Perceiving Similarities and Differences in Listening to a Piece of Music

A re-reading of a real-time listening model in relation to ANTICIPATION

Irène DELIEGE

URPM - Centre de Recherches musicales de Wallonie
Université de Liège
16 place du Vingt Août, 4000 - Liège (Belgium)
Fax 32 2 660 10 13 ; e-mail Irene.Deliege@ping.be

Abstract

Several experimental studies have demonstrated that the formation of groups, as defined by the Gestalt laws, applies to music listening. However, other organizational principles must be considered when music is played in real time. On this particular point, I have proposed a model based on a *mechanism of cue abstraction*, where two fundamental principles - the principle of *similarity* and the principle of *difference* - come into play in listening over long time spans. A number of experimental procedures based on the hypothesis of cue abstraction in relation to the concept of Anticipation are described, but particular emphasis is given on the role of similarity-difference perception. Starting from cue abstraction, a categorization process of the musical material is immediately established on the basis of exact or varied repetitions of the abstracted structures. Progressively, all these more or less varied reiterations elaborate a memory imprint : the memory does not retain all the precise components of the material, but establishes some sort of summary which comprises the principal attributes of a set of percepts. The results of different experimental approaches that followed these methodological principles are summarised.

Keywords: Cue abstraction, segmentation, grouping, categorisation, imprint

An overview of a model of cognitive processes involved in listening to a piece of music is given here. First of all, listening is considered as a schematization process based on a cue abstraction where two fundamental principles - the principle of *similarity* and the principle of *difference* - come into play in listening over long time spans. The basic lines of this model were proposed about ten years ago together with several experimental procedures to test the validity of its different aspects. Given the opportunity of this

International Journal of Computing Anticipatory Systems, Volume 4, 1999
Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-9600179-5-1

conference centered primarily on *anticipation*, some new visions of the model are suggested by developing an additional interpretation of some of its aspects.

Referring to Daniel Dubois (1998), "To anticipate means to realise beforehand, to foresee, to look forward, to act in advance, to prevent, to forestall" (p.3, vol.2). This is, indeed, nothing new as it pertains to usual definitions of the word. But, two different aspects of the concept - externalised anticipation and internalised anticipation (ibid. p. 4 & 5) - are further specified. *Externalised anticipation* refers to external events which can be anticipated and the author gives as an example "to take an ombrella when one is anticipating raining weather". This kind of anticipation is dependent on the environment: the weather may be sunny, so the future is potentially *multiple* and the realisation at each moment collapses all the possibilities to only *one*. Thus the present and the past are actually *unique*, meanwhile the future is always potentially multiple, but *a posteriori* there is only one realised present state. *Internalised anticipation* on the other hand, deals with the *memory of the future*. For example, to plan a lecture for a conference one year in advance, which requires to write a paper, to prepare the trip, etc. During that year one should thus memorize the event and all practical commitments to fulfill. Such kind of anticipation creates its own future events and manages to meet what has been anticipated: the realisation being practically a certitude.

Given the definitions of these two aspects of the concept of anticipation, which one should be most appropriate in relation to music listening ? At a first glance, to listen to music refers to an external event: a piece played in a concert hall, for example. On the other hand, to plan the concert in your agenda is clearly an internalised anticipation, but as soon as you are sitting on your seat in the concert room, the real-time listening process is dependent on an external event: the performance of the piece.

The second point of this introduction focuses on the cue abstraction underlying the schematization process in listening to music, based on the perception of similarities and differences. A study on rhythmic groups (Deliège, 1987), intending to test Lerdahl and Jackendoff's grouping preference rules (1983), made clear that even though the perception of music has its specificities, it also shares important psychological constants used in other perceptual processes. Indeed, at the basis of their grouping rules, there are also perceptual laws put forth in vision by the gestalt theory: proximity and similarity. Starting from this remark, possible relationships with other fields of psychological investigation were examined, principally with psycholinguistics, in order to exploit ideas outside the actual perception of music. Certain similarities between text and music processing clearly exist in the fact that both extend over long temporal spans. To grasp the contents of a discourse, for example, one does not need to adhere to the literal text but rather to certain salient elements: we are building reductions, or simplifications of the whole so as not to overload memory. These selections and reductions are the starting points for a schematization process.

Listening to a musical piece results, just as in listening to discourse, in the elaboration of a sort of schema. But given that music, unlike a text, does not refer back to a direct semantic support, it was necessary to study what kinds of tools such a schema might need to become effective. This reflection led me to the hypothesis of *cues abstracted during listening* (Deliège, 1989, 1991). Cues are conspicuous points which become fixed in the memory because of their pertinence and repetition. A cue, as defined by Charles Peirce (1974-78), always contains rare but striking attributes linking it to what it calls attention to, so that the latter may be recognized. It thus acts primarily as a reference point, a simple and effective way of processing larger quantities of data" (Deliège, 1989). For example, to recognise a motif in whatever piece only small cues suffice : in terms of anticipation, when abstracted and stored, the cue enables the listener to anticipate a longer sequence on the basis of only a few recognized features. In other

words, the function of cues is to generate abbreviations of longer units, thereby reducing the weight of information to be stored and allowing to anticipate broader sequences. Cues contain the invariants of discourse, which point them out as being the starting points for categorisation processes of new entries. Two main principles are here hypothesized : the principle of **same**, which permits a grouping of groups to be extended for as long as a same type of invariant is recognized, and **different**, which records contrasting features to establish necessary boundaries for segmentations. After a boundary, new cues will be abstracted to reorganise anew the abbreviation process for further anticipation of the information.

A *Segmentation procedure* enabled to observe how these two principles are at work during listening. Among others, Berio's *Sequenza VI*, a piece from the contemporary repertoire, was chosen as stimulus. It was predicted that subjects would not, or only seldom, segment as long as a same *cue* is perceived. The repetition of more or less similar material is for the listener an intuitive indication that the ongoing period of the piece is not closed. This prediction was supported by the results: much more subjects segmented at the boundaries of sections where strong contrasting structures were perceived. Internal segmentations were recorded but did not contradict the cue-abstraction hypothesis as they were less frequent and induced by rather local contrasting structures: these small units were regrouped in a larger one at a higher hierarchical level.

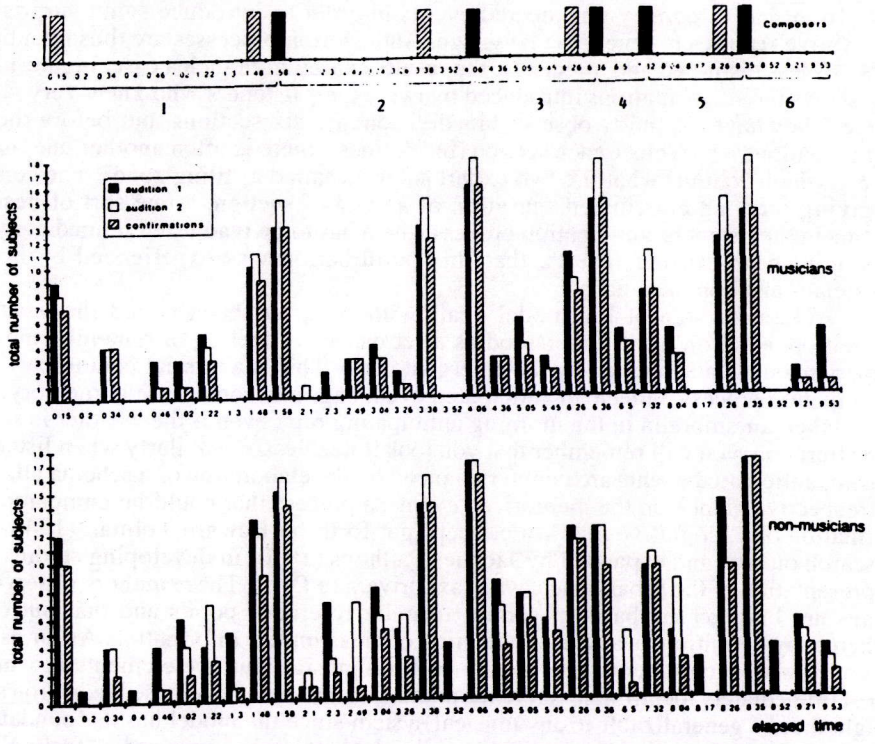

Fig. 1. Segmentations made by composers, musicians and non musicians in L. Berio, *Sequenza VI* for viola solo.

It should be noticed that there is no difference between Musicians and Non Musicians results, which seems to indicate that the segmentation process could be of a rather automatic nature, not greatly influenced by training (cf. fig.1).

In relation to the Anticipation concept, how could these results be analysed? Two different listening sessions were required in order to compare the collected segmentations and appreciate their degree of stability. In both listenings, subjects had to push a button of the computer to indicate the segmentations they were perceiving. Clearly in such a situation, the listener is facing a potentially multiple future (cf above, the definition of an "externalised anticipation"), but each time the button is pushed, all the possibilities are reduced to only one segmentation and so on until the end of the piece of music. What is interesting in these results, from an anticipation point of view, lies especially in the comparison of the performances of the two listenings. Almost no difference was observed, which seems to indicate that the anticipation process based on the abstracted cues is rather stable and established on an intuitive knowledge of what should be done. Exactly as one has an intuitive knowledge of the meteorologic features that are suggesting that an ombrella might be useful later in the day, here the point is that musical regions, inside which the musical materials acknowledge structural solidarity, should not be destroyed, nor fragmented, by major segmentations.

To investigate the segmentation processes, only real music compositions, pertaining thus to the artistic domain, were employed. Yet, an artistic message, by definition, has to convey unexpected events in order to introduce some surprising or innovative features to renew the language. Anticipation processes are thus intentionally more or less disturbed and, precisely, the results exhibit here such a case. Figure 1 on the top shows the segmentations introduced by two expert listeners who knew very well the piece. They have explicitly observed its division into six sections, but before the main segmentations which close each section (black lines), there is often another one (hatched lines), which delimits what the two expert subjects called a "tiling zone". The composer is giving there an anticipated sensation of an end of section. Some sort of hesitation occurs in the listeners' anticipation process which has to be readjusted immediately. The figure 1 shows, in the results, that this perturbation was experienced by both the musicians and non musicians.

The next step of the model deals with what has been called the *mental line* formation, a notion to be understood as a retrospective feeling or remembering of the experienced events labelled by the emergent cues. This is a normal counterpart of any anticipation process. Indeed an anticipated event engraves some trace in memory: if you have taken an ombrella in the morning anticipating rain, even if the weather is sunny in the afternoon, you will remember that you took it needlessly. Similarly when listening to music, anticipated events are counterbalanced by the elaboration of a schema allowing a retrospective glance on the memorized events, a process that could be compared to the formation of a "cognitive map", a concept put forth by Edward Tolman (1948) in his research on rats, and expanded by Jacques Pailhous (1970) in developing studies on the representation of the urban space, with taxi drivers in Paris. These authors observed that plans are built on the basis of some particular reference points and that out-coming schema, by definition, is a type of reduction of the complete information. As far as music is concerned, such schemas are largely built on musical surface elements in which the most immediately identifiable reference points lie ; the idea of cue-based reduction schema might thus be generalizable to any musical system since the model has its foundations in general cognitive mechanisms. Consequently, the listener has acquired a "feeling" of the location of the structures in the duration of the work and the time elapsed between events. Each cue is an entry label which summarizes a longer sequence. Accordingly, it was

hypothesized that the temporal course of a piece could be symbolized by a *mental line* along which the structures would progressively take place.

To investigate this proposal, subjects listened to a piece and were then presented with segments to be localized in the duration of the piece. Three different versions of the Mental Line procedure have been applied. In the first, on the *Sequenza VI* by Berio, where six main sections were recorded in the segmentation results (see figure 1), subjects were informed that they would be presented with about fourty segments to be located in their original section. All these subjects had participated three-four weeks before in the segmentation experiment. The results were significantly above chance and slightly better for musicians. In the second version, on Boulez's *Eclat*, the subjects were receiving 15 segments to be located on a horizontal line divided in 15 boxes which symbolized the complete duration of the piece. They participated just before in the same session, in a segmentation experiment of the piece and were only informed about the mental line task at the end of the segmentation experiment in order to observe how the cues act retrospectively in memory for time elapsed. Table 1 shows that the mode of the frequency distribution of each segment location coincided better with the actual position of the segments in musicians' responses. Errors were more frequent for segments located in the middle of the piece.

TABLE 1

The table shows the segments in the order in which they appear in the piece and the mode for each category of subjects. The figure in bold is the principal mode. The ordinary figures represent a secondary mode, i.e. the most frequent response after the principal mode. The numbers separated by a hyphen are the cases where 2 extracts were equally located in that position.

Segments	Mode Musicians	Mode Non-musicians
1	**1**	**1**
2	**2**/5	**2**/12
3	**6**/4	**11**/12-5
4	**5**/2-9	**3**/7
5	**5**/7	**12**/5-3
6	**6**	2-5
7	**9**	**9**/7
8	7-8-9-10	4-6
9	**4**/9	**3**
10	**12**/7-3	**10**/5-8
11	**11**/10	**9**/11
12	**12**	**9**/12-10
13	**13**	13-6
14	**14**	**15**/14
15	**15**	**15**/14

In the third version of the procedure, subjects had to reorder the complete material of a short piece, the *Tristan cor anglais solo* by Wagner, divided in 7 segments. Here again errors are few for the first and last segments as shown, in table 2. In addition, in contrast to the segmentation processes, which are of a rather automatic nature, it was here observed that training and familiarisation had a strong influence on subjects' sense of retrospectiveness. New factors were studied : number of previous familiarisation listenings (3 or 5) and 4 different categories of subjects were tested: Musician Students (MS) and Music Professors (PM); Undergraduate Students (NMS) and University Researchers (UR).

TABLE 2

Primary modes (in bold) and secondary modes of the positions attributed to each of the 7 segments by the different groups of subjects, as well as globally. After 3 hearings a majority of only the music students (MS) and professional musicians (PM) attained the correct order, but after 5 hearings the other subjects achieved correct results too.

	segm	mode primary/secondary		mode primary/secondary
NMS3	1	**1** / 2	NMS5	**1** / 2
	2	**3-5** / 4-6		**2** / 3
	3	**2** / 1		**3** / 2
	4	**4** / 3		**4** / 5
	5	**3-4-5** / 6		**5** / 3-4
	6	**6** / 5		**6** / 2-4-5
	7	**7** / -		**7** / -
MS3	1	**1** / 2-6	MS5	**1** / -
	2	**2** / 5		**2** / -
	3	**3** / 1-2		**3** / 4
	4	**4** / 3-6		**4** / 5
	5	**5** / 4		**5** / 6
	6	**6** / 3		**6** / 3-4
	7	**7** / -		**7** / -
UR3	1	**1** / -	UR5	**1** / 3-5
	2	**6** / 3		**2** / 3
	3	**2** / 3		**3** / 1-2
	4	**4** / 3		**4** / 1-2-5
	5	**4-5** / 3		**5** / 1-2-3-4
	6	**6** / 2-3		**6** / 4
	7	**7** / -		**7** / -
PM3	1	**1** / -	GLOB	**1** / 2
	2	**2** / 6		**2** / 5
	3	**3** / 2		**3** / 2
	4	**4** / 3-5		**4** / 3
	5	**5** / 4		**5** / 4
	6	**6** / 5		**6** / 4-5
	7	**7** / -		**7**

Table 2 shows the effect of familiarisation hearings on the mode level. After 3 listenings, only Musicians (students and professors) rebuilt correctly the piece. But after 5 listenings, the mode became correct for the 4 groups of subjects. Thus in terms of anticipation, both experimental procedures described until now, could be summarized as follows : (i) the sense of an externalised anticipation, outlined with the segmentation procedure in real-time listening, seems to elicit a rather automatic behaviour, not influenced by external factors: expert and naïve subjects performed in a similar way; (ii) on the contrary, considering the counterpart of the anticipation process, where the capacity of remembering is concerned, performances were strongly improved by training and familiarity with the perceived information. This is not to say that naïve subjects have no sense of retrospectiveness at all, but their memorized schema is not that accurate.

Other procedures investigated subjects' categorisation aptitudes underlying the anticipation processes in segmentations and in mental line achievements. From an anticipation point of view, two of these procedures seem to be most interesting : (i) the classification of structures on the basis of the abstracted cues; (ii) the formation of what has been called *imprints*.

A piece built on two main motifs, the *Allegro Assai* of the *Sonata for Violin Solo in C major* by Bach was employed in the experiments. Starting basically from the cue abstraction, the categorization process of the musical material is established on the basis of exact or varied repetitions of the abstracted structures. As put forth by Deirdre Wilson and Dan Sperber, "the representation and the represented object being two different things, they cannot, indeed, share all their properties : only some salient properties suffice. ... If I am summarizing for you an article that I have just read... you will in no way confuse the content of the summary with the one of the article itself" (1992, p. 227). In the same way, speaking of memory for temporal events, Fraisse notes "...not everything from our past experience is transferred into memory. A large part is not fixed. There is a large discrepancy between the immediate richness of a perception and what we can recall a few seconds later" (1967. p 167).

It was thus necessary to appreciate the subjects' ability to recognize very quickly a cue when reappearing in whatever varied form. All the different variations of the two main motifs of the piece were selected and presented three times in a random order in the experimental sessions. The results showed that the accuracy of the responses was above 90% for non-musicians and reached 100% for musicians. A similar experiment was later realized with 9-to-11 years old children. Easy pieces from Schubert and Diabelli were employed. The results reached between 75 and 90 percent. These very high performances allow us to assume that the classification of musical motifs based on abstracted cues involves highly reliable automatic processes. It is probably for this particular reason that no difference was observed in adult musicians and non musicians responses when studying their segmentation behaviour. It is even possible that these mechanisms are present very early in life. As pointed out by Andrew Woodfield "The child is already equiped, at birth, by several micro-features detectors" (1992, p. 277) and Roger Lécuyer observed that three or four months old infants have categorisation aptitudes (1994, p. 196).

Finally, another effect of cue abstraction and repetition lies in the *imprint* formation. When listening to whatever piece, listeners hear all the different variations of the motifs. This does not mean that all the precise components of all these variations will be stored. On the contrary, some sort of summary giving an "average value" - the *imprint* - will be automatically established which comprises the main coordinates of the ensemble. The background of this concept is to be found in the notion of *prototype* developed in categorisation studies, the prototype being the most typical exemplar that

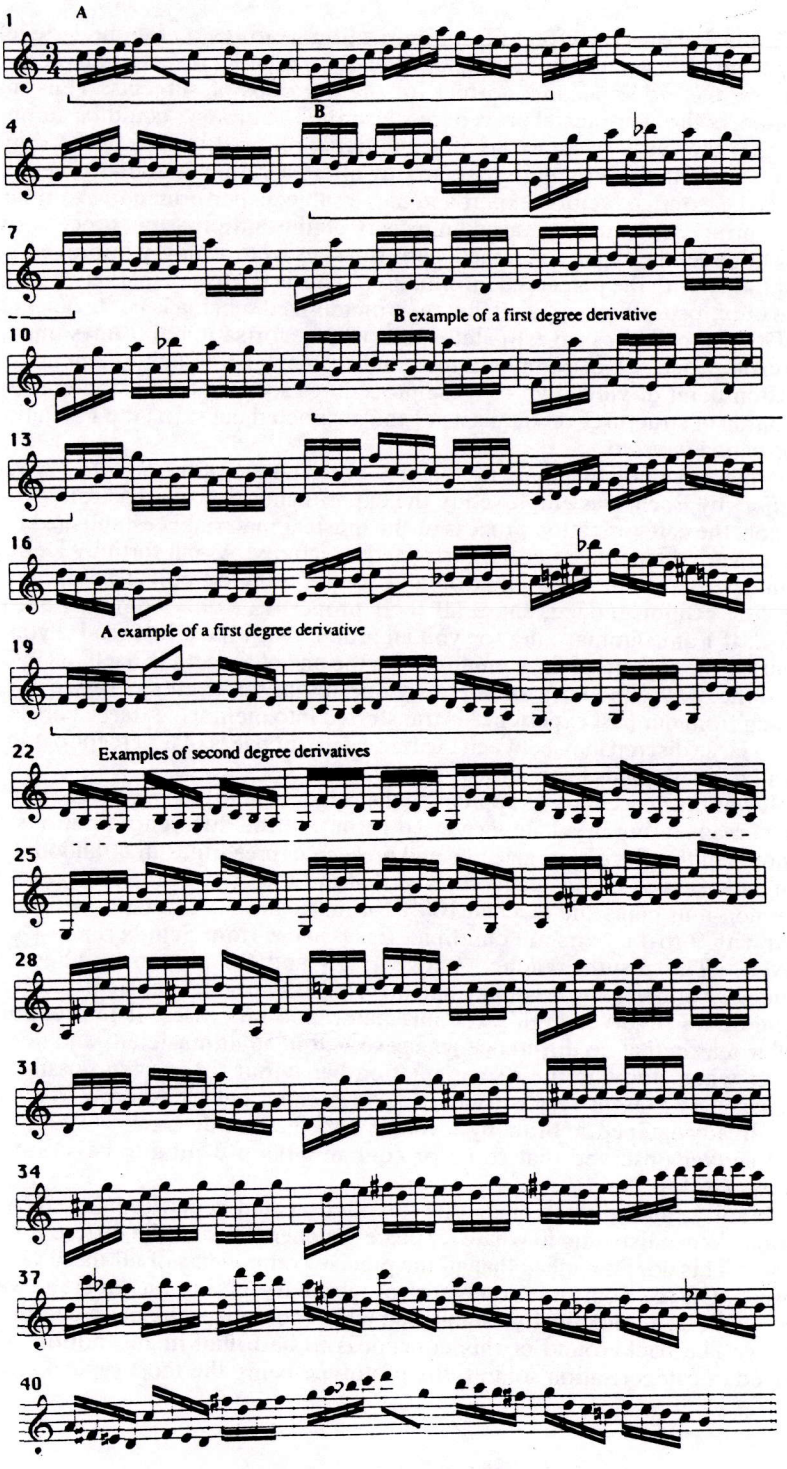

Fig. 2. J.S. Bach, First section of the *Allegro Assai* of the *Sonata* for violin solo in C major

comes to mind for a given category, for instance a blackbird or a sparrow for the category of birds, but not an ostrich, nor a turkey.

The procedure was inspired by prototype methodologies. Subjects first received an acquisition phase and listened to the initial section of the piece by Bach (the 42 first measures). Then a recognition phase followed. This time they were presented with (i) items already heard in that first section; (ii) items belonging to the non heard section; (iii) items with rhythmic modifications.

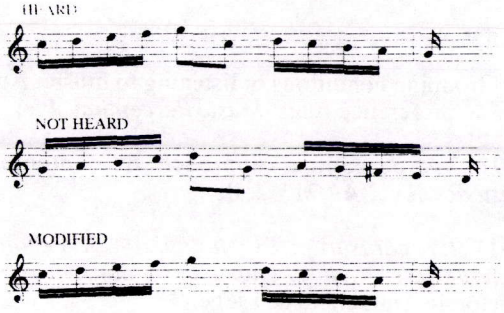

Fig. 3. Examples of heard, not heard and modified sequences

Thirty-two subjects (musicians and non-musicians) took part in the study. In order to detect an imprint effect, they had to decide whether the items had been heard or not before and to indicate their level of certainty in their responses. The idea behind this kind of procedure is that items very similar to those heard in the acquisition phase (thus here during the listening of the first part of the piece) would be erroneously felt as already been heard. Consequently, it was hypothesized that **modifications** or items too distant from the generated "average value" of the imprint, should be well detected.

As expected, for a great proportion, non-heard items were accepted as already been heard and the modified sequences were well rejected. This last result, particularly, is essential in relation to the stylistic characteristics of a piece : it shows that as soon as an imprint is stored, it acquires a value of an analyzer of the content of the work making a difference between what is and what is not the norm regarding the style of that work.

Conclusion

This re-reading considered some aspects of a model of real-time listening to music that was progressively elaborated over the last ten years. This model assumes that in order to experience the form of a piece of music, the listener elaborates a mental schema of the given work built on the basis of abstracted cues and broadly governed by the principles of **same** and **different**. It was observed that music listening is also supported by anticipation processes, especially those which govern *externalised anticipation* as defined by Dubois. As such, the *schematization* of the musical discourse underlined by the *abstracted cues* is similar to the schematization arising in text comprehension. Though, in

absence of semantic content in music, a particular stress was given on the role of structures - the *abstracted cues* - coming from the musical surface. Consequently, to demonstrate how a mental representation is progressively organized, research on listeners' music understanding, potentially relevant for any repertoire, should take into account the role of musical surface structures and their incidence on anticipation processes in real time listening. *

References

Deliège, I. (1987). 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, empreinte - dans la représentation mentale de l'oeuvre*. Thèse doctorale, Université de Liège.

Deliège, I. (1993). Mechanisms of cue extraction in memory for musical time. In I. Cross & I. Deliège (Eds), Proceedings of the 2d symposium *Music and the Cognitive Sciences*. Cambridge, septembre 1990. *Contemporary Music Review*, vol. 9, 1993, 191-207.

Deliège, I. (1995). Cue abstraction and schematization of the musical form. *Scientific Contributions to General Psychology*, **14**, 11-28.

Deliège, I. (1996). Cue abstraction as a component of categorisation processes in music listening. *Psychology of Music*, **24** (2), 131-156.

Deliège, I. (1998). Wagner "Alte Weise" : une approche perceptive. *Musicae Scientiae*, special issue 1998, 63-90.

Dubois, Daniel (1998). Introduction to computing anticipatory systems. *International Journal of Computing Anticipatory Systems*, vol. 2, pp. 3-14.

Fraisse, P. (1967) *La psychologie du temps*. Paris, Presses universitaires de France.

Lécuyer, R. (1994). Nouveauté et organisation des connaissances. In R. Lécuyer, A. Streri & M.G. Pêcheux (Eds) *Le développement cognitif du nourrisson*. Paris, Nathan Université, tome 1, troisième partie.

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

* The author is grateful to Christina Anagnostopoulou for her revision of the English .

Pailhous, J. (1970) *La représentation de l'espace urbain.* Paris, Presses universitaires de France.

Peirce, Ch. S. (1978), *Ecrits sur le signe*, Paris, Seuil; traduits de l'anglais et commentés par G. Deledalle (1974), *Collected Papers* , Cambridge, Mass., Harvard University Press.

Tolman, E.C. (1948). Cognitive maps in rats and men. *Psychological Review*, **55**, 189-208.

Wilson, Deirdre & Sperber, Dan (1992). Ressemblance et communication. In D. Andler (Ed.) *Introduction aux sciences cognitives.* Paris, Gallimard, coll. Folio-essais, pp. 219-238.

Woodfield, A. (1992). Un modèle à deux étapes de la formation des concepts. In D. Andler (Ed.) *Introduction aux sciences cognitives.* Paris, Gallimard, coll. Folio-essais, pp.273-290.