change in harmony that occurs at that point in each two-bar group would seem to imply a segmentation that affords these features primacy within the grouping structure of the piece. This segmentation leads to unequal-sized groups, some being anacrustic and some crustic; however, it was felt that this feature might have a positive benefit for subjects in differentiating one group from another.

Moreover, this division of the piece is in accordance with that which conventional music theory might propose on the basis of the tonal, phrasal and metrical functions of the elements within each group and sub-group (see, e.g., Prout, 1893). Such a division is also in accordance with more recent and more formal theories such as that of Lerdahl and Jackendoff (1983), both in terms of their Grouping Structure (in, for example, selecting as groups patterns that are structurally parallel, such as segments 5 and 6 within the third four-bar phrase of the piece) and their Time-Span Reduction (in preferring cadential structures as group determinants, as under Time-Span Reduction Preference rules 6, 7 and 8). Such a segmentation also receives empirical support from Deliège (1987b), who found that both musicians and non-musicians tended to form group boundaries between points of harmonic change; as noted above, a harmonic change occurs between each of the two-bar sub-groups that had been postulated on the basis of analytic or theoretical considerations. Whilst a reliance on any single source of evidence might seem to permit of an undue degree of ambiguity in choice of segmentation, the convergence of the partitioning generated by "conventional wisdom", recent theory and empirical evidence would seem to give validity to the segmentation imposed.

<u>Equipment</u>: The same equipment was used as in Experiment 1.

<u>Procedure</u>: The subjects were informed that they would hear a short piece for piano, followed by the same piece but presented in eight segments in random order. Their task was to indicate the location of these segments in the original piece. They listened to the piece three times before hearing the segments. The piece was symbolised by a horizontal line of eight boxes representing its temporal development. On hearing the first segment, subjects were asked to write the number "1" in the box that they felt to be correct for that segment, on hearing the second to write the number two in the appropriate box, and so on. They were not obliged to fill all the boxes, and were allowed to insert a segment between two segments already positioned if they felt this was appropriate so that responses were not limited by those previously made. This procedure followed that used in previous studies (Deliège, 1991, 1993), and was employed in order to ensure that judgment of position of any specific segment was independent of judgments made in respect of other segments, allowing the application of parametric statistics in the analyses of results. Only the final ordering of segments figured in the analysis of results.

After completing the task, they listened to the "piece" they had created. When listening to their piece, they were asked to push a button on the Yamaha rhythm box to indicate when they felt there was an error. They were then asked to produce a verbal account of their reasons for noting errors, which was taken down by the experimenter. The task was then undertaken a second time. This time, the subjects were allowed to listen to the piece as often as they liked before hearing the segments, which were presented in a different random order.

<u>Results and comments</u>

As the experiment required the subjects to locate the segments on a horizontal line, it was necessary to analyse the location accorded by the subjects to the different segments. One sample t-tests were conducted on a variable defined as "distance", i.e. the difference between the real position of the segment and its mean position attributed in the subjects' responses. In this analysis, a correctly located segment would have a mean distance of 0, the mean distance could thus vary between 0 and 7. This analysis is reported in Table 2 which shows the segments in their correct temporal order from 1 to 8, and gives for each segment: 1) the primary mode, i.e. the location most frequently chosen by the subjects; 2) the secondary mode, i.e. the second most frequent location; 3) the mean location; 4) the value of t for each segment (7 df, 2-tailed) and its significance in respect of one sample t-tests conducted on the variable "distance" (in this analysis, a correctly located segment would have a mean distance of 0). As segments 5 and 6 were identical, a segment localised to the fifth or in the sixth position was given a "location" of 5.5.

In order to give a clear picture of the patterns of subjects' responses, Tables 3a and b show for each segment and in each trial the frequency of attributed location according to subjects' responses.

	First trial					Second trial				
Seg.	Mode	Mean	Mean distanc	t-value	p	Mode	Mean	Mean distance	t-value	p
1	**1**/3	2.19	1.19	2.10	.07	**1**/7	3.31	2.31	2.37	.04
2	2-7	3.94	2.31	2.99	.02	**2**/8	4.69	2.44	2.77	.02
3	2-3-5.5	4.50	2.00	3.09	.01	**2**/4	3.06	1.06	4.43	.003
4	**5.5**/4	4.87	1.62	3.39	.01	2-3-4-5.5	3.62	1.12	4.02	.002
5	**5.5**/4	4.56	1.56	2.78	.02	**5.5**/4	5.12	1.00	2.49	.04
6	**8**/1-7	5.69	2.44	4.53	.002	**5.5**/7	6.19	0.68	1.94	.09
7	7-5.5	5.25	2.00	3.09	.01	1-3-5.5	3.75	3.25	4.23	.003
8	**8**/5.5-4	5.75	2.25	3.13	.01	**8**	5.87	2.12	1.97	.08

Table 2: For the eight segments in both trials, the primary (bold numbers) and secondary modes of location, the mean location chosen for each segment, the mean distance from the correct location and the the results of t-tests conducted on segment "distance" (the difference between the real position of the segment and its mean position attributed in each subject's responses) with significance levels, for the results of Experiment 2 (non-musician subjects, Version A of Schubert's Valse).

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1(4)	2(2)	3(2)	5/6(4)	5/6(3)	8(3)	7(2)	8(3)
3(2)	7(2)	2(2)	4(2)	4(2)	1(2)	5/6(2)	5/6(2)
2(1)	1(1)	5/6(2)	1(1)	3(1)	7(2)	2(1)	4(2)
5/6(1)	3(1)	7(1)	8(1)	1(1)	5/6(1)	3(1)	3(1)
	4(1)	8(1)		8(1)		4(1)	
	5/6(1)					8(1)	

Table 3(a): The eight possible locations of segments (bold characters).within Experiment 2 (non-musicians, mental line, first trial, Version A of Schubert's Valse) Numbers below indicate which segments were localised by the subjects to each position, as well as the number of subjects (in parentheses) who attributed this location to the segment.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1(4)	2(3)	2(4)	4(2)	5/6(4)	5/6(5)	5/6(2)	8(5)
7(2)	8(2)	4(2)	5/6(2)	4(2)	7(2)	1(2)	1(1)
3(1)	3(1)	3(1)	2(2)	3(1)	8(1)	3(2)	2(1)
5/6(1)	5/6(1)	5/6(1)	3(2)	8(1)		4(1)	4(1)
	7(1)					7(1)	

Table 3(b): The eight possible locations of segments (bold characters).within Experiment 2 (non-musicians, mental line, second trial, Version A of Schubert's Valse) Numbers below indicate which segments were localised by the subjects to each position, as well as the number of subjects (in parentheses) who attributed this location to the segment.

First trial: As shown in Tables 2 and 3(a) and (b), segments 1, 4, 5, and 8 gave rise to a primary mode. However, the mode for segment 4 was 5.5 - thus not the actual location - with 4 as the secondary mode. For segments 5 and 8, the mode appeared to correspond to the actual location, but the weakness of these modes (only 3 responses for both locations) is supported by the significant results for the t-tests on the "distance" variable for those segments. Except for segment 1 all mean distances were significantly different from 0.

Second trial: Compared to the first trial a general improvement in performance was observed. Seven segments out of eight gave rise to a mode, which corresponded to the actual location in five cases (for segments 1, 2, 5, 6, and 8). Moreover it must be noted that this time the modes which were recorded had at least four or five responses. Nevertheless, the mean distances reached significance, except those of segments 6 and 8.

The previous results were further analysed by comparing the mean distances between the actual locations and the recorded responses for each segment. A single repeated-measures ANOVA was carried out on the results of each trial, taking segments as within-factor. A two-way repeated-measures ANOVA compared the results of the two trials.

First trial: Distance from the actual location did not appear to vary as a function of segment ($F<1$). This analysis confirms the conclusions drawn from the t-tests which were significant for seven segments out of eight and marginally significant for segment 1 ($p=.07$).

Second trial: The effect of segment was marginally significant ($F(7,49)=2.07$, $p=.064$). This finding can be attributed to the results for segments 6 and 7; segment 6 was considered to have occurred at its actual location, while segment 7 was mis-located as occurring earlier in the piece. The results of this analysis point up features evident in the frequency distribution shown in Table 2. While in the first trial there was no mode for segment 6, a primary mode corresponding to the actual location emerged in the second trial; indeed the t-test was highly significant in the first trial but did not reach significance in the second (see Table 2). For segment 7, while there was no mode in either trial the distribution of responses was more dispersed in the second trial, as evidenced by the values of the t-tests and the corresponding probabilities (see Table 2).

Between-trials analyses: Despite the latter results, the two-way ANOVA did not produce a significant interaction between segment and trial, but only a marginal effect ($F(7,49)=1.889$, $p=.09$). Other effects were non-significant (for segments: $F(7,49)=1.218$, $p>.31$, for trials: $F<1$).

Cluster analyses were run on the data (complete-linkage method, farthest neighbour). The outputs of the process were tree-diagrams showing the "nested proximities" of the segments in the subjects' judgments, indicating the degree to which any segment was related to all other segments in the subjects' judgments of location within the piece. The resultant tree diagrams are shown in Figures 3(a) and 3(b) below.

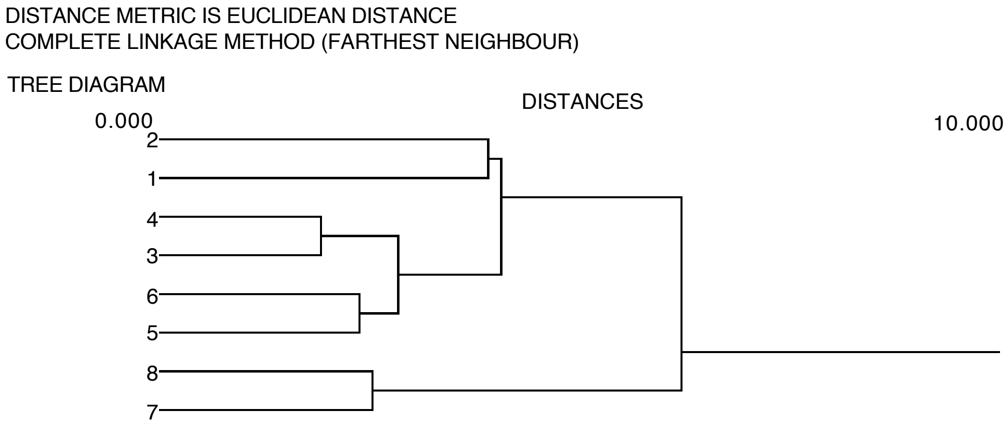

Figure 3a

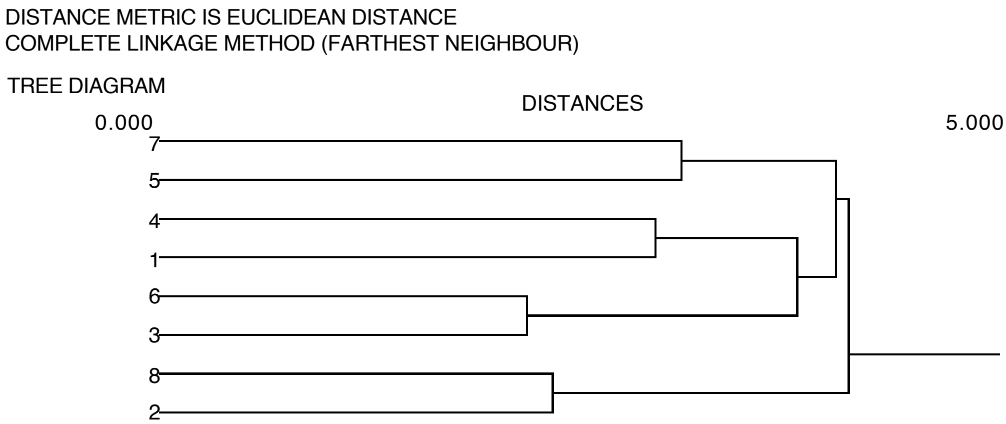

Figure 3b

Figures 3(a) and 3(b) show data obtained from the data of Experiment 2 by a clustering method (complete linkage, farthest neighbour). They illustrate the inter-segment proximities experienced by the non-musician subjects after the first hearing (Figure 3a) and after the second hearing (Figure 3b) of Version A of Schubert's Valse.

These trees indicate that the subjects' representations of the piece had improved in accuracy on the second hearing. In the tree-diagram of the responses in the second trial (Figure 3b), segments that exhibit greater proximity are, in general, adjacent or near-adjacent. In both tree-diagrams subjects appear to divide the piece into two clusters; in Figure 3(b) more segments are integrated into a large cluster, which can be read as an indication of an increase in the subjects' tendency to represent the piece in cognition as an entire entity. Finally, the distances between the segments in Figure 3(b) - representing their perceived similarity in location in the piece - are smaller, again indicating an enhanced representation of the piece as a whole.

Discussion

Altogether, these results indicate that non-musicians' representations of the temporal course of music are characterised by a degree of uncertainty. The mean location attributed to segments by the subjects shows a tendency to concentrate the segments toward the centre of the piece. While its actual span can be considered as ranging from 1 to 8, the observed span was located between 2.375 and 6.562 in the first trial and from 2.5 to 6.625 in the second. Thus the observed span tended to be shorter than the real duration of the piece. Similar results were obtained by I. Deliège (1991, 1993) in respect of a piece by

Boulez, particularly for non-musician subjects, and by Clarke and Krumhansl (1990) for musicians. Given that segments 1, 4, and 8, i.e. the main points of articulation of the tonal scheme, were in general incorrectly located, the tonal functions of the segments in respect of the global structure of the piece did not appear to impact significantly on subjects' responses. Rather than structural features predominating, superficial characteristics seem to have been more important determinants of subjects' responses. The results of the first trial indicate the possible operation of a primacy effect classically found in memory for unstructured lists (see, e.g., Postman and Phillips, 1965), in that accuracy of localisation was greatest for the first segment (distance from correct location being only marginally significant). However, this weak effect is not present in the results of the second trial; here, accuracy is greatest for segments 5 and (particularly) 6.

These findings might be accounted for in terms of cue-abstraction processes; segments 5 and 6 are associated with cues IV, V and VI (see Table 1), are identical, and their material is thus presented more frequently than any other in the piece. While it appears more difficult to account for other aspects of these findings in terms of cue-abstraction processes, it may be hypothesised that this arises from the abstraction of cues that are not localised <u>within</u> segments but instead occur <u>between</u> segments (see Table 1, particularly cues I and II), leading to a degree of confusion in manipulating regions of the piece in memory. This argument finds support in the fact that for several segments the attributed location often migrates to adjacent positions; for instance, in first trial, segment 2 was localised to positions 1, 2, or 3 five times. Another instance is evident in the results for segment 4; in the first trial, the mode for segment 4 was 5.5 with 4 as the secondary mode. On the second trial, the mode for this segment was 3 with 4 as secondary mode. Altogether, these results indicate that the central position of this segment was perceived but that its position in relation to the exact mid-point of the piece was not clearly identified.

These results appear to show that subjects seem to have registered in memory the approximate location of events in the piece rather than their precise chronological order. The non-musician subjects' performance improved over the two trials, a result consistent with the results of other studies (I. Deliège, 1991; 1992c; 1995; in press) in that the correspondence between the localisation of segments by subjects and their actual positions improved in the second trial even if the degree of "mis-localisation" remained significant. This improvement in performance with number of hearings may be explained in terms of either the nature of the abstracted cues or the degree to which they were bound to specific locations within the piece. Such an improvement was also observed in a study on categorisation performances for non musicians, but occurred to a lesser degree for high trained professional musicians (Deliège, 1991; in press). Musician subjects studied by Krumhansl (1995) did, however, yield improvement with repeated hearings. However, the musical competence of her subjects might be thought of as not comparable to that of professional musicians. Despite the fact that one of her subjects had 20 years tuition, most had consdierably less (median number of years of instruction=12). Such contrasting results suggest that more empirical studies are needed to elucidate the relationship between musical expertise and the benefits drawn from repeated hearings. This might focus in particular on the way in which listeners differing in level of expertise memorize a piece. Some data already available suggest that the level of expertise influences the accuracy of the memorisation of the cues.

Previous studies have shown that although non-musicians were able to abstract cues and did not need several hearings to do it (I. Deliège, 1993; 1995), the cues they abstracted were less potent than those abstracted by musicians (I. Deliège, 1989, 1990, 1992a & c). Cues abstracted by musicians were longer and more closely tied to structural attributes of the piece than those abstracted by non-musicians, which tended to be shorter and more likely to derive from more superficial characteristics of the music (cf Pollard-Gott, 1983; I. Deliège, 1993). Cues abstracted by non-musicians may be thus less stable and more susceptible to degrade rapidly in memory than those employed by musicians (despite the fact that, when asked in the cue-localisation experiment to explain what were the cues in a given piece, subjects were generally able to describe them properly and the most salient cues were often confirmed at a second listening[vi]).

Alternatively, the improvement in subjects' performance from first to second trial might be dependent on the degree to which subjects experienced cues as being tied to specific locations in the piece. In the present experiment, subjects knew that their task was to reorder segments from the piece they had heard and that they had to remember not only the segments but also their location. It could be argued that, for these non-musician subjects, the links between already-formed cues and their locations become weaker as new cues were encoded. Thus non-musicians could need more hearings to be better able to remember effectively the location of the cues that they had encoded, particularly when similar cues were present at different places in the piece (as was the case in the current experiment).

In general, there seems little evidence in these results for the sorts of sensitivity to tonal function that might be expected on the basis of the findings of Bigand (1990, 1993a and b) or Krumhansl (1995). However, some support for the emergence of such a sensitivity with increased exposure might be found in the change in locations attributed to segments 7 and 8 from the first to the second trial. The increased dispersion of responses for segment 7 in the second trial displays some coherence in that following the second hearing of the piece, segment 7 was more likely to be localised to positions such as one or three with which it shares functional tonal characteristics than to the locations of functionally dissimilar segments. Nevertheless, it must be emphasised that these segments also share very similar surface cues (see below, subjects' comments). Again, the recency effect evident in the enhanced performance for segment 8 in the second trial (where distance is not significantly different from actual location) might be taken to indicate the emergence of a sensitivity to structural attributes of the piece with repeated hearing (given that presentation of more structured material has been shown to lead to an increased recency effect in other domains of memory, see, e.g., Glanzer and Razel, 1974).

Against this there is little evidence from the comments of subjects themselves that they were aware of any such sensitivity. After each completion of the mental line task, subjects were asked why they had the impression of having made errors and how they thought they had produced their piece. On the first trial, three subjects stated that they had used contrasts of register (they tried to "repérer les graves et les aigus" or remarked that at the beginning the accompaniment was not very obvious, but that the high notes "sont plus claires"). The repetition of segments 5 and 6 was mentioned by three subjects, as well as the similarity between segment 7 and segment 1 ("ça recommence comme au début", and, interestingly, "ce n'est pas tout à fait les mêmes notes mais c'est presque la même chose"). Two subjects were totally unable to explain how they proceeded to construct their pieces. The answers to the questions as to why the button of the rhythm-box had been pressed (indicating a perceived error) were not very informative, tending to take the forms of remarks such "ça ne sonnait pas bien".

On the whole, these remarks indicate that these non-musician subjects had little or no explicit access to any cognitive schemata embodying declarative musical knowledge of the structural function of the components of a piece of tonal music. Nevertheless it remains possible that they might have been drawing implicitly on such knowledge in making judgments (cf Johnson-Laird, 1983, p. 466). However, given that subjects appeared to be

---

[vi] However, I. Deliège (1992c) showed that the superficial character of a cue was predictive of its efficiency and stability in memory, the presence of, for example, clear rhythmic grouping organization and/or accents being more immediately relevant. However, "global" melodic contours that were not easily divisible into small rhythmic groups were less efficient (cf Edworthy, 1985).

retaining relatively superficial characteristics of the piece in memory, there would seem to be little necessity to postulate an effect of any tacit declarative knowledge of tonal relations. The next experiment, using the puzzle procedure, was intended to shed more light on this sensitivity.

Both non-musicians and trained musicians were used in this experiment in an attempt to clarify the degree to which the cognitive representations of pieces of tonal music formed by non-musician subjects depended on superficial or structural features. The puzzle procedure was conceived as a counter-example of the mental line (I. Deliège & Mélen, in press); if the function of temporal landmarks of the cues collected during listening enable subjects to elaborate a mental representation of the piece, it should be very difficult, at least for non-musicians, to build a piece starting from separated chunks, without having received any reference listening before. In order that the same material could be used for both subject groups, it was necessary to evolve an experimental procedure that would yield neither a ceiling effect for the musician group nor below-chance responses from the non-musicians. Accordingly, all subjects were required to recreate the most coherent piece that they could using the collection of segments employed in the previous experiment. It was expected that non-musicians' responses would be likely to be determined by the immediate, superficial characteristics of each segment, leading to relatively incoherent pieces; musicians were expected to be able to apply their declarative knowledge of tonal function so as to at least maintain structural coherence even if they did not re-create the original piece.

EXPERIMENT 3 - PUZZLE PROCEDURE

Version A

Method

Subjects: Ten musicians (between the ages of 19 and 21, mean 20.7) and ten non-musicians (between the ages of 18 and 22, mean 20.6) took part in the experiment. The musicians were formally-trained within the Western classical tradition (mean number of years of formal musical training=12 with a maximum=16 and a minimum=10) and were graduate and undergraduate students in the Faculty of Music at the University of Cambridge. The non-musicians were students of Psychology at the University of Liège. None of them had received formal music tuition.

Material: The same piece by Schubert and the same segmentation as in Experiment 2 was used.

Equipment: The same equipment as in Experiment 2 was used except for the Yamaha RX-5 rhythm box.

Procedure: The subjects were required to create the most coherent piece of music they could using the eight sub-groups or segments. They did not listen to the piece prior to undertaking the task. Icons symbolising each segment were displayed simultaneously on the screen as "chunks" under Performer software; all icons were identical other than in being labelled "a", "b", "c", etc., (the label in no way denoting the temporal position of each segment within the piece). To create a piece, they simply needed to move these icons and arrange them in a linear order (in a "Song" window under Performer). Subjects were first familiarised with the computer and shown how to arrange the segments themselves (how to move, erase or to replace an icon). The labels that indicated the original ordering of each segment in the original piece were masked and the segments displayed in a random arrangement that was different for each subject, who were told that they must use all the segments but may only use each segment once. They were told that the segments made up an entire piece (and not simply a part of a piece); no information about the composer, style or the historical period of the original piece that could have influenced the subjects in performing their task was given in the instructions. They were allowed to listen to individual segments as often as they liked before and during the creation of a piece, and could listen to their piece as often as they liked during its construction. They were allotted thirty minutes to complete the task and were permitted to make as many changes as they liked during this time.

Results

The same methods were used to analyse the results of this experiment as were used for those of Experiment 2. Table 4 gives primary and secondary modes, means, mean distances from correct location and the results of t-tests conducted on the distance from correct location for each segment. Table 5(a) shows for each segment the frequency of positions within the piece given by musician subjects, while Table 5(b) gives the same information for non-musicians.

	Musicians					Non-Musicians				
Seg.	Mode	Mean	Mean Distanc	t-value	p	Mode	Mean	Mean Distance	t-value,	p
1	**1**	2.00	1.00	1.63	.13	**7**/5.5	5.15	4.15	6.18	.0002
2	**2**	3.75	1.75	2.19	.05	**7**	5.45	3.65	5.77	.0003
3	7-5.5	5.50	2.70	6.19	.0002	**1**	3.15	1.55	4.04	.002
4	**4**	4.00	0.60	2.34	.04	**8**/4	5.45	2.05	3.60	.005
5	5.5-3	3.85	1.95	3.83	.004	1-3-5.5	4.40	2.40	4.81	.001
6	**5.5**	4.25	1.25	2.68	.02	**2**	3.10	2.40	4.81	.001
7	**7**/5.5	5.55	1.45	2.24	.05	**5.5**	3.80	3.20	4.55	.001
8	**8**	7.15	0.85	1.36	.20	**5.5**	5.50	2.50	4.66	.001

Table 4: Primary (bold numbers) and secondary modes of location, mean location chosen for each segment, mean distance from the correct location and t values with significance levels for the distance from correct location for each segment, for the eight segments for both musician and non-musician subjects in Experiment 3 (Version A of Schubert's Valse).

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
7(4)	7(3)	1(3)	8(4)	1(2)	2(4)	5/6(4)	5/6(6)
5/6(3)	4(2)	3(2)	4(3)	3(2)	4(2)	1(2)	8(2)
3(2)	8(2)	4(2)	2(1)	8(2)	5/6(2)	2(2)	2(1)
1(1)	1(1)	2(1)	3(1)	5/6(2)	1(1)	3(1)	4(1)
	5/6(1)	5/6(1)	5/6(1)	2(1)	3(1)	7(1)	
	3(1)	7(1)		7(1)			

Table 5a (non-musicians) shows the eight possible locations of segments (bold characters).within Experiment 3 (Version A of Schubert's Valse) Numbers below indicate which segments were localised by the subjects to each position, as well as the number of subjects (in parentheses) who attributed this location to the segment.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1(7)	2(6)	7(4)	4(6)	5/6(3)	5/6(5)	7(5)	8(8)
3(2)	8(2)	5/6(4)	2(1)	3(3)	2(2)	5/6(3)	2(1)
7(1)	4(1)	2(1)	3(1)	1(2)	3(1)	3(1)	5/6(1)
	5/6(1)	3(1)	5/6(2)	4(1)	4(2)	1(1)	
				7(1)			

Table 5b (musicians) shows the eight possible locations of segments (bold characters).within Experiment 3 (Version A of Schubert's Valse) Numbers below indicate which segments were *localised by the subjects to each position, as well as the number of subjects (in parentheses) who attributed this location to the segment.*

Non-musicians: The difference between the actual location of the segments and the chosen location in the responses was highly significant for all segments (see Table 4), seemingly demonstrating that non-musician subjects possessed little capacity to produce coherent tonal structures. However, Table 5(a) appears to indicate that these subjects tended to produce structures in which segments that are odd-numbered in the original piece are placed in odd-numbered locations, even-numbered segments appearing in even locations, but without much apparent regard for the preservation of tonal coherence across consecutive odd-even segment-pairs. The order produced by seven of the subjects are very similar; these subjects put a cadential segment at the mid-point of their piece and began the second part with a tonic segment. However, this fact alone does not suffice to demonstrate an ability to confer a well-formed structure on the piece. In fact, of all the non-musician subjects only three used segment four as the cadential segment at the mid-point of the piece. Moreover, segment eight is seldom put at the end of the piece (by only two out of the ten non-musician subjects). The general distribution of the frequency of placement of segments at each position in non-musicians' pieces shows that they tend to use cadential segments early in the piece, resulting in a dearth of cadential segments within the latter half of the piece. A positive point was observed in their responses, that is that no subject begun his (or her) piece with a cadential segment, showing that a sense of cadence appears to be functional for non-trained Western ears (Bharucha & Todd, 1991). In general, non-musicians did not know what to do with segments 5 and 6; these segments were isolated in the pieces of six out of the ten subjects, mostly being positioned at the end of the piece. It appears that these subjects localised the other segments and, unsure of the position or function of segments 5 and 6, placed these at the end of the piece.

Musicians: Table 5(b) shows that in the musicians' pieces the modes of most segments coincided with the correct position. Nevertheless, only the t-tests on distances for segment 1 and 8 were non-significant, indicating that musicians were aware of the initial and terminal functions of segments 1 and 8, but were less able to judge the potential functions of other segments. Table 5(b) shows that the tendency in the non-musicians' pieces (the positioning of odd-numbered segments at odd-numbered locations, etc.) is again evident, but more markedly so. These results are somewhat surprising; given that these musicians had received extensive training in harmony and in the classical repertoire it was expected that their capacity to "re-create" an archetypal tonal work of the classical period would be highly developed, but this appears not to have been the case. (In fact, the only musician subject who re-created the original piece did so well within in the time allotted, although he expressed some dissatisfaction with his creation and attempted to "improve" it by continuing to re-order the segments through the rest of the experimental session !)

Cluster analyses were run on the data from musicians and non-musicians separately (complete-linkage method, farthest neighbour), outputting tree-diagrams that show the "nested proximities" of the segments in the subjects' judgments, and indicate the degree to which any segment and segment-grouping is related to all other segments and segment-groupings in the subjects' judgments. The resultant tree-diagram for musicians is shown in Figure 4(a) below, and that for non-musicians in Figure 4(b).

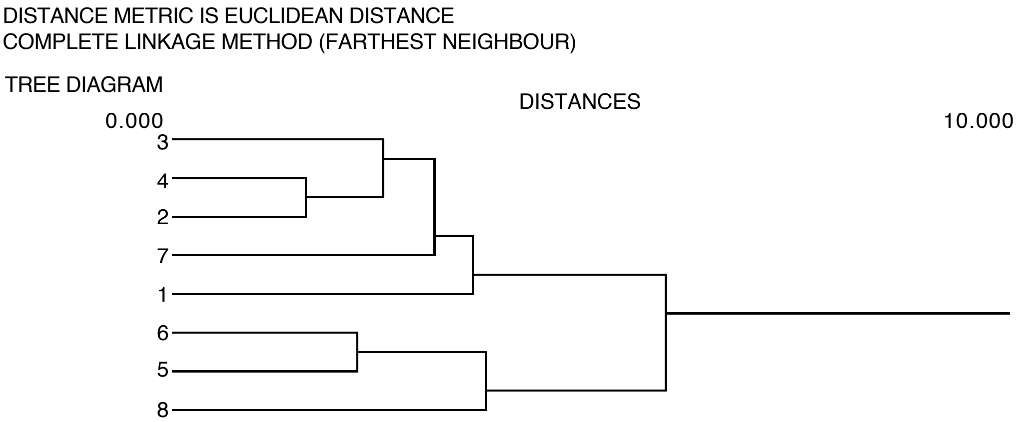

Figure 4(a)

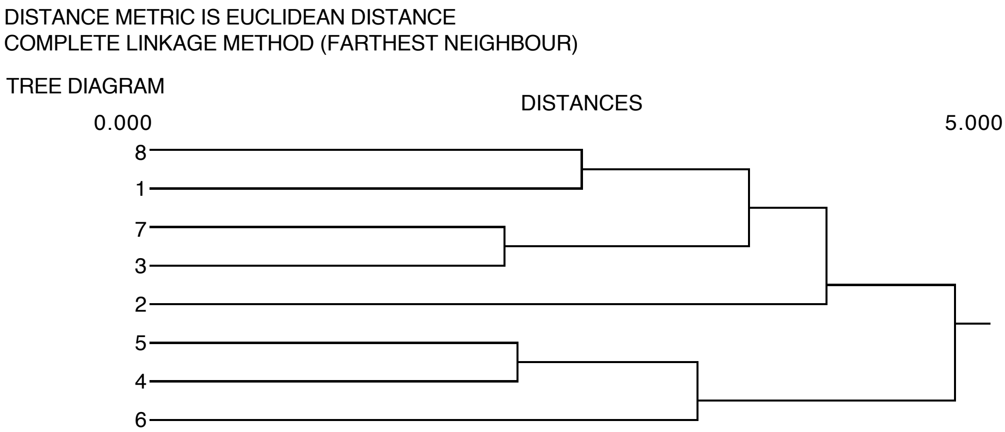

Figure 4(b)

Figures 4(a) and (b) were obtained from the data of Experiment 3 by a clustering method (complete linkage, farthest neighbour). They show the inter-segment proximities experienced by the musician (Figure 4a) and non-musician (Figure 4b) subjects for Version A of Schubert's Valse.

The tree-diagrams for musicians and for non-musicians are markedly different. The musicians tended to group segments into the pairings that appeared in the original piece, and segment-pairs are grouped coherently so as to conform to the phrase-structure of the original piece. The structures produced by non-musicians, conversely, had few if any of the characteristics of the original piece. However, the cluster analyses confirm the interpretations of the data and results given Tables 4, 5(a) and (b) above. Musicians were better able to judge appropriate locations for markers of larger-scale tonal structure (such as segments one and eight), while musicians, and to a much lesser extent, non-musicians tended to order segment-pairs in groupings that constituted "odd-even" pairs (musicians' pieces being more likely to preserve the odd-even segment-pair groupings of the original piece than were non-musicians').

On the basis of these results it appears that even non-musicians might be sensitive to tonal function; their slight tendency to use odd-numbered segments in odd-numbered locations, and to follow these with even-numbered, cadential, segments, seems to indicate that they may be sensitive to the tension-creating function of the prolongations generally present in the odd segments and the tension-releasing function of cadences (even segments), as would be implied by the results of Bigand (1990). However, their disregard for the maintenance of tonal coherence in odd-even segment pairs indicates that any putative sensitivity to tonal function is local and short-range.

An alternative explanation for these findings is, however, available. In this piece by Schubert the even-numbered, cadential segments close on notes of a relatively long duration (in the upper part). Thus rhythmic structure and harmonic structure can be seen as acting together to establish groupings within the original piece. Relative duration has been shown to act as a powerful determinant of grouping, even for young infants (see Krumhansl & Jusczyck, 1990; Jusczyck & Krumhansl, 1993), while metrical organisation appears to be a primary component of non-musicians' memories of musical works (Sloboda and Parker, 1985) and hence might act as a principal constraint on their manipulations of tonal materials. It could be argued that rhythmic and metrical characteristics of segments had acted as more powerful cues for non-musicians than did putative harmonic function in determining the appropriate ordering of segments; segment-pairs in non-musicians' pieces may thus have closed on even-numbered segments not because of the cadential function of such segments but because adherence to such a strategy would best preserve a periodically-grouped and metrically regular structure through the piece. In order to test this hypothesis a further experiment was conducted, again using the puzzle procedure, but excising the anacrustic figure that initiates each odd-numbered figure so as to enhance any possible grouping effect due to temporal factors. This editing was undertaken so as to preserve as much as possible of the materials of the original piece while strengthening rhythmic and metrical grouping cues. It was expected that if non-musicians' tendency to group odd and even segments was based on temporal rather than structural factors, they would exhibit this tendency even more strongly with the altered musical materials.

Version B

Method

Subjects: The same conditions were observed for the selection of the subjects as for Version A of the piece. There were 10 student musicians from Cambridge (between the ages of 19 and 23, mean 20.8) and 10 student non-musicians from Liège (between the ages of 18 and 20, mean 18.7) took part in the experiment. None of them had taken part in the experiment employing Version A.

Material, Equipment, Procedure: The B version of the piece - produced by excising the three-quaver upbeat figure occurring at the outset of each odd-numbered segment - was employed. The same equipment and procedure were used as for the original version of the piece.

Results

Non-musicians: All segments were given mean locations that differed significantly from actual position within the B version of the piece. Subjects' responses were similar to those based on the materials of the original piece in that they tended to use of all segments other than 5 or 6 at the outset of their pieces, placing these latter segments at the end of the piece. From Table 7(a) it can be seen that there was again a tendency for odd- and even-numbered segments to appear at, respectively, odd and even locations, although this Table does not unequivocally show that this tendency was more pronounced than in respect of the materials of the original piece.

	Musicians					Non-Musicians				
Seg.	Mode	Mean	Mean Distanc	t-value	p	Mode	Mean	Mean Distanc	t-value	p
1	**1**	2.90	1.90	2.11	.06	1-3-7	3.85	2.85	3.52	.006
2	2-5	3.75	1.75	3.00	.01	**8**	5.00	3.20	3.40	.007
3	**3**	3.10	0.90	2.65	.02	3-5.5	3.75	1.55	3.60	.005
4	**4**	4.20	0.40	1.92	.08	**4**/8	4.70	1.70	2.94	.01
5	**7**	4.90	1.80	4.07	.002	7/4-5.5	5.00	1.70	3.90	.003
6	**5.5**	5.00	0.50	1.37	.20	**5.5**/1-4	4.35	1.95	3.51	.006

7	3	4.00	3.00	4.31	.001	**5.5**/2-4	3.95	3.05	4.98	.0008
8	**8**	8.00	0.00	/	/	**5.5**	5.40	2.60	5.75	.0003

Table 6: Primary (bold numbers) and secondary modes of location, mean location chosen for each segment, mean distance from the correct location and t values with significance levels for the distance from correct location for each segment, for the eight segments for both musician and non-musician subjects in Experiment 3 (Version B of Schubert's Valse).

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1(3)	8(5)	3(3)	4(4)	7(4)	5/6(3)	5/6(3)	5/6(8)
3(3)	2(3)	5/6(3)	8(3)	4(2)	1(2)	2(2)	2(1)
7(3)	1(1)	2(2)	3(2)	5/6(2)	4(2)	4(2)	8(1)
5/6(1)	3(1)	7(1)	1(1)	1(1)	2(1)	1(1)	
		1(1)		2(1)	7(1)	3(1)	
					8(1)	7(1)	

Table 7a (non-musicians) shows the eight possible locations of segments (bold characters) within Experiment 3 (Version B of Schubert's Valse). Numbers below indicate which segments were localised by the subjects to each position, as well as the number of subjects (numbers in parentheses) who attributed this location to the segment.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1(6)	2(5)	3(5)	4(7)	7(4)	5/6(8)	3(3)	8(10)
7(3)	5/6(5)	2(2)	3(1)	1(1)	2(1)	1(2)	
2(1)		5/6(2)	5/6(2)	5/6(2)	4(1)	5/6(2)	
		1(1)		2(1)		7(2)	
				3(1)		4(1)	
				4(1)			

Table 7b (musicians) shows the eight possible locations of segments (bold characters) within Experiment 3 (Version B of Schubert's Valse). Numbers below indicate which segments were localised by the subjects to each position, as well as the number of subjects (numbers in parentheses) who attributed this location to the segment.

Musicians: From Table 6 it can be seen that segments 1, 4 , 6 and 8 tended to be placed in the correct locations (t-values all mean distances for other segments being significant). All subjects used segment 8 to close their pieces. These results reinforce the conclusions of the previous experiment in showing that musicians were aware of the global tonal functions of segments in the piece, at least in respect of the primary articulators of tonal function (the cadence to the dominant at the centre of piece, and final perfect cadence to the tonic). On the other hand, musicians' placement of segments other than 1, 4, 6 and 8 was variable; Table 7(b) shows that there was very little consistency in choice of segment to follow the cadence to the dominant in position 4. Nevertheless, musicians again tended to locate odd segments at odd locations and even at even positions. On the whole, results from the experiment using the original materials and from the present experiment point in similar directions

Again, cluster analyses were run on the data from musicians and the non-musicians separately (Complete-linkage method, farthest neighbour). The resultant tree-diagram for musicians is shown in Figure 5(a) below, with that for non-musicians in Figure 5(b).

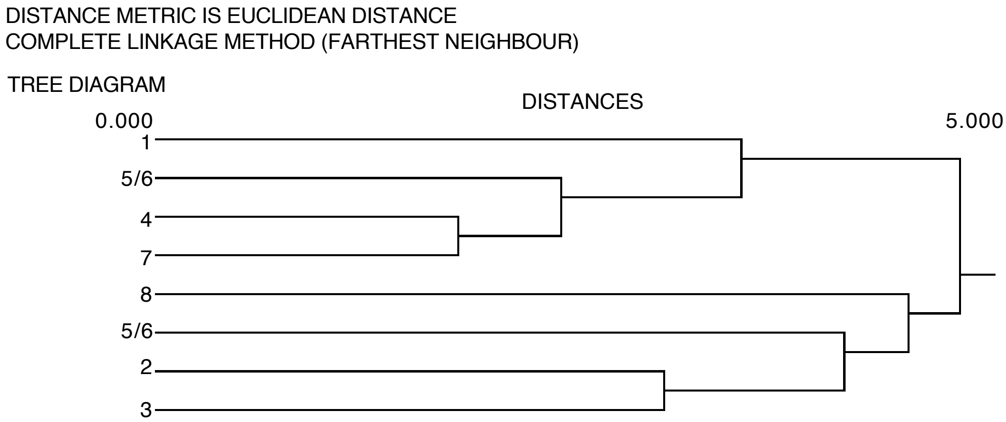

Figure 5(a)

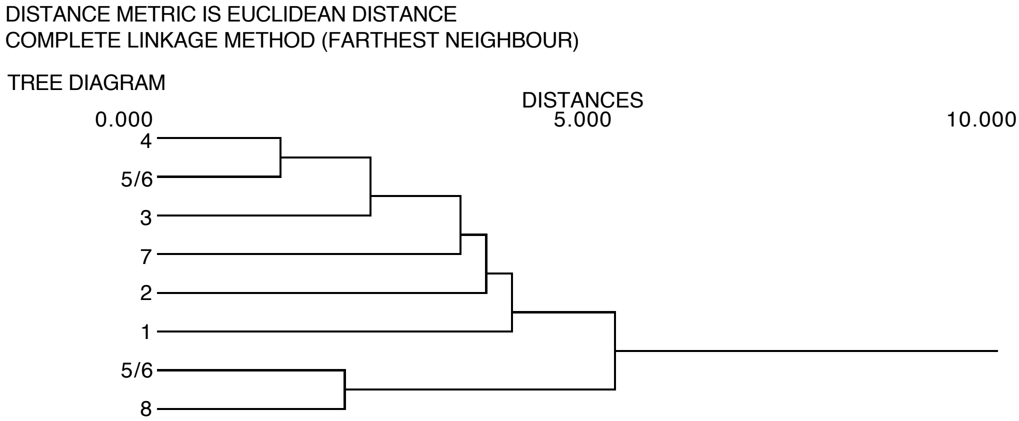

Figure 5(b)

Figures 5(a) and 5(b) were obtained from the data of Experiment 3 by a clustering method (complete linkage, farthest neighbour). They show the inter-segment proximities experienced by the musician (Figure 5a) and non-musicians subjects (Figure 5b) for Version B of Schubert's Valse.

Removal of the anacrustic figures appears to have considerably reduced the musicians' ability to "re-generate" the relationships within the piece as originally structured (relations between segments 1, 3 and 7 and all other segments appeared particularly unclear). However, their tendency to produce structures that were tonally normative at higher structural levels appeared to remain unaltered (compare with Table 7(b) above), as did their ability to produce coherent pair-groupings. Non-musicians' performance also appears to show a slight improvement at the level of producing coherent segment pair-groupings.

Comparison between original and B versions: A three-way mixed ANOVA, two (musical training) by two (experiment) by eight (segment position), taking the last variable as a within factor, was carried out. A main effect of musical training was observed ($F(1,36)=19.17$, $p<.0001$), mean distances from correct location being greater for non-musicians. Results for the original and B Versions of the piece did not differ significantly ($F(1,36)=1.224$, $p>.27$). There was a significant interaction between segment position and segment number ($F(7,252)=4.409$, $p<.0001$); Tukey tests showed that segments 1, 2, and 7 were the most distantly located from original positions across all subjects' pieces ($p<.05$). This can be accounted for in terms of the responses of non-musicians; the significant interaction between segment and training ($F(7,252)=2.526$, $p<.016$) showed that non-musicians tended to "mis-locate" segments 1 and 2, and (to a lesser extent) 7 by the greatest margin while musicians maximally "mis-located" segment 3 (their most accurate placements being those for segments 4 and 8). The interaction between segment and version was not significant ($F < 1$).

Discussion

Overall, performance of musicians and non-musicians was markedly different in both experiments. Contrary to hypothesis, the use of the musical materials of the B Version in which metrical and rhythmic grouping cues were strengthened produced little clear evidence of an increase in non-musicians' tendency to produce odd-even segment pairs. Thus neither of the puzzle experiments yielded unequivocal information about whether non-musicians' possessed or could access declarative knowledge of tonal-harmonic function.

It was decided to compare the results of the original puzzle experiment with those of the mental-line experiment in order to explore whether the different experimental procedures had led to different types of performance that might shed light on this issue. In the mental-line experiment, subjects had judged the position of each segment within the piece in the context of the representations that they had formed from having heard the piece; in the puzzle experiment, subjects had had to rely on their declarative knowledge of musical structure together with any cues present in the segments themselves in creating an appropriate ordering. Given that the evidence in the mental line experiment points to subjects having relied on characteristics of the musical surface in abstracting representations, it can be argued that if results for that experiment indicate a lower likelihood of producing odd-even segment pair groupings than in the puzzle experiment, subjects in the puzzle experiment were likely to have been relying on some declarative knowledge of tonal function in creating their pieces.

<u>Comparison between mental line and puzzle experiments</u>: This comparison was only performed on non-musician results and in relation to version A of the puzzle procedure, as version B did not involve the anacruses of the piece. Two 2-way ANOVAs were carried out with <u>experimental procedure</u> (mental line versus the original puzzle) as a between factor and <u>segment position</u> as a within factor. The first compared results of the first trial of the mental line experiment with the non-musicians' results from the first puzzle experiment, while the second compared the second trial of the mental line to the first puzzle experiment. The main effect of experimental procedure was significant in both analyses: respectively, $F(1,16)=8.696$, $p<.01$, and $F(1,16)=8.833$, $p<.01$, as shown in Figures 6(a) and (b). The main effect of segment was not significant for the comparison of mental line first trial with puzzle ($F(7,112)=1.059$, $p>.30$), but was significant for the comparison involving the second trial ($F(7,112)=3.977$, $p<.006$). The interaction between the two factors (segment and experimental procedures) was significant for the comparison between the first trial and the puzzle ($F(7,112)=2.138$, $p<.05$) but not for the other comparison ($F<1$). The significant result in the comparison of the first trial with the puzzle can be attributed to the fact that the mean distance from correct location for segment 1 was particularly large for responses within the puzzle experiment compared to those attributed in the mental-line first trial (4.75 vs 1.188, see Tables 2 and 4 above).

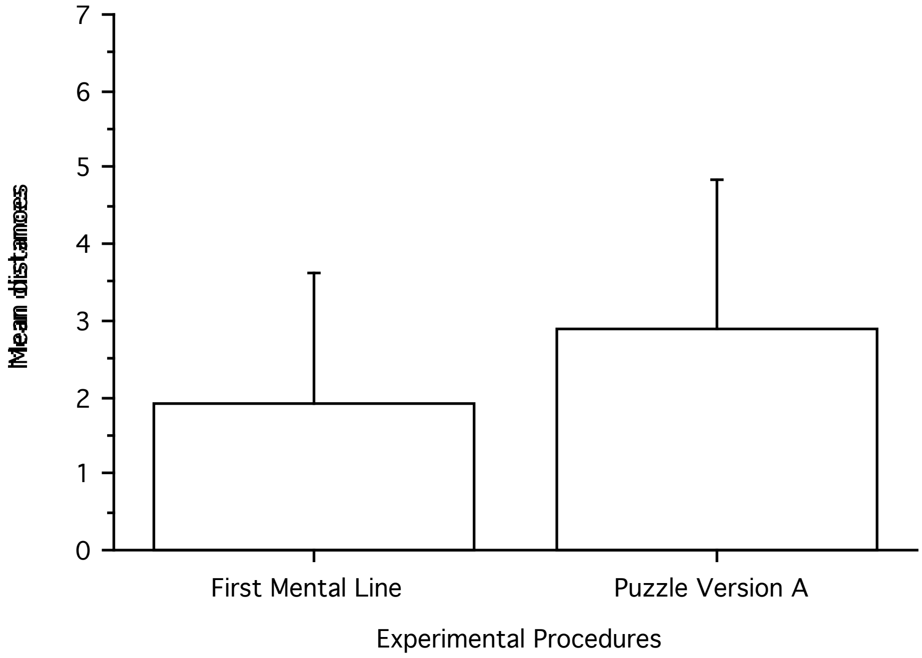

Figure 6(a)

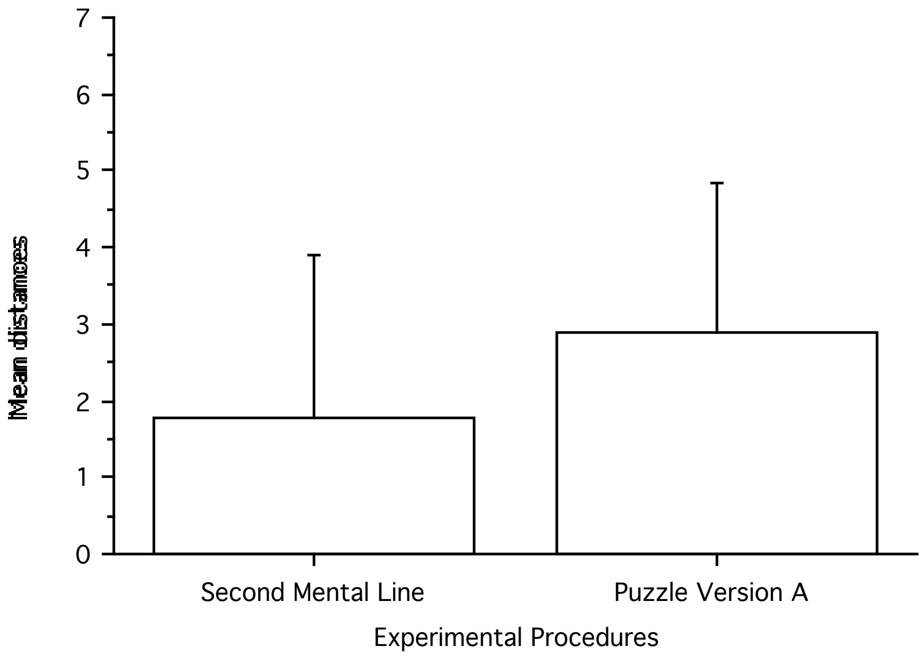

Figure 6(b)

Figures 6(a) and 6(b) were obtained by comparing the results of Experiment 2, Mental Line Procedure [First hearing, figure 6(a), Second hearing, Figure 6(b)] and those of the non-musician subjects in Experiment 3, Puzzle Procedure.

On the whole, the distances between the positions accorded to or chosen for segments and their correct locations were greater in the puzzle experiment than in either trial of the mental line experiment (see Tables 2 and 4). This indicates that, as expected, the function of segments within the piece was far less clearly defined for subjects in the puzzle experiment than in the mental line, where the piece was heard in its entirety. However, other than this difference there is little to shed light on the issue of the basis for subjects' performance in the puzzle experiments; it can be seen from a comparison of Tables 3(a) and (b) and Table

5(b) that there is no apparent difference in the frequency of use of odd segments at odd locations and even at even positions between either mental line trial and the puzzle results.

Consequently, it is not possible to decide on the basis of the results of the puzzle experiments whether durational/metrical features or tonal-harmonic functionality played more of a role in non-musician subjects' responses in ordering and grouping segments. However, given that few if any tonal-harmonic cues were reported as having played a role in non-musicians' subjects judgments in Experiment 1, it must be judged more likely that durational and metrical factors played the primary role in determining the ordering and grouping of segments by the non-musicians in both puzzle experiments.

Altogether, the results tend to indicate that non-musicians' responses were governed by characteristics of the musical surface; the tonal-harmonic information embodied in the segments of the piece seems likely to have been inaccessible to these subjects. The musicians appear to have been sensitive to such information, being able to project tension-relaxation relationships over the (admittedly quite brief) time-span of the piece (see also Dibben, 1994). The pieces produced by the non-musicians were not wholly unacceptable, but even in the best cases were unbalanced and tend to be far from the classical frame. They tended to use-up the "cadential" segments early in the piece, and were also generally confused by segments 5 and 6 (i.e. segment 6 appears to have been seldom construed as a repetition of segment 5, and subjects did not know where to place those segments). The musicians' pieces generally exhibited some structural coherence, but even these subjects also appeared confused by the "anomalous" segments 5 and 6.

It is clear that the musicians were sensitive to tonal structure, but only in respect of larger-scale formal articulation; their knowledge about the syntactical and structural function of segments (which may well be based on the type of declarative knowledge structure outlined in Krumhansl, 1990, or in Bharucha and Todd, 1991) obviously aided the construction of their pieces. Nevertheless, this sensitivity was not wholly reliable, and in any case, irrespective of such sensitivity, it appears evident that factors other than harmonic structure underlay the judgments of both musicians and non-musicians; for example, the fact that the alteration of the rhythmic structure that produced the B version material seems to lead to an enhanced capacity for the musicians to locate segment four in its correct position is perhaps an indication that even for the musician subjects, durational and metrical factors were powerful cues in selecting and ordering segments.

Conclusions

As already stated, most previous work on the perception of musical form (such as that reported in Serafine, Glassman & Overbeeke 1989, or in Dibben, 1994) has been conducted employing musician subjects (see also Clarke & Krumhansl, 1990; Krumhansl, 1995). The difficulty of conducting experiments on the musical capacities of non-musicians is pointed up by the fact that those few studies that have investigated non-musicians' perceptions of musical works either have resulted in broad accounts of very general propensities (e.g., Sloboda and Parker, 1985) or have employed fragmentary musical materials and have extrapolated from the patterns of sensitivities to these incomplete materials in proposing theories applicable at the level of complete pieces (e.g., Bigand, 1993a). The experiments reported here sought to employ complete musical pieces and to ensure that subjects' responses and activities during the experiments were minimally constrained, while their response tasks were not dependent on explicit knowledge of the task domain. Accordingly, the results reported here can be thought of as providing information about processes that might be implicated in everyday music listening.

On the whole, all the results for non-musician subjects (and many of the data from the musicians) are indicative of a reliance on elements of the musical surface in listening to and in manipulating the materials of complete tonal pieces (for similar conclusions, see Clarke and Krumhansl, 1990 and Krumhansl, 1991b, 1995). This finding seems to reinforce the idea that the primacy afforded to harmonic structure in Lerdahl and Jackendoff's theory may only be operational for musicians, and that those processes that appear common to musical listening for musicians and non-musicians (and that therefore are likely to be operational in and to be reinforced by processes of acculturation (see Sloboda, 1985) are based on factors other than tonal-harmonic syntax. While such factors can be accounted for in terms of Lerdahl and Jackendoff's 1983 theory (one of its strengths being its treatment of almost all musical parameters), they can be accounted for perhaps more parsimoniously in terms of the concepts of cue-abstraction and of schemata of order suggested respectively by Deliège and by Imberty.

In summary, Experiment 1 (the cue-localisation experiment) appears to confirm that non-musicians relied largely on elements at the level of the musical surface in forming some cognitive representation, exercising a much more restricted access to aspects of tonal function. Subjects were mostly consistent between the two trials of the experiment in their performance, the cues that they selected being primarily determined by features of the musical surface such as registral change, density, temporal discontinuities and repetition. Their responses in this experiment do not appear to be mediated by the differentiation between "objective" and "subjective" cues (respectively, identifiable in the musical score and largely the same for each listener, and non-identifiable in the score and variable from listener to listener) suggested by Imberty (1991; 1993). Subjects could clearly delineate the events that they selected as cues, and were also able to provide specific descriptions of the sounding effect that they perceived that could be stored in memory and help them to establish a schema of the piece. The locations of the reported cues within the score were always unambiguous, and the descriptions provided by those subjects who selected the same cues in both hearings of the experiment did not differ in content; all abstracted cues in this experiment are classifiable as "objective" in Imberty's terms. This does not invalidate his distinction, as the fact that no effect that could be traced to this differentiation was observed in the present experiment may be a result of the brevity and simplicity of the piece employed; it is possible that a more complex piece would have offered more opportunities for subjects to abstract "subjective" cues.

Experiment 2 (the mental line experiment) showed that non-musician subjects had difficulty in localising events in a just-heard piece when required to provide an explicit account of the time-course of those events in the piece. Evidence for the effect of repeated listening on subjects' performance is provided in the cluster analyses; these illustrate a clear tendency for the piece to come to be represented in cognition as a complete entity with repeated hearing, as well as a tendency for segments to be localised more accurately after repeated listenings. However, no clear evidence of any sensitivity to tonal function was observed in the results of either trial of the experiment.

Experiments 3a and 3b (the puzzle experiments) confirmed this observation and showed that non-musicians exhibited little if any sensitivity to the tonal functions of segments in respect of a complete piece. Trained musicians, as expected, were more capable of employing declarative information about tonal relations in manipulating blocks of tonal musical material, this being evident in the greater consistency with which they positioned those segments that outlined the tonal plan of the piece. However, the simplification of Schubert's Valse into its Version B form did not lead to enhanced performance by either musician or non-musician subjects, nor did it provide any evidence of reliance on the harmonic identity and function of segments by non-musicians. Comparisons carried out between the results of Experiment 2 and Experiment 3a (non-musicians only) also failed to reveal any evidence of such a reliance for non-musicians.

The puzzle results (Experiments 3a and 3b) appear to indicate that sensitivity to tonal-harmonic structure and function derives largely from formal musical training, a finding that appears in conflict with those of Bigand (1990, 1993a and b) although it seems consonant with views outlined by Krumhansl (see Krumhansl, 1990, p. 9). The finding that superficial aspects of the music were significant for both musicians and non-musicians reinforces the conclusions of a previous study (Mélen & Deliège, 1995); in a melodic recognition task it was found that when melodies were manipulated so as to hinder access to the underlying harmonic structure recognition was weak, but when better access to surface cues was provided, performances were enhanced in terms of both correct recognition and speed of response in both musicians and non-musicians.

In terms of the ideas expressed by Imberty (1985) it can be argued that the materials of one and the same piece appear to give rise to different "schematas of order" that are largely dependent on listeners' previous musical training. This difference is characterised by the strength of the "vectorial tensions" (Imberty, 1985) elaborated in listening to a piece. Though even for musicians, the features of the musical surface are evidently operational, they are able by virtue of their formal training, to develop long "vectorial spans" loaded by an important information weight. Non-musicians' vector formations are considerably smaller, inducing a more important parcelling of the complete content of the piece. Thus the structural stability of the salient auditory events of a piece that Imberty (1991) holds to be a central attribute of tonal music may well be largely inoperational for non-musicians. These remarks are consistent with results observed in categorisation processes (Deliège, in press) showing that cues memorized by musicians contain generally longer musical structures enabling them more efficiency in establishing relations between musical structures during listening.

The puzzle experimental procedure was devised as a complement to the mental line procedure (Deliège et al, 1994), in which the cues that were abstracted by subjects during listening could function as temporal "landmarks". This role could not be played by any cues that might be abstracted by subjects on their first encounter with the musical materials within the puzzle procedure. If cues are indeed a primary functional component in musical listening, it was expected that the absence of specific prior referents would lead non-musicians (at least) to have difficulty in manipulating the tonal musical materials which were presented to them in the experiment in random orders. The results of the puzzle experiments showed that non-musicians' declarative knowledge of and familiarity with tonal musical organisation is at best fragmentary, and does not furnish enough resources to rebuild even a very simple piece. Rather more surprisingly, even trained subjects - whose results were somewhat closer to the original piece - were far from consistent in their application of a "correct" tonal syntax. Thus reference to cues previously stored in memory seems to be useful for both categories of subjects, declarative knowledge of tonal musical syntax and structure seeming to provide an inadequate basis for the manipulation of tonal materials.

The research reported here can be taken to provide a demonstration of the effectiveness of the cue abstraction process and of the validity of the hypothesized mechanism of cue abstraction. Together with previous studies (Deliège, 1989, 1991, 1992c, 1993, in press; Deliège & El Ahmadi, 1990) the present results suggest that if a general theory of musical cognition is intended to account for the emergence of a progressively elaborated mental representation in real-time listening and to be applicable to a wide range of musics and degrees of musical experience, its focus must be on the structures of the musical surface.[vii]

[vii] This research was supported by a grant from the British Council in Belgium, Commissariat Général aux Relations Internationales, and Fonds National de la Recherche Scientifique to Irène Deliège and Ian Cross. The authors would like to thank Carol Krumhansl and three anonymous reviewers for their helpful comments on an earlier version of this paper.

REFERENCES

Bharucha, J., & Todd, P. (1991). Modelling the perception of tonal structure with neural nets. In P. Todd & G. Loy (Eds.) Music and connectionism. Cambridge, MA: The MIT Press.

Bigand, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. Psychology of Music, **18**, 1, 45-59.

Bigand, E. (1993a). The influence of implicit harmony, rhythm and musical training on the abstraction of "tension-relaxation schemas" in tonal musical phrases. Contemporary Music Review, 9, 123-138.

Bigand, E. (1993b). Contributions of music to research on human auditory cognition. In S. McAdams and E. Bigand (Eds). Thinking in sound: the cognitive psychology of human audition (pp. 231-277). Oxford: Oxford University Press.

Clarke, E.F. (1982). Timing in the performance of Erik Satie's "Vexations". Acta Psychologica, **50**, 1-19.

Clarke, E.F. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross and R. West (Eds). Musical structure and cognition (pp. 209-236). London: Academic Press.

Clarke, E.F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), Generative processes in music (pp 1-27). Oxford: The Clarendon Press.

Clarke, E.F. (1993). Generativity, mimesis and the human body in music performance. Contemporary Music Review, 9, 207-219.

Clarke, E.F., & Krumhansl, C. (1990). Perceiving musical time. Music Perception, **7** (3) 213-252.

Cook, N. (1987). The perception of large-scale tonal closure. Music Perception, **5** (2) 173-196.

Deliège, C. (1989). De l'oeuvre musicale comme discours et comme texte. In A. Helbo (Ed.). Text & Theory, Belgian approaches in semiotics. Semiotische berichte, Jg. **13**, 2, 3, 137-149.

Deliège, I. (1987a). Le parallélisme, support d'une analyse auditive de la musique : Vers un modèle des parcours cognitifs de l'information musicale. Analyse musicale, **6**, 73-79.

Deliège, I. (1987b). Grouping conditions in listening to music : An approach to Lerdahl & Jackendoff's grouping preference rules. Music Perception, **4** (4), 325-360.

Deliège, I. (1989). A perceptual approach to contemporary musical forms. Contemporary music review, **4,** 213-230.

Deliège, I. (1991). L'organisation psychologique de l'écoute de la musique. Des marques de sédimentation -indice et empreinte- dans la représentation mentale de l'oeuvre. PhD Thesis, Université de Liège, unpublished.

Deliège, I. (1992a). Perception et analyse de l'oeuvre musicale : points de rencontre. Analyse Musicale, **26**, 7-13

Deliège, I. (1992b). Paramètres psychologiques et processus de segmentation dans l'écoute de la musique. In R. Dalmonte and M. Baroni (Eds). Secondo convegno europeo di analisi musicale (pp.83-90). Italy: Universita' degli studi di Trento.

Deliège, I. (1992c). Recognition of the Wagnerian leitmotiv. Jahrbuch für Musikpsychologie, **9**, 25-54.

Deliège, I. (1993). Mechanisms of cue extraction in memory for musical time. Contemporary Music Review, **9**, 191-207.

Deliège, I. (1995). The two steps of the categorization process in music listening. An approach of the cue extraction mechanism as modular system. In R. Steinberg (Ed.), Music and the mind machine. Psychophysiology and psychopathology of the sense of music (pp. 63-73). Heidelberg: Springer Verlag

Deliège, I. (in press). Cue abstraction as a component of categorisation processes in music listening. Psychology of Music.

Deliège, I., & El Ahmadi, A. (1990). Mechanisms of cue extraction in musical groupings : A study of perception on Sequenza VI for viola solo by L. Berio. Psychology of Music, **18** (1), 18-44.

Deliège, I., & Mélen, M. (in press). Cue abstraction in the representation of musical form. In I. Deliège & J.A. Sloboda (Eds.), Perception and Cognition of Music, London: Lawrence Erlbaum Ass.

Deliège, I., Mélen, M., Stammers, D. & Cross, I. (1994) Musical schemata in real-time listening. In I. Deliège (Ed.), Proceedings of the 3rd ICMPC (pp. 271-272). Bruxelles: ESCOM Publications.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. Music Perception, **12** (1), 1-26.

Dowling, W.J. (1989). Attention and knowledge in music cognition. Proceedings of the First International Conference on Music Perception and Cognition (pp. 325-330). Kyoto, Japan.

Fredrickson, W. E. (1995). A comparison of perceived musical tension and aesthetic response. Psychology of Music, **23** (1), 81-87.

Gabrielsson, A. (1987). Action and perception in rhythm and music. Royal Swedish Academy of Music, No. 55.

Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. Journal of Verbal Learning and Verbal Behaviour, **13**, 114-131.

Handel, S. (1989). Listening. Cambridge, Mass. M.I.T Press.

Imberty, M. (1981). Les écritures du temps. Sémantique psychologique de la musique. Tome 2 Paris, Dunod.

Imberty, M. (1985). La Cathédrale Engloutie de Claude Debussy : de la Perception au Sens. Revue de Musique des Universités Canadiennes, **6**, 90-160.

Imberty, M. (1991). Comment l'interprète et l'auditeur organisent-ils la progression temporelle d'une oeuvre musicale ? Psychologica Belgica, **31** (2) 173-196.

Imberty, M. (1993). How do we perceive atonal music? Suggestions for a theoretical approach. Contemporary music review, **9**, 325-338.

Jackendoff, R. (1991). Musical parsing and musical affect. Music Perception, **9** (2), 199-230.

Johnson-Laird, P.N. (1983). Mental Models. Cambridge: Cambridge University Press.

Jusczyck, P.W. & Krumhansl, C.L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. Journal of Experimental Psychology: Human Perception and Performance, **19**, 627-640.

Kintsch W., & van Dijk, T.A. (1975). Comment on se rappelle et on résume des histoires. Langages, **40**, 98-116.

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

Krumhansl, C.L. (1991a). Music Psychology: Tonal structures in perception and memory. Annual Review of Psychology, **42**, 277-304.

Krumhansl, C.L. (1991b). Memory for musical surface. Memory and Cognition, **19** (4), 401-411.

Krumhansl, C.L. (1995). Perceptual approaches to the analysis of musical structure. Paper presented at the ESCOM Symposium organised at the VIIIth Conference of the European Society for Cognitive Psychology, Rome, sept. 1995.

Krumhansl, C.L. & Jusczyk, P.W. (1990). Infants' perception of phrase structure in music. Psychological Science, **1**, 70-73.

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

Lerdahl, F. (1989). Atonal prolongation structure. Contemporary music review, **4**, 65-88.

McAdams, S., & Bigand, E. (Eds). (1993). Thinking in sound: the cognitive psychology of human audition. Oxford: Oxford University Press.

McAdams, S. (1993). Recognition of sound sources and events. In McAdams, S., & Bigand, E. (Eds). Thinking in sound: the cognitive psychology of human audition. (pp. 146-195). Oxford: Oxford University Press.

Mélen, M., & Deliège, I. (1995). Extraction of cues or underlying harmonic structure: which guide recognition of familiar melodies? European Journal of Cognitiive Psychology, **7**, 81-106.

Nielsen, F.V. (1983) Oplevelse af musicalsk spoending (The experience of musical tension). Copenhagen: Akademisk Forlag.

Oura, Y., & Hatano, G. (1988). Memory for melodies among subjects differing in age and experience in music. Psychology of Music, **16**, 91-109.

Oura, Y., & Hatano, G. (1991). Identifying melodies from reduced pitch patterns. Psychologica Belgica, **31**, 217-237.

Palmer, C and Vandesande, C. (1993). Units Of Knowledge In Music Performance. Journal of Experimental Psychology: Learning, Memory and Cognition, **19** (2), 457-470.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. Cognitive Psychology, **15**, 66-94.

Postman, L., & Phillips, L.W. (1965). Short-term temporal changes in free recall. Quarterly Journal of Experimental Psychology, **17**, 132-138.

Prout, E. (1893). Musical Form. London: Augener.

Ratner, P. (1980). Classic music: Expression, form and style. New York: Schirmer.

Serafine, M.L., Glasmann, N., & Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. Music Perception, **6** (4) 397-430.

Sloboda, J.A. (1985). The Musical Mind. Oxford: Oxford University Press.

Sloboda, J.A., & Parker, D.H. (1985). Immediate recall of melodies. In P. Howell, I. Cross and R. West (eds). Musical structure and cognition (pp. 143-167). London: Academic Press.

Stoffer, T.H. (1985). Representation of phrase structure in the perception of music. Music Perception, **3** (2)., 191-220.

Sundberg, J., & Lindblom, B. (1976). Generative theories in language and music descriptions. Cognition, **4**, 99-122.

Sundberg, J., Friberg, A., & Fryden, J. (1991). Common secrets of musicians and listeners: An analysis-by-synthesis study of musical performance. In P. Howell, I. Cross and R. West (eds). Representing musical structure (pp. 143-167). London: Academic Press.

Tighe, T.J., & Dowling, W.J. (eds). (1993). Psychology and music: The understanding of melody and rhythm. Hillsdale, N.J.: Lawrence Erlbaum Ass.