Musicae Scientiae
Discussion Forum 4A, 2007, 235-280

Melodic similarity as a determinant of melody structure

SVEN AHLBÄCK
Royal College of Music, Stockholm

• **ABSTRACT**

This paper presents an approach to the analysis of melodic similarity as a determinant of melody structure, which has been applied successfully as a computational system of analysis for the prediction of segmental structure in a large sample of monophonic melodies of diverse cultural origin. (Ahlbäck 2004) Although the impact of similarity in the determination of segmental structure in music is generally acknowledged, methods based on similarity have been criticized with regards to the difficulty of formalizing criteria and threshold values for structurally significant similarity (e.g. Cook 1987). It is herein maintained that, since similarity is a fundamentally relative concept and categorization of similarity and difference relates to a given context, segmentation based on similarity requires a context-sensitive parametric measure of similarity.

The proposed method of analysis is based on common psychological principles such as gestalt psychological principles, human perceptual and cognitive limitations regarding temporal frames of attention, simultaneous category handling and cognition of temporal proportions.

The core hypothesis of the model is that melodic segmentation at a structural level is primarily established by similarity, in particular repetition of melodic content, and dissimilarity, in particular discontinuity of melodic processes. It is discussed how the influence of melodic similarity on segment structure is dependent upon general features of musical structure, such as metrical structure. This notion is supported by the results of an experiment, which indicates that listeners do not make use of repetitions of pitch sequences for melodic segmentation when these are in conflict with the perceived general beat structure.

The influence of metrical context is handled within the method, as well as the means of allowing for different levels of similarity through a categorization of different types and degree of similarity. This is illustrated by an example analysis of a melody in which the segmental structure is determined by melodic similarity of different types, the result of which is evaluated by a listener test. The result supports the notion that structurally significant similarity is relative and contextual and indicates that this is possible to model formally within a rule-based, style-independent method of analysis.

1. Introduction

1.1. The contextual relativity of similarity

The central theme of this article is based upon a fundamental observation: Similarity is contextual and relative. For example, in the context of the species of the earth, all humans are similar and constitute a category, while in the context of the human race men and women are categorically different. However self-evident this observation may appear, it should not be underestimated; categorization by similarity is related to the perspective of the observer.

There are, however, examples of similarity between musical works being treated in an absolute sense within fields as different as those of copyright law and musicology. Generally, similarity regarding the phenomenal structure of musical works is believed to reflect common musical origin. This can be problematic, since one might ask in what sense and at what level of similarity is it safe to assume common origin of musical ideas.

Moreover, if we actively search for similarity we may be gradually extending our similarity category the larger the sample is that we examine, if we do not relate to a category of non-similarity. One example of this can be found in a study by the important Swedish musicologist Carl-Allan Moberg. Moberg aimed to establish a connection between two similar melodies, a Swedish lullaby and a twelfth or thirteenth-century French *lai*. In his effort to connect these two melodies, the category of variants of the same melody is gradually extended. Figure 1 (ex. a-d) shows four melodies which he incorporated into the category of variants of the same melody.

The most striking phenomenal similarity between the examples given by Moberg can be found between the reference melody (a) and the French *lai* (d), but are the two other melodies (b and c) really members of the *variant* category? It might well be that these melodies have a common origin in the French medieval melody. However, from the evidence presented by Moberg it is impossible to judge whether the very general features they have in common, such as the initial ascending fifth interval, the fifth frame and the common elements of form, are features typical for an entire repertoire of melodies or if they are specific to this melody as opposed to other melodies within the same corpus. To answer these questions the context of the similarity categorization in terms of the sample of melodies must be defined.

If we extend Moberg's sample to include other European folk music traditions we will find a great number of melodies which are as phenomenally similar to the lullaby as the variants given by Moberg (see e.g. fig. 1, ex e) [1]. The category will expand to

(1) The Example e (fig. 1) is a version of a Finnish melody, the "Kalevala melody", which also has served as a reference melody for variant studies. The two categories overlap, as observed by Väisänen (1954), among others. As Väisänen remarks, an assumption of common origin based on similarity between the two melodies becomes absurd since it would suggest close connections between socio-culturally, historically and geographically very distant cultures.

Figure 1.

The examples a - d were categorized as variants in C.A. Moberg (1950: 25-45). No 1, a common lullaby melody (Vyssa lulla…), no 49 and no 63, common Swedish song-dance folk melodies. no 99, Lai from J-B Beck: Gesammmelte Studien. These examples are four out of a total of 101 variants presented by Moberg. Example e is a Finnish "runo song" melody, from A.O. Väisänen (1954:39-55).

the extent where it becomes meaningless to speak of variants of "the same" melody, but rather of an "elementary musical concept" (Ling 1989: 56).

Moberg's categorization of variants demonstrates that if we are looking for similarity, i.e. forming a category by sameness, there is almost no limit as to how general similarity can be considered significant. It depends on the diversity within the sample we are looking at; the more diverse the context, the more general similarity may be recognized.

This reasoning is also relevant regarding similarity as a structural determinant in melody cognition; in what sense and at what level can melodic similarity be perceived as determining melody structure, i.e. indicating segmentation of melody? In the case of analysis of melody structure, however, the melody per se constitutes the "sample", thus forming the context for similarity categorization.

1.2. The role of similarity in melody structure — General questions

Similarity as a determinant of musical structure is generally acknowledged but is also the subject of debate (see for example Cambouropoulos 1998:39, Lamont & Dibben 2001: 246). In the socalled method of *paradigmatic analysis* by Ruwet and Nattiez

(Ruwet 1966, 1977, Nattiez 1990) similarity has a special position as a structural determinant. According to this method of analysis, musical structure is defined mainly by repetition of similar musical structures. Two structures are considered *equivalent* when they are sufficiently similar, according to the model. It has been criticized, for example by Middleton (see also Cook, 1987) for the absence of a formalization of similarity in relation to equivalence.

> "Which parameters are to be regarded as pertinent and on what grounds? What are the criteria for a judgement that two entities are sufficiently similar to be considered equivalent? If intuitive, such decisions would seem to solve the segmentation question before the event. […]." (Middleton 1990: 183)

The problem of at what level melodic similarity is structurally determinant is also addressed by Cambouropoulos in *Towards a General Computational Theory of Musical Structure* (1998). He criticizes the application of the method of paradigmatic analysis by Nattiez from the viewpoint that all musical sequences are similar in some respect.

> "[…] the paradigmatic technique proposed by Ruwet whereby relationships between musical sequences are established mainly because of recurrence and repetition (with or without variants). But can such relationships be established in a true neutral manner […])? It is suggested herein that if similarity (i.e. not merely exact repetition) is taken into account in an analysis of surface structure, then analysis at neutral level becomes unwieldy because any two musical sequences are similar in some respect. Analysis at the neutral level is useful only if guided by some sort of heuristics — for instance, based on general cognitive principles." (Cambouropoulos 1998: 9).

In his criticism Cambouropoulos points to the relativity of similarity judgment, but he also calls attention to the problem of determining at what level of similarity two musical sequences may be considered structurally equivalent.

The criticism of the paradigmatic analysis thus points to some important questions regarding the role of similarity in melodic structuring:
- Which musical parameters have significance for melodic similarity?
- What are the criteria for judging that two entities as sufficiently similar to be considered equivalent? Is there a threshold value for structurally determinant melodic similarity?
- If perceived similarity is relative and contextual, is it possible to model it formally?

In the following, a new approach to the problem of melodic similarity as a structural determinant will be proposed, namely to employ a dynamic and contextual similarity measure with a given number of parameters within a formalized method of analysis. Different aspects of the influence of melodic context on similarity recognition will be considered, such as the influence of context in terms of the general level of homogeneity vs. diversity of melody structure, the influence of context in terms of the general structural features of melody and the influence of context in terms of the temporal order of category formation[2].

Similar rule-based approaches to structural analysis by means of similarity can be found in some other relatively recent models, in particular within the previously mentioned interesting work by Cambouropoulos (1998, 2001, 2005). Important concepts are presented also in some models that are not directly concerned with melodic segmentation, in particular within pattern-finding models proposed by Lartillot (2004) and Conklin and Anagnostopoulou (2001) and within a model for metrical analysis based on melodic parallelism by Temperley & Bartelette (2001).

The analytical model of melodic similarity which will be presented in this paper was developed within a broader study (*Melody Beyond Notes, A Study of Melody Cognition,* Ahlbäck, 2004), in which a general computational model of melodic surface structure is proposed. Briefly, the fundamental question posed in that study was whether it is possible to construct a relatively style-independent theory of the cognition of melody structure from cognitively based assumptions of the relationship between phenomenal and conceived melody structure. This resulted in the development of a formalized rule-based model that was able to make testable predictions of cognitive implications of phenomenal structure. This was achieved by a computer implementation of the model. In order to study the implications of phenomenal melody structure in its most limited sense — as pitch change over time — the only input to the model was pitch and duration. This was in the form of Standard Midi Files, comparable to a "piano roll" representation of melody. The method of analysis involves context-sensitive quantization, melodic pitch categorization, metrical analysis and phrase/section analysis of non- and quasi-metrical melody structures (see fig. 2). The model was evaluated by applying it to melodies of different stylistic origin, and also through comparisons with listener experiments. The model was tested on more than 200 melodies.

Measure of melodic similarity holds a central role within both the metrical analysis and, significantly, within the phrase and section analysis, both for the segmentation as such and for the analysis of relationships between segments and hierarchical levels of segments.

This article focuses on melodic similarity as a cue for melodic segmentation in metrical music. In the article, I will investigate some of the theoretical foundations for the measure of similarity. I will then investigate how similarity measurement is employed in the proposed method of analysis of melodic surface structure, specifically regarding melodic segmentation within the Metrical Phrase and Section Analysis (see fig. 2). This involves presenting both the general grouping principles on

(2) An interesting general discussion of the influence of context on similarity categorization in music can be found in Cambouropoulos (1998; 2001).

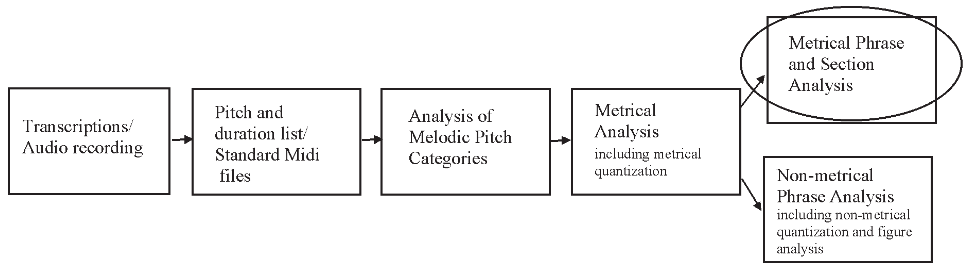

Figure 2.

Outline of the computer implementation of the proposed method of analysis of melodic surface structure (the MODUS model for analysis of melody structure, Ahlbäck 2004). The current paper concerns similarity measure within the enclosed module.

which the model is founded (section 2) and the parts of the analytical model which concern segmentation by similarity (section 3). The latter includes an outline of the entire model of phrase analysis (section 3.1), discussion of the theoretical basis for segmentation by similarity within the model (section 3.2), a presentation of similarity classification (section 3.3) and an evaluation of structurally significant similarity within the model (section 3.4). The presentation includes references to the computer implementation of the theoretical model, and the discussion presupposes that the initial steps of the analysis are performed, including the metrical analysis. It should be noted, however, that essentially, the same method of melodic segmentation based on melodic similarity is also employed within the module Metrical Analysis (see fig. 2, Ahlbäck, 2004: 99-168). Section 4 of this article presents an example of a metrical phrase and section analysis, together with a listener evaluation of the segmentation performed by the computer model. In the concluding part of this article (section 5), the results of the application of the model to test material will be briefly discussed.

The theoretical concepts presented below constitute parts of a general theory of cognition of melodic surface structure, which is presented in its entirety in the aforementioned study. The scope of the present article necessitates that some concepts are only mentioned briefly.

2. General principles for melodic segmentation

The segmentation or identification of structural units within a melody can be regarded as a process of grouping melodic events. Within the current model this process is assumed to be governed by general principles that can be expressed in terms of the Gestalt laws (see e.g. Wertheimer, 1923). There are other well-established models of musical structure that have a similar approach, such as "*A*

Generative Theory of Tonal Music" proposed by Lerdahl & Jackendoff (1983). Since the focus of this paper is melodic similarity I will not discuss the herein proposed structural principles in relation to these other models.

A fundamental hypothesis is that the principles which govern the experience of grouping as well as any aspect of structure, can be reduced to just one basic principle or one question — the same / not the same (c.f. Deliège, 1989: 214-215; 2001: 236) This opposition can be formulated as specific contradictories along different parameters, with important applications for melody structure, similarity — dissimilarity (the same/not the same shape or qualities), continuity — discontinuity (the same/not the same stream), proximity — distance (the same/not the same location), recurrence — change (the same/not the same events). Primary grouping is dependent on perception of either the positive or the negative side of this basic principle: events perceived as "the same" are both defined as units and are united by similarity while "not the same" indicates boundaries between units. The first case refers to grouping by similarity, continuity, proximity and recurrence and the other to grouping by dissimilarity, discontinuity, distance and change; in other words the positive and negative sides of the same/not the same.

On the basis of the above distinction, as well as upon general principles of cognitive economy, coherence and significance (see footnote[3]) and experiments (Ahlbäck, 2004: 188), I have proposed a set of gestalt theory-based principles for melodic grouping. The core of this model is the classification of the different grouping principles according to the role they play in the cognition of grouping structure in melody. The model postulates that groups are primarily formed by categorization into "the same" and "not the same", i.e. grouping by sameness and difference, designated *primary grouping principles*. It further postulates that groups can be formed by implication based on properties of grouping determined by primary grouping principles, reflecting the general coherence principle. These *secondary grouping principles* refer to grouping by good continuation, which involves both selection of grouping and implicative grouping by means of periodicity and symmetry. The third class, *tertiary grouping principles*, is defined as involving perceptual and cognitive selection of grouping (which may also be expressed as grouping preference), reflecting both the principle of cognitive economy and the principle of significance and intelligibility. To this category belongs grouping by perceptual salience, corresponding to the figure-ground gestalt principle, and

(3) Structuring is assumed here to be governed by three inter-related general principles; (1) The principle of *reduction of information* or *cognitive economy*, i.e. that we are consistently and spontaneously trying to reduce the amount of information that needs to be processed through categorization and filtering of information; (2) The principle of *coherence*, i.e. that we seek for all-pervasive, general principles that can be thought to rule structural processes; (3) The principle of *significance* and *intelligibility*, i.e. that we are consistently looking for significance and intelligibility in the structure. (See Ahlbäck, 2004: 37ff; cf. Cambouropoulos, 1998: 27,33-34)

grouping by structural integrity, corresponding to *prägnanz* and *closure* gestalt principles. The primary grouping principles are thus thought to create grouping, the secondary to implicate grouping and the tertiary grouping principles refer to grouping prominence and selection.

2.1. Primary grouping principles (I)

The primary grouping principles govern the basic formation of groups in the sense that they create primary grouping indications. The primary grouping principles are divided into grouping by sameness and difference.

Grouping principle I.A Grouping by sameness — similarity, continuity and proximity

Sameness both unites and defines. It unites in the sense that close and similar events tend to be grouped together. It defines in the sense that similar events and recurrent series of events tend to be perceived as groups.

The *uniting* aspect refers to grouping of events based on shared properties, such as grouping based on "same" pitch or "same" duration. One example of this principle is given in figure 3. The *defining* aspect, on the other hand, refers to grouping of series of events based on recurrence, grouping based on "same" pitch changes, "same" duration changes etc. The "same" pitches are *united* by pitch similarity, while the "same" series of pitches are *defined* as a group by recurrence.

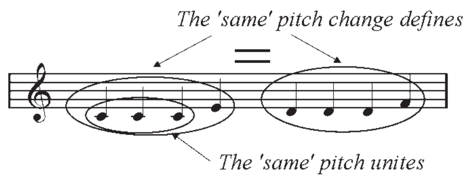

Figure 3.

Grouping by sameness (grouping principle I.A) is exemplified by grouping of events with the "same" pitch and grouping by recurrent "same" pitch change Grouping indications are shown as enclosed areas.

Grouping principle I.B Grouping by difference — dissimilarity, discontinuity, and distance

Difference separates, i.e. indicates a group boundary. This implies that events or groups of events which are relatively dissimilar or distant with regards to pitch or duration or are separated by a discontinuity or change, are indicated as belonging to different groups (see example, fig. 4).

This principle mirrors the *uniting* aspect of grouping by sameness. However, there is no mirroring of the *defining* aspect of grouping by sameness, grouping by

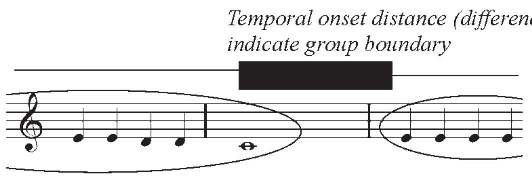

Figure 4.
Example of the grouping by difference principle (grouping principle I.B). Temporal onset distance indicates group boundary. Grouping indications are shown as enclosed areas.

recurrent order of events, since the opposite of order is chaos. As will be discussed in section 2.5, grouping by sameness and difference are different with regards to the power of structural implication.

2.2. SECONDARY GROUPING PRINCIPLES (II)

Secondary grouping principles govern the formation of groups by implication and the selection of groups by prominence. This implies that groups can be formed by implication based on properties of grouping determined by primary grouping principles, which reflects the general coherence principle[4]. But it also implies that grouping indications, which comply with the general properties of the structure, will be perceptually more prominent, i.e. more likely to be recognized, than groupings which conflict with the previous grouping structure.

Grouping principle II.A Grouping by good continuation/constancy
Group size, group start etc., which is coherent with previous grouping is prominent and can be implicative in the case of grouping by periodicity (see example, fig. 5).

This reflects the gestalt principle of *good continuation.* Implicated grouping means that if a grouping is indicated repeatedly, for example, by recurrence, the grouping may be conceived to continue even without phenomenal support by primary grouping indications (see fig. 5 and section 3.2.2, fig. 13).

Grouping principle II.B Grouping by symmetry
Symmetrical grouping is more prominent than asymmetrical grouping and can be implicative in the case of grouping by hierarchical symmetry (see example fig. 6).

This means that a grouping is more likely to be recognized if it complies with a symmetrical structure, which reflects the general principles of coherence and cognitive economy. It further implies that if a structure exhibits hierarchical symmetry, symmetrical grouping may be conceived even if there is no phenomenal support by primary grouping indication.

(4) This means that the secondary grouping principles have the power of implicating new groupings without phenomenal support from primary groupings, which is not true for the tertiary grouping principles.

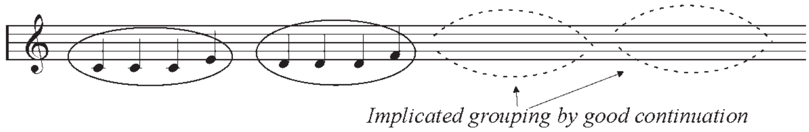

Figure 5.

Grouping by good continuation (II.A) is exemplified by grouping implications based on periodicity (displayed by dotted enclosures).

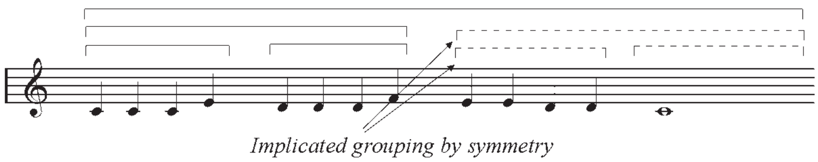

Figure 6.

Grouping by hierarchical symmetry (II.B) is indicated by dotted brackets, which mirror the primary groupings in the first half of the structure.

2.3. Tertiary grouping principles (III)

The tertiary grouping principles govern the selection of groups by prominence. The tertiary grouping principles, like the secondary grouping principles, lack the power of giving primary indications of grouping, but concern the structural prominence of grouping indications originating from primary grouping principles. Thus, they are involved in the perceptual selection of grouping or, in other words, influence which possible grouping indication will actually be recognized.

Grouping principle III.A Grouping by perceptual salience

Groups delineated by perceptually salient, e.g. articulated, events are prominent (reflecting the figure-ground Gestalt principle). This means that indication of grouping, in which the most articulated event marks the boundaries of the groups will be more likely to be recognized than an alternative grouping in which the most articulated event is within the group (example, see fig. 7). Furthermore, an alternative grouping indication in which the first event in the group is the most articulated (e.g. by phenomenal accent[5]), represents the most evident impulse (e.g. by frequency of onsets[6]) or has the most grave articulation (e.g. by duration[7]) is prominent.

(5) The term phenomenal accent is borrowed from Lerdahl & Jackendoff (1983: 17)

(6) Frequency of onsets refers to the relative density of events in terms of interonset period. The assumed principle is that the first tone onset in a series of tone onsets with relatively short interonset periods is perceived as articulated, receiving an impulse quality.

(7) The assumed principle is (in analogy with linguistic terminology) that the graveness of an articulation refers to the perceptual prominence of the event, in terms of e.g. duration, pitch or

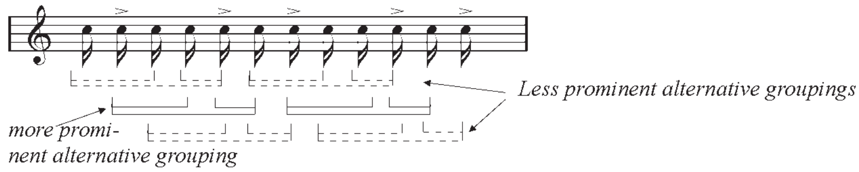

Figure 7.

Grouping by perceptual salience (III.A) is illustrated by alternative hierarchical groupings related to articulation, where the most prominent grouping — between most articulated, hence salient events — are marked by unbroken brackets. The grouping alternatives originate from primary grouping principles, sequence repetition of "the same" accent structure.

Grouping principle III.B   Grouping by integrity and contrast

Discrete grouping in the sense that the content of the group contrasts to the group boundaries, is prominent, The structural integrity of a grouping indication is related to the degree to which it can be considered discrete in the above sense, which reflects the gestalt principles of *Prägnanz* and *Closure*.

This means that a grouping indicated by primary grouping principles is more likely to be recognized the greater the contrast between within-group and between-group events, and the more evident the within-group contrast is (with regards to any dimension of melody structure) (fig. 8).

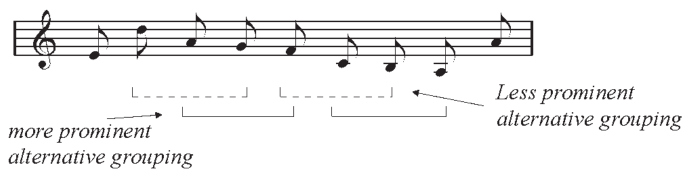

Figure 8.

Grouping by structural integrity (III.B) is exemplified by alternative sequences where the most prominent grouping displays the greatest contrast between within-group pitch intervals (steps) and between-group pitch intervals (leaps) of the two alternative groupings. The alternative groupings originate from grouping indications by primary grouping principles, i.e. sequence repetition.

volume. A tone that is relatively long, relatively low or relatively loud is assumed to be perceived as perceptually more prominent and as having greater power to mark the boundary of a group than less prominent events.

2.4. The relationship between grouping by sameness and grouping by difference

As is exemplified above, there are two different dimensions of the grouping by sameness, the defining and uniting aspect of sameness. Since the latter aspect, that similar elements are grouped together, requires the contrast to non-similarity to be structurally definite, the delineation of groups formed by similar elements requires segmentation by "difference". This is, however, not the case with the other aspect of grouping by "sameness", the definition or formation of a unit/group by recurrence or repetition, where the identification of a group is not possible by change alone. Thus, the most important aspect of grouping by "sameness" is grouping by the *recurrence* of something, i.e. a recurrence of a perceived change along a structural dimension. In melodic terms, that "something" may be an array of events involving pitch-change and/or change of duration of events, a melodic shape or process (see fig. 3). Such an array of events defined as an entity by similarity through immediate recurrence is designated here as a *sequence*[8]. Conversely, the most important aspect of grouping by "difference" is a *change* of something, which implies a group boundary[9]. The change may imply the change of melodic properties such as change of pitch set, melodic direction, pitch interval size, tone/rest and duration[10].

3. Melodic segmentation by means of melodic similarity

3.1. General design of the model of melodic segmentation

The hierarchy of principles of grouping structure is implemented in the current model in a three-tiered method of analysis of melodic segmentation:

(8) The concept of sequence repetition may, in this significance, be regarded as a special case of the general concept of pattern recurrence, implying an immediate, contiguous repetition of an array of eventswhich are conceived as being structurally equivalent. A pattern, on the other hand, may be identified by recurrence without any relationship between the instances of the pattern or in a very general sense even without any recurrence at all. Sequence repetition may result in segmentation, which in this case refers to the perception of a division of a melody into substructures. The special significance of sequence repetition with regards to segmentation in comparison to pattern recognition in general is that a sequence repetition determines a certain time-span, which is not necessarily the case in pattern recognition.

(9) The dissimilar implications of similarity and discontinuity are remarked upon by Deliège (2001: 236) who cites George Vignaux: "Similarities allow us to group objects together, differences allow us to set them apart".

(10) One may consider the grouping by difference principle more *local*, since it considers the change of "the same" conceived as a continuum — it does not require any knowledge of the extent of "the same". Likewise, the grouping by "sameness" principle can be considered a more *global* principle, since the classification of "the same", requires knowledge of all the members of a group. However, both principles apply to all hierarchical levels, global as well as local.

1. Segmentation by sequence repetition (Primary grouping principles I.A) evaluated by discontinuity/change (I.B) together with secondary and tertiary grouping principles.
2. Segmentation by structural discontinuity on local and global level (Primary grouping principles I.B), evaluated by secondary and tertiary grouping principles.
3. Segmentation by good continuation, i.e. periodicity, and grouping implication by symmetry (Secondary grouping principles, II.A and II.B).

The order of the above three steps of the analysis is determined by the nature of the different grouping principles. Most importantly, sequences can include structural discontinuities, which in isolation would be perceived as group boundaries. Consequently, segment implication by sequence repetition is superordinate to segment implication by discontinuity. Segment implication by good continuation and symmetry is derived from primary grouping, which is why it must be analyzed after primary grouping indications. Simultaneously, the secondary and tertiary grouping principles are involved in the evaluation of different grouping implications which originate from primary grouping principles (see section 2.4, III.A and III.B).

As can be understood from this, indication of segmentation by melodic similarity plays a central role in the method of analysis, through the segmentation by sequence repetition called *sequence analysis*. In the following, we will focus entirely on this part of the segmentation analysis. If sequence repetition occurs, it is assumed to be the most important means of structural indication on a global level, since it can organize many individual elements and can provide precise indications of group boundaries. On the other hand, the phase of the sequence must be determined by structural discontinuity/change.

The sequence analysis is based on the identification of contiguous melody structures that are conceived to be structurally *equivalent*, with regards to different properties of melodic surface structure (see below, section 3.2.1). This implies that all possible sequences at any level of structural generality are considered. The structural prominence of the sequences is then evaluated with regards to degree and type of similarity, structural integrity, structural coherence and intelligibility by application of the secondary and tertiary grouping principles, together with the grouping by discontinuity principle.

3.2. What is similar enough? A theoretical basis for measuring similarity

3.2.1. The need for a parametric similarity measure related to dimensions of melodic surface structure

The primary task for the analysis of melodic segmentation is to identify the sequences, recurrences of phenomenal structures that are cognized as being structurally equal. — But what is similar enough to be regarded structurally equivalent?

According to the general principle of cognitive economy (Ahlbäck, 2004: 37, Cambouropoulos 1998: 27), we are consistently and spontaneously trying to reduce the amount of information that needs to be processed by generalization and simplification of structure. This implies that any aspect of a structure which is repeated may serve as the basis for cognition of melodic grouping. The more phenomenally chaotic the structure, the lower the threshold for what is regarded as structurally equal or equivalent. Conversely, in a well-ordered structure, more general similarity between structures will be obscured by more specific similarity.

This implies a classification of types of melodic similarity, ranging from more specific to more general similarity (described below in section *3.3*), based on different dimensions of melodic surface structure (Cf. Conklin & Anagnostopoulou, 2001; Lartillot, 2004). These dimensions include aspects of temporal structure such as rhythm, meter and form and aspects of pitch structure such as pitch contour, register and pitch set (Ahlbäck, 2004: 40-43). Most importantly, the dimensions of melodic surface structure can, according to this model, be assigned to groups of events as well as individual events, forming a global surface structure (see section 3.3, fig. 19-20). In addition, the different types of melodic similarity have to be classified by their level of structural prominence, in order to evaluate the structural significance of the sequencing of different aspects of structure (see section 3.4).

3.2.2 Significance of similarity in relation to metrical structure

What then, is required for a repeated structure to be recognized as structurally significant? In the current model this is assumed to be strongly determined by the general properties of the structure, in particular the temporal structure. This means that a repeated sequence will be recognized as structurally significant only when it complies with the fundamental temporal structure of the melody. In the current model, the metrical structure is assumed to provide the fundamental temporal grid in relation to which melodic events are conceived whenever a metrical structure exists (see Ahlbäck, 2004: 41-2, 73-82). In other words, the metrical structure creates a general basic temporal context for the melody structure. The most important level of the metrical temporal grid is, according to this model, the central beat level. This implies that events on beats are structurally more prominent than events that occur between beats.

As an example, consider the two constructed melodies displayed in figure 9 that were used in an experiment (see Ahlbäck, 2004: 410ff).

The two melodies were presented to eighteen test participants in deadpan performances, played with identical articulation and duration of tones by a computer[11]. The experimental task was twofold: to notate a metrical structure by grouping notes

(11) The test group consisted of 18 persons. There were equal numbers of men and women, and the age range of participants was from 18 to 72 years. All participants were musicians with knowledge of common Western notation. Thirteen of the participants were music students at the Royal College of Music in Stockholm, studying folk music, jazz, classical and popular music. The other participants were amateur performers and professional musicians/composers. The stimuli in the tests were deadpan performances of the melodic excerpts created in Finale 2002 software on a Macintosh computer, transformed into audio files using the "Grand piano" sound patch of the synthesizer module Roland JV-1010. The velocity parameter was set to a moderate level (44).

Figure 9.

Melody examples, stimulus versions a and b. M.M. (eighth note) = 240 The repeated sequences are indicated by brackets. (Ahlbäck 2004:411).

into beats and measures and to segment the melody into phrases[12]. The *stimulus version a* was presented and interpreted first, giving structural implications for the interpretation of the second melody[13].

The melodies were constructed in order to imply a 3/8 central pulse level, by means of sequence repetition and discontinuity within the first half of the melody (see fig. 12). The second half of both melodies contained a perfect repeated pitch sequence, marked by brackets in the notated example above. In the case of the first melody, the implied beat structure was not congruent with the repeated sequence. The question to be considered was whether the sequence would be recognized as structurally determinant even though it interfered with the implied beat structure, showing in notated phrases or a change of metrical structure. In the case of the second melody the sequence concurred with the implied 3/8 beat structure but interfered with the implied 9/8 measure structure.

(12) The test participants were given a score with equally spaced eighth notes, similar to those in fig. 9 but with the system breaks at other positions. The test participants had no time to study the score before listening to the musical examples. The instructions were as follows: 1. Listen to the recording without looking at the score. 2. Listen to the recording and try to follow the music with the score. 3. Mark out in any way you find suitable (rings, brackets, beams etc.): pulse, time and measures, phrases. The subjects were allowed to listen to the stimuli several times, but were instructed to notate the first structural conception that came to mind.

(13) The reason for presenting the melodies in the same order for all test participants was partly to examine whether the predicted symmetrical phrase structure of stimulus version a would influence the interpretation of the stimulus version b, partly to avoid influence of the predicted recognition of sequence repetition in stimulus version b in the cognition of stimulus version a.

In both cases the listeners unambiguously notated the intended 3/8 pulse indication, which exemplifies that a metrical structure can be inferred even from dead-pan performances of music with pitch change as sole perceptual cue. Would this metrical conception affect the structural interpretation of the second half of the melodies?

In the case of the first melody the results showed that the majority of the listeners did not recognize the grouping structure implied by the sequence, but preferred a segmentation which was concurrent with the metrical structure at measure level, as indicated by slurs in the notation in fig. 12. In the case of the second melody the sequence was recognized by most participants as structurally determinant, showing mainly in the indicated phrases (fig. 10-12).

The result of this test can be interpreted as demonstrating that even an exact repetition of a pitch sequence may not be recognized as structurally determinant if it is not congruent with the fundamental metrical structure[14]. This implies that even an exact pitch repetition may not be *similar enough* for it to be structurally determinant when presented in a non-congruent metrical context. This is an example

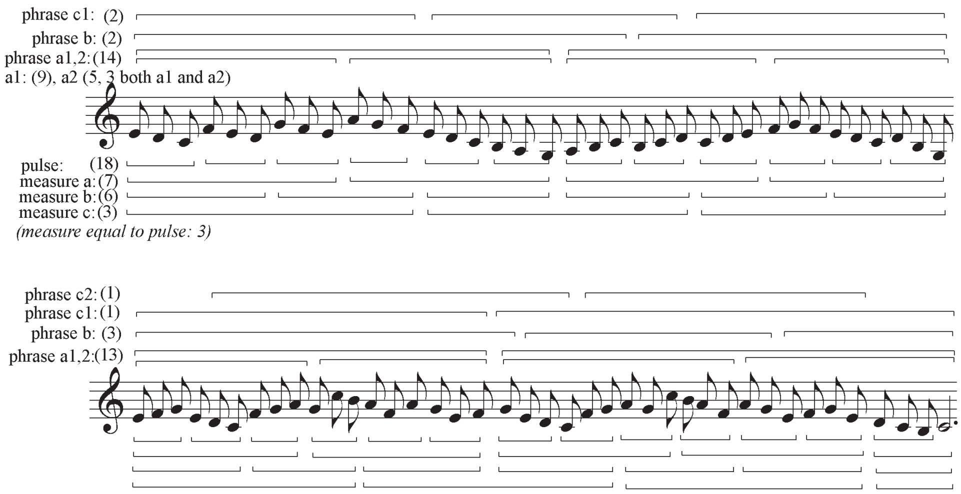

Figure 10.

Test results for stimulus version a. Notated interpretations are marked by brackets below (for metrical structure) and above (for phrase structure). Numbers within parentheses denote the number of test participants who notated the interpretation that follows the number. The hierarchical phrase interpretation a1, a2 was notated by three of the subjects.

(14) This is not to say that true repetitions — perfect sequences — may not be recognized when they do not correspond to beat structure. There are plenty of musical examples of melodic variation that exhibit phase shift of melodic gestalt with regards to meter (in-phase to out-of-phase or vice versa).

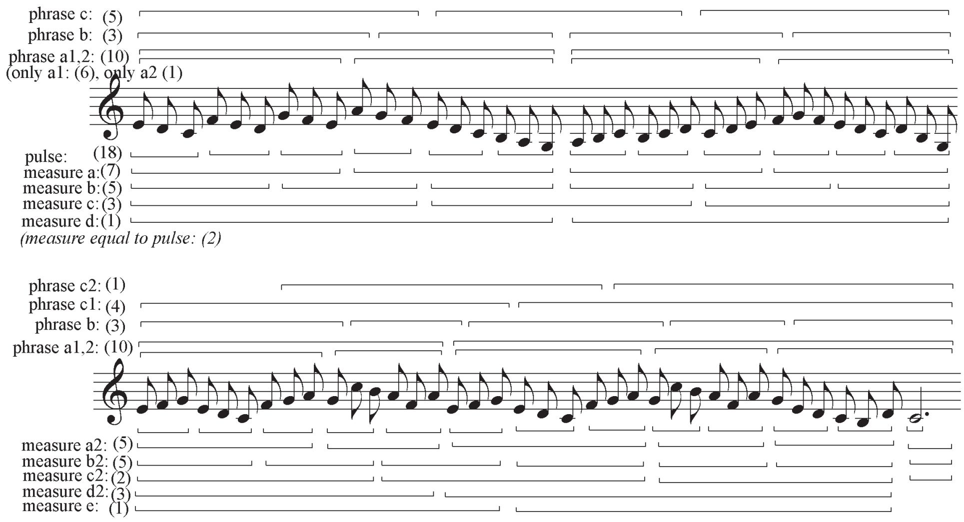

Figure 11.

Test results for stimulus version b. Notated interpretations are marked by brackets as in fig. 11. Numbers within parentheses denote the number of test participants who notated the interpretation that follows the number.

Figure 12.

Summarized results of listener evaluations of both the metrical and phrase structure of test examples. (Ahlbäck 2004: 411ff). Notated metrical structure displayed by time signature and correspondent beaming. Most common phrasing denoted by slurs.

of how melodic context, in terms of the basic structure, influences similarity recognition. The second test example shows that a sequence that concurs with the metrical structure at beat level but interferes with implied measure level structure might be conceived as indicating a change of metrical structure at measure level. Thus, the results support the assumption that when present, metrical structure serves as the basic temporal grid of structural prominence, with the power to obscure grouping indications even by true repetition of phenomenal structures.

If fundamental metrical structure determines the prominent temporal positions in metrical music, we can also assume that similarity in pitch change between beat positions can be sufficient to recognize a repeated sequence as structurally identical. Thus, it is assumed in the current model that pitch and duration change between beat positions is the most structurally determinant level of pitch and duration change in metrical music[15].

Support for this assumption can be found in examples of melodic variation in music, where structural equivalence of the repeated structures is presupposed. The following example of melodic variation is given by Quantz in his "*Versuch eine Anweisung die Flöte traversière zu spielen*" (Quantz, 1752/1966).

1

Figure 13.
From Quantz (1752/1966: 141), chapter 13, "Of Extempore Variations on Simple Intervals".

(15) In non-metrical melody the temporal points of structural prominence can be denoted as structural event salience, mainly determined by interonset interval and offset-onset interval. For examples of analysis of structural similarity in non-metrical music, see Ahlbäck, 2004: 483ff.

All but two of the melodic variations of the original scale motif, preserve the original pitch transition between the initial notes of the beats, c-d-e. In the two variations that differ, the initial note of the second beat is transposed by one octave. This indicates that the pitch change between beats was conceived by Quantz to be structurally prominent.

The above discussion demonstrates the primacy of fundamental metrical structure in metrical music. Therefore, the current model is based on the analysis of pitch and duration patterns between groups of events that correspond to beats, *beat-groups* [16]. Consequently, pitch and duration change is primarily encoded beat-wise in the sequence analysis.

3.2.3. How much needs to be similar in order for it to be structurally significant?

The question put forward in the title of this section — what is similar enough? — may also be interpreted as referring to how much of a sequence needs to be repeated in order for it to be structurally determinant. As has been pointed out by e.g. Lartillot (2004), successive similarity links events together and thus is of special importance in the formation of temporal structures. The similar part of the sequence may thus be interpreted as the number of successive similar events (changes or elements) or the temporal extent to which the similar part is needed for a sequence repetition to imply structural equivalence. Within the current method, the answer to this question is two-fold: it relates to (i) how large a part of the sequence is being repeated and (ii) to the temporal and categorical extent of the repeated part of the sequence.

The first point, (i), regards the prominence of a sequence repetition as being related to the structural integrity of the segment, which can be expressed as the degree to which the repetition unambiguously defines the structure at a certain hierarchical level (at a given temporal position). This is partly related to how large a part of the implied temporal structure is addressed by the sequence repetition, i.e. is repeated. The basic threshold for a significant temporal structure defined by sequence repetition in this respect, is interpreted in the model as a situation where the sequences are not more dissimilar than similar, in terms of structural equivalence at corresponding temporal positions. This implies a basic repetition to be regarded structurally significant if the equivalent part between the original and the repetition is at least as great as the difference, which means a basic threshold of 50 % similarity [17]. When more than half the sequence is repeated the differences may be

(16) In the analysis of phrase and section structure, this beat level concurs with the central pulse or tactus level. In the metrical analysis, for which the primary aim is the analysis of the tactus level, the reference "beat" level is a lower level metrical level. The concept *beat-group* refers to a primary group delimited by a beat period. N.B. Non-metrical music is treated differently.

(17) The first instance of a sequence repetition is herein called *original*, while the second instance is called *repetition*.

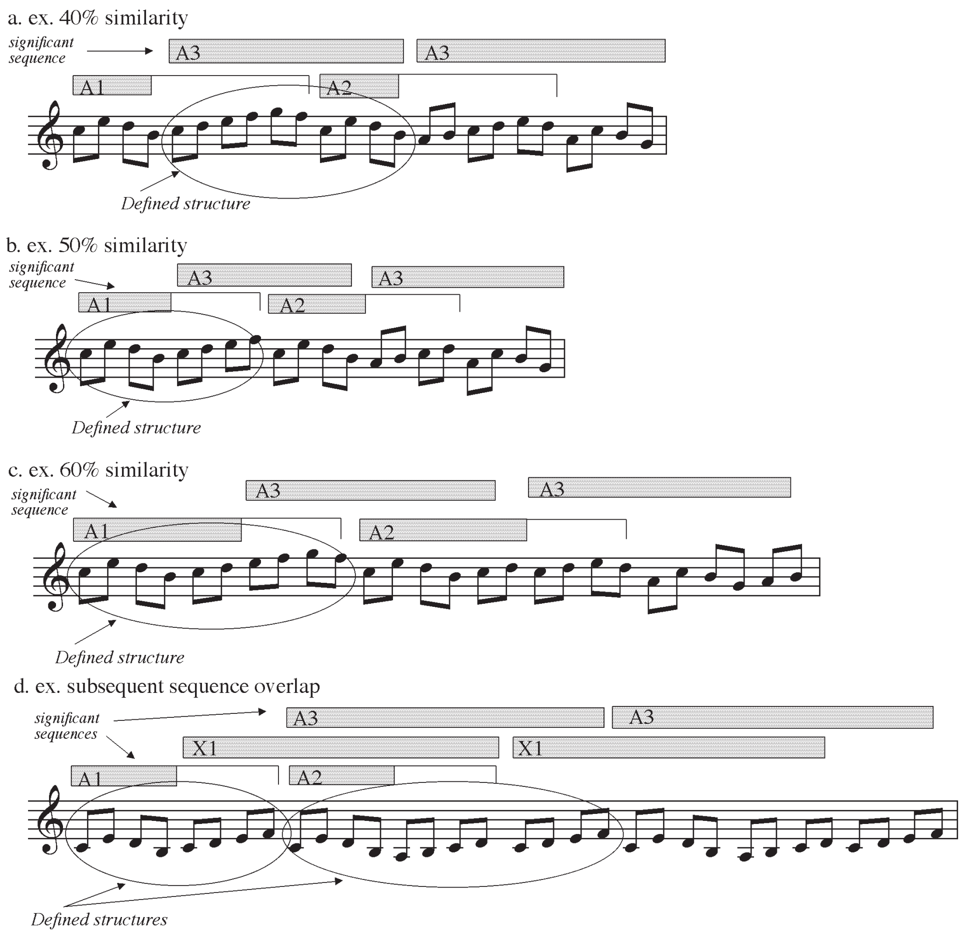

Figure 14.

Illustration of the 50 % limit, which is one of the two alternative basic requirements for structural significance by sequence repetition. Sequences are illustrated by brackets. Shaded areas designate repeated parts of sequence. Enclosed areas denote defined structures by sequence repetition. In ex. a only 40 % of A1 is repeated in A2, which makes the A1-A2 sequence non-significant. Thus, the overlapping sequence repetition A3-A3 will be significant. In ex. b 50 % and in ex. c 60 % of the original sequence A1 is repeated in A2, which makes the A1-A2 sequence repetition significant. Ex d exemplifies continuous sequence overlap: The sequence rep. A1-A2 makes the X1-X1 true repetition non-significant, thus defining A1 as a segment since 50 % of the A1 sequence is repeated in A2. A2 becomes a part of A3 through sequence repetition, which exemplifies that only the original and not the repetition becomes a structural unit by sequence repetition.

considered insignificant. The original is confirmed as a temporal structure, a segment, since no overlapping sequence repetition at the same hierarchical level can *uniquely* address a larger part of the implied temporal structure (see fig. 14). However founded on logical considerations, the relevance of the threshold is also supported by results from listener studies regarding interpretation of sequence overlapping (Ahlbäck, 2004: 192ff and 441ff).

However, the formation of structures in the listening process is also ruled by our fundamental perceptual and cognitive capacities, as is stated in the second point (ii) above. In the general model of melody cognition which this method of analysis is based upon, it is assumed that structural information is processed within certain temporal scopes of attendance, designated "parallel nows". The fundamental span of immediate attention is called a "*basic now*" and is confined to the temporal limitations given by the perceptual (or psychological) present[18] (Fraisse, 1978) and the capacity of simultaneous category handling within short term memory (Miller, 1956), herein designated the categorical present[19]. This implies a limitation to the span of structural units, and indeed, studies of phrase durations in music have provided support for the existence of a perceptual present (Fraisse, 1982). This suggests that the absolute duration of and number of elements within a repeated sequence will influence the structural prominence of the sequence, since the identity of a structural entity will be determined within the scope of the *basic now*[20] (cf. Cambouropoulos, 2005). This means that sequence repetitions in which the similar part extends beyond the limits of the *basic now* may be structurally significant even if the repetition does not involve more than 50% of the original sequence.

The influence of the two factors of the extent of the repeated part in relation to the whole of the sequence and of the extent of the repeated part in relation to the *basic now*, constitutes a basis for a quantification of structural prominence of repeated sequences.

(18) There is substantial converging psychological evidence that indicates the perceptual reality of a time-span at which stimulations can be perceived at a given time, without the intervention of rehearsal during or after stimulation (Fraisse, 1978: 205). The temporal extent of this timespan, which is here termed the perceptual present, is generally estimated to be between 3-8 seconds (Clarke, 1999: 476). Most frequently, a value around 5 seconds is mentioned (Fraisse, 1982).
(19) According to Miller we are able to separate 7 ± 2 categories of a unidimensional stimuli within "the span of immediate memory": "I would propose to call this limit the span of absolute judgment, and I maintain that for unidimensional judgments this span is usually somewhere in the neighborhood of seven." (Miller 1956: 90). I am here asserting that this general rule may apply to categorization at all levels, delimiting the number of subcategories within a given category. This rule applies to the process by which a representation of a structure is acquired from the perceptual present, the span of immediate attention.
(20) Cambouropoulos has adopted a similar view in a recent paper (Cambouropoulos, 2005).

Figure 15.
M.M (quarter note) = 96. Illustration of sequence repetition, which is significant since the repeated part of the sequence spans a basic now, providing it is not overridden by a more complete sequence repetition at the same or a higher hierarchical level.

3.2.4. Summary of theoretical fundaments

To summarize, the method of sequence analysis is built on some theoretical concepts that can be regarded as hypothetical answers to the questions that were put forward in the beginning of the article; what is structurally significant melodic similarity? Which musical parameters have significance for melodic similarity? What are the criteria for judging two entities as sufficiently similar to be considered equivalent? Is there a threshold value for melodic similarity?

• Similarity is employed in the cognitive formation of melodic substructures, i.e. melodic segmentation, primarily by the special case of similarity designated *sequence* repetition. This is defined as the determination of a structural unit by the repetition of a melodic process, i.e. a perceived change in a structural parameter which is conceived to be contextually prominent. In terms of melody structure, sequences can be defined as contiguous arrays of events, involving change of pitch and/or duration, which are conceived to be structurally equivalent.
• Structurally determinative similarity can encompass any aspect of melodic surface structure, determined by the general dimensions of melodic surface structure. This involves for example, categorical perception of pitch and duration change between individual events as well as groups of events.
• Melodic segmentation is assumed to be a spontaneous process driven by the principles of cognitive economy, coherence and significance, which implies that recognition of any level of similarity regarding pitch and/or temporal structure may be structurally significant.
• The structural prominence of a sequence repetition is assumed to be governed primarily by the secondary and tertiary grouping principles: continuity, symmetry, perceptual salience and structural integrity. This implies for example, that sequences with more specific similarity are more structurally determinant than sequences of more general similarity. Similarly, sequences will also be prominent if they are

consistent with previous grouping, create symmetrical division, concur with accent structure, are structurally unique and are structurally discrete.

• The structural prominence of a sequence is also determined by relative prominence of different general aspects of melody structure. In particular, metrical structure is assumed to influence the recognition and salience of sequences so that sequences that are congruent with metrical structure have higher structural prominence. Similarly, pitch patterns that are similar with regards to the general pitch organization in the melody — scale structure — are more structurally prominent than pitch change patterns which are different with regards to pitch categorization.

• The structural prominence of a sequence is related to the number of temporally prominent events it can uniquely address. This implies that longer sequences determined by discrete pitch change are more prominent than shorter sequences. This further implies that longer sequences at higher hierarchical levels rule above shorter sequences at lower hierarchical levels.

• The structural prominence of a sequence relates to how many constituting structural changes or elements are similar between the original and the repetition. The general threshold of structural equivalence is that the repetition should not be more dissimilar than similar, implying that at least half of the constituting elements or changes should be structurally equal in the repetition for the units to be structurally equivalent. However, extensive sequences with similarity that extend the limits of the perceptual and categorical present are assumed to fulfil the fundamental requirements for structural implication and are regarded in the analysis (see Ahlbäck, 2004: 291), with reduced structural prominence.

• Besides sequence repetition, structures may also be formed by similarity with previously established structures. This reflects that the formation of categories is a temporal process where previously identified entities will be a reference for subsequent identification of structures according to the general principle of cognitive economy. This implies that once a structure is defined by e.g. sequence repetition, subsequent structures will be compared to this entity, giving higher structural prominence to passages which are similar to previous structures. The defined structures constitute the category *global melodic motifs* within the current model[21].

3.3. WHAT IS SIMILAR? — CLASSIFICATION OF MELODIC SIMILARITY WITHIN THE MODEL

In the current model sequences are identified, as previously mentioned, through valuation of melodic similarity according to different properties of melodic surface structure. This implies a categorization of different types of melodic similarity, reflected in a classification of types of sequences at different levels of structural significance. Similarity is basically valuated with regards to pitch content, pitch contour and

(21) c.f. the *imprint* concept introduced by Deliège (Deliège, 2001: 238)

interonset/duration patterns. Similarity within each dimension is valuated varying from more specific to very general similarity[22]. The main structural constraint of the sequence concept is that a similarity must be recognized at correspondent temporal positions for the repetition to be structurally significant.

There are two main types of sequences based on successive pitch similarity, sequences based on similar pitch content and sequences based on similar melodic contour in terms of transitions between Melodic Pitch Categories (see below). Both of these belong to the same class, the highest order of sequences. These sequences are labelled as pitch or interval-sequences, **PI-sequences**, which reflects that the structural significance of the sequence repetition is regarded as equivalent for similar pitch content and similar melodic contour. This implies that the structural implications of transposition and repetition are not regarded as being categorically different. The basis of the measuring of similarity regarding melodic contour is the concept of *Melodic Pitch Categories* (MPC) (Ahlbäck, 2004: 45ff). *Melodic Pitch Categories* represent the structurally determinant level of pitch change in melody and can be regarded as the "bricks" of pitch, which we use to form, perceive and cognize pitch change within a melody[23]. This is typically manifested in a scale structure. Melodic Pitch Category is similar to the concept "scale degree", except that in general usage scale degree seems to imply a reference to a tonic, which does not have to be present for a Melodic Pitch Category to be conceived (see e.g. Dowling & Harwood, 1986: 152).

(22) The designations of the sequence types which represent different types of similarity originate from the computer implementation of the model. In this paper the original designations are changed in some cases to make the presentation easier to follow, and so differ from the original descriptions in Ahlbäck (2004). What is here denoted as PI-sequences corresponds to NIL-sequences in Ahlbäck (2004). Similarly D-seq. corresponds to T-seq. and B-seq. to I-seq. For all other types, the designations are kept unchanged. The sequence types are designated by a code, e.g. PI 3 or R 4. The first item refers to the type of sequence, while the second item refers to the level of structural prominence within the sequence type.The use of the type designations in this article is solely to make it easier to follow the presentation.

(23) The MPC concept is derived from the assumption that the pitch change pattern within the melody determines the categorization of pitch in melody. Basically this means that pitches that are structurally separated by melodic motion- appear in immediate succession — and belong to different pitch categories, while neighboring pitches (with regards to pitch) that occupy identical position within the melody structure, e.g. follow and are followed by the same pitches, are cognized as variants of the same pitch category. This implies that two melodic passages which differ only with regards to interval size, e.g. played in "major" or "minor" mode, are conceived as "the same" melodic passage in terms of MPC change — a major-minor variation of the same melodic pitch categories — in spite of the actual pitch variation. This is strongly supported by different experimental studies, which demonstrates for example the inability among general listeners to separate between perfect transpositions and tonal imitations (see e.g. Dowling & Harwood, 1986). For a more thorough presentation of the concept, see Ahlbäck, 2004: 45ff.

PI-sequences are defined by successive beat-wise similarity from the first event in the group (see fig. 16 and 17). The similarity of a PI-sequence is categorized into different levels representing structural specificity and prominence of the sequence. They range from ***specific pitch set similarity, level 5*** which designates exact pitch repetition — perfect similarity with regards to pitch content, to ***specific pitch contour similarity, level 4*** which designates repeated interval pattern. The more general levels of pitch similarity include ***general pitch set similarity, level 3,*** which designates similarity regarding pitch content (pitch set) in terms of highest/lowest pitches of successive beat-groups — same pitch set but different order; and ***general pitch contour similarity, level 3*** which designates similarity in terms of pitch distance and melodic direction between single beats and between pairs of beats. The *level 3* categorization has considerable flexibility regarding pitch distance and melodic direction within beats.

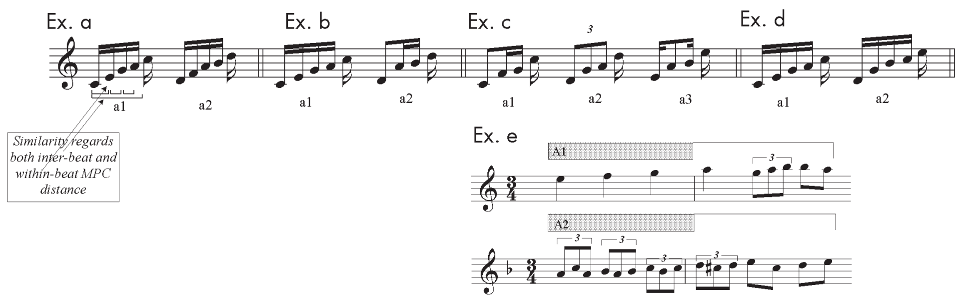

Figure 16.

Examples of type PI similarity, level 4 and 3, **MPC Contour similarity**. *Beat-groups a1, a2, and in ex. c also a3 are compared and exhibit MPC contour similarity (a-c level 4, d level 3), successive pitch contour similarity from first beat-group of the repeated sequence.*

a. Equal MPC contour (including inter-beat and within-beat distance), equal rhythm (duration pattern) but pitch interval differences.

b. Equal MPC contour between corresponding rhythmic within-beat positions, equal inter-beat distance, rhythmic difference.

c. Equal MPC within-beat and inter-beat contour, division of beat in equal number of events.

d. Equal MPC contour except for inter-beat and within-beat differences within categories steps and leaps, similar rhythm.

e. Musical example of MPC contour similarity, PI level 4, over three successive beats (shadowed section) between original and transposed variation displayed on different systems in Polska after Bengt Bixo (Andersson, O. 1926: 18).

Figure 17.

Example of type PI-similarity, level 5 and 3, **pitch set similarity**. *Beat-groups a1 and a2 are compared and exhibit type PI pitch set similarity (a-c level 5, d-e level 3), successive pitch set similarity from first beat-group of the repeated sequence.*

a. Equal pitch set and pitch contour including inter-beat pitch and within-beat pitch on corresponding rhythmic positions, rhythmic difference (different duration pattern).

b. Equal pitch set and rhythm but different pitch order.

c. Start-variation. Equal pitch set and order except for initial pitch, initial rhythmicdifference

d. Pitch set extremes. Equal lowest pitch and register change, but rhythmic and within-beat pitch set difference

e. Pitch set extremes. Equal highest pitch and register change, but within-beat pitch set difference.

f. Musical example of max-min general pitch set similarity, PI, level 3. Equal sequence is shadowed and displayed on different systems.

R-sequences (reversed similarity) and ***O-sequences*** (overlap of first beat) can be said to mirror PI sequences since they are based on successive pitch-set and/or contour similarity, and concern the same levels of similarity. However, R-sequences regard structural equivalence by similar ending and O-sequences similarity from the second beat in the sequence, allowing for start-variation (see fig. 18).

B-sequences (general pitch contour similarity implied by marked sequence boundaries, see fig. 23) and ***C-sequences*** (general pitch contour similarity, see fig. 22) designate sequences with the most general similarity regarding melodic contour and pitch content. These do not require successive similarity on the same level between individual beats, and can be regarded as representing the more general pitch similarity at specificity levels 1 and 2 (cf. PI-sequence type). Successive similarity is required, but exceptions are allowed, and similarity can vary between very specific to more general. Identification of sequences is therefore based upon total similarity value, global similarity and how discrete the sequence boundaries are in relation to the content of the sequence. Especially C-sequence type similarity is very important within the model since it allows for determination of similar melodic contour in spite of more elaborate melodic variation.

C- and B-type similarity are based on the encoding of pitch change between beat-

Figure 18.

Example. R sequence. Sequence repetition identified by successive end similarity at level 5 — equal pitch set. Shadowed areas represent significant similarity between sequences A1-A2. R-sequences replicate the PI-sequences similarity types and levels, but have lower structural significance.

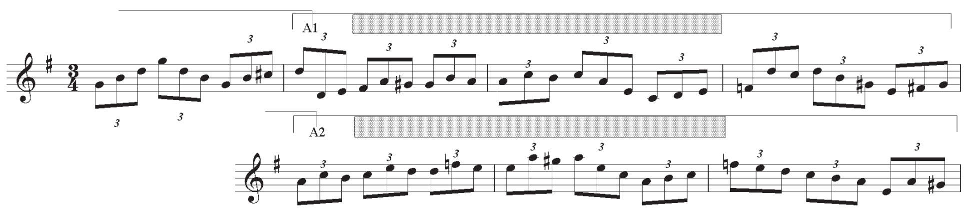

Figure 19.

Example of pitch contour sequence repetition (A1-A2) with start variation in excerpt from Obligatto melody in "Jesu Bleibet Meine Freude" in Cantata no. 147 by J.S. Bach, identified in listener evaluation and computer model (see Ahlbäck 2004:483ff). Sequence repetition A1-A2 identified by successive similarity at level 3 — general pitch contour similarity, from second beat-group of sequence. Shadowed areas represent significant similarity. O-sequences replicate the PI-sequences similarity types and levels. The overlapping sequence boundaries represent the two alternative interpretations, given both by listeners and computer modeling (see Ahlbäck 2004: 436ff).

groups as a change of register. This means that differences between beat-groups with regards to MPC change and pitch set *within* beat-groups are more or less disregarded. Instead, beat-groups are treated as wholes where the total MPC difference between beat-groups is measured basically in terms of MPC register set-difference and relative-pitch-mean difference (see fig. 20 and 21). The melodic contour is regarded as similar at two instances when registers change, general change between initial pitches of beats and the relative mean pitch value of beats concur. In addition the global changes of pitch register, melodic direction and duration must correspond for a C sequence repetition to be recognized.

Since C- and B-type similarity is based on a quantification of the total successive similarity between two passages, the specificity level of each pitch change between beat-groups is measured by a numerical value which contributes to the total similarity value (see fig. 22). The values designate different types and levels of

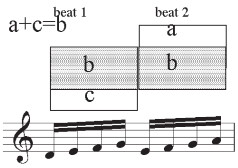

Figure 20.

Illustration of the register intersection limit for pitch change between beat-groups. If pitch register intersection is <= 2/3 it may be regarded as a change of pitch, otherwise it will be regarded as a repetition.

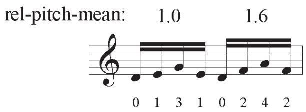

Figure 21.

Illustration of quantification of pitch change based on relative pitch mean index. The mean of the relative MPC values for beat-groups are displayed above system with relative MPC values below. The initial MPC of the beat-group is counted twice. For the first beat: (0 + 0 + 1 + 3 + 1) / 5 = 1. Approved pitch change >= 1/3.

Figure 22.

Example of C-type similarity, i.e. non-successive pitch register change similarity, for two variations of the same melody (Andersson: 1926: 36). Arrows above beats designate local change of pitch contour to succeeding beats based on local change of pitch (MPC) register and pitch transition between initial notes. Similarity rating values are displayed between systems.

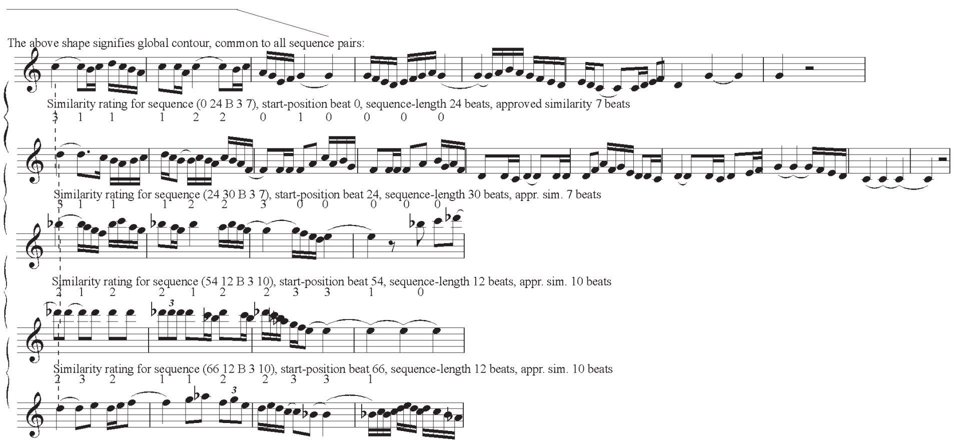

Figure 23.

Example of B-sequence similarity. Excerpt from solo part of Boléro by Ravel (Ravel 1928). Segments are classified as pairwise similar at general level while segmentation is mainly determined by global discontinuity at segment boundaries (extreme interonset intervals and global register changes). Similarity values between systems refer to C similarity measure (for explanation of sequence codes (0 24 B 3 7) etc. see fig. 28).

similarity, where 3 indicates PI level 4 or 5 similarity, 2 indicates general pitch contour similarity on C level (according to the measure presented above) or PI level 3 similarity, 1 indicates global similarity (see X sequences below) and 0 indicates compatible global structure. The equivalence classification is based on the number of successive similar beats and the total similarity value and requires in addition similar global pitch contour, global register shifts and global duration pattern changes (see further Ahlbäck, 2004: 276ff).

X-sequences: (global similarity, see fig. 24) are based on similarity with regards to global structural changes such as global change of melodic direction, global shift of register and global rhythmic breaks. They require discrete sequence boundaries in relation to the content of the sequence. X similarity is determined by similarity between *groups* of beat-groups, consisting of alternatively 2-, 3 or 4 beats depending on at which level the most persisting change takes place. Since X similarity is determined only by continuity and change, it covers for example melodic inversion as a structurally implicative factor. X similarity has the lowest structural significance of the different types of similarity (sequence types)

D-sequences (duration pattern sequences) are based entirely on successive rhythmic similarity between beats (see fig. 25). Like PI-sequences, the type of similarity is classified into different sub-types, representing levels of structural specificity. ***Level 1*** concerns similarity regarding duration of events in relation to the beat. It is based on a classification of beat-groups into divided beats, undivided beats, tied beats and beats with no onset at the beat start. This classification, denoted *dur-*

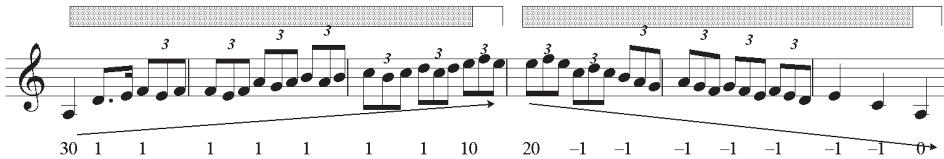

Figure 24.

Example of X sequence repetition determined by global melodic direction. The melodic example is divided into two segments of nine beats with continuous global melodic direction. Numbers below the system refer to the encoding of global melodic direction. 30, 10 and 20 refer to indication of change of global direction, while 1, -1 and 0 refer to upwards, downwards and no change of melodic contour between groups of beats. Shadowed areas indicate similar part of sequences.

Figure 25.

Examples of D type similarity (Duration pattern similarity).

a. D3 similarity — values designate successive rhythmic relationships between events within and between beats. A1 — A2 is categorized as equal, long(l)-short(s), ssl, ss, l

b. D2 similarity. Division of beat into equal number of events. Numbers refer to level of division.

c. D1 similarity. Equal with regards to categories divided (1) and undivided (2) beats. Numbers between systems refer to categories

class, is assumed to be the most general level of rhythmic structure in a metrical melody. ***Level 2*** adds a classification of division of the beat to the *dur-class* concept. This means that beats divided into two events are regarded different to beats divided into three events etc., but similar within the category regardless of the actual pattern of duration within the beat. ***Level 3*** designates actual rhythmic similarity in terms of patterns of relative duration between *beat-group*s. ***Level 6*** is employed only when there are beats of different length in the structure and designates similarity in terms of beat length.

3.4. EVALUATING THE STRUCTURAL SIGNIFICANCE OF MELODIC SIMILARITY.
A melody may contain many overlapping possible sequence repetitions of different sequence-types at different levels of specificity of similarity (see section 4). In principle, all these possible sequence repetitions are considered in the sequence analysis. They are then evaluated successively based on structural discontinuity (section 3.1), the aforementioned principles that govern structural prominence of sequence repetition (section 3.2.3) and a valuation of the structural prominence of the different sequence types (section 3.3).

The scope of this article does not allow for a detailed description of this evaluation process (see Ahlbäck, 2004: 288ff). However, an overview of the principles involved and some of the main points regarding how melodic context is involved in the sequence evaluation will be given below.

The basis of the evaluation is that each sequence repetition is given a basic prominence value, which is the sum of the discontinuity values[24] for the implied segment boundaries, reflecting that sequence phase is primarily determined by the *grouping by difference principle* (section 3.1). This basic value is then influenced by different factors relating to the structural prominence of the sequence, determined by the quality of the sequence and the context in which it appears. The product of the basic value and the prominence factors represents the structural prominence of the sequence. The factors are:

- *Sequence-type* factor, reflecting the perceptual salience and specificity of the similarity.
- *Sequence-length* factor (in terms of metrical units, beats), reflecting the temporal prominence of the sequence repetition.
- *Sequence-completeness* factor, based on how much of the sequence is repeated, reflecting the salience of the sequence.
- *Distribution of boundary indications* factor, reflecting the grouping by perceptual salience principle.
- *Sequence prägnanz* factor, based on, for example, a valuation of rhythmic and pitch pattern uniqueness of the sequence in its local context, reflecting the grouping by structural integrity principle.
- *Degree of overlapping* factor, reflecting the grouping by structural integrity principle.
- *Latticity* factor, based on the prominence of structural division by sequence repetition and discontinuity within the implied sequence, reflecting the grouping by

(24) The discontinuity value for each sequence boundary (start-, mid and end-position) is based on the group boundary indications by the difference principles, such as e.g. pitch or interonset distance.

structural integrity principle. The *latticity* factor (my own term) refers to the degree to which a structure is divided into coherent, prominent substructures. The cognitive assumption behind this factor is that the more prominent the substructural level, the less prominent the super-structural level.

• *Previous sequence similarity* factor, based on similarity and congruence with previous sequences at the same hierarchical level, reflecting the contextual influence of good continuation.

• *Match with defined structures* factor, based on the similarity to already defined segments, reflecting influence of context by temporal order of category formation.

• *Defined start-position* factor, based on segment boundaries defined by previous segments, reflecting the influence of context by the principle of good continuation.

• *Metrical position* factor, reflecting the contextual influence of metrical structure (an aspect of general musical structure.)

• *Symmetrical-division* factor, reflecting the contextual influence of symmetry.

As is evident from the above list, the influence of melodic context on similarity recognition is involved in a number of these factors. The influence of context, in terms of the general level of homogeneity vs. diversity in the melody, is basically modelled by the influence of the sequence-type factor. This factor is based on a ranking of the structural prominence of different sequence-types. The fundamental assumption behind this valuation is that the more specific and multidimensional the similarity, the more structurally prominent it will be. In addition, it is assumed that sequence repetition that involves pitch change is structurally more determinant at larger temporal scopes than rhythmic similarity. The maximal influence of the different sequence types can be read from Table 1.

By means of this factor, the sequence evaluation will adapt to the level of specificity of the sequence repetitions in the melodic context. In a context with no true repetitions, sequences determined by more general similarity will prevail, while in a context with more specific repetitions, the sequences determined by more general similarity will be disguised.

The influence of melodic context in terms of the general structural features of melody is inherent in the sequence classification, which describes pitch and duration change in relation to metrical structure and MPC structure. In addition, the *metrical position factor* further enhances this relationship.

The influence of context in terms of the temporal order of category formation is modelled primarily through the successive sequence evaluation, which gives priority to earlier sequences before subsequent sequences (section 3.2.3). Moreover, temporal order of category formation is also modelled through the *Previous sequence similarity* factor, *Defined start-position* factor and the *Match with defined structures* factor.

Table 1

Sequence type factor. Table shows maximum enhancement factors for different sequence types and levels, also dependent on minimum strength in sequence boundary indication. The type and level designation refers to types described in section 3.3. (R, longer D and X sequences do not receive any type/level factor enhancement)

Sequence type	Max. factor	Restriction
PI 5	4.0	
O 5	3.0	
D 1		extending 3-4 beats
PI 4 / O 4	2.0	
D 6		extending 2-12 beats
D 3		extending 5-12 beats
B 3		
PI 3	1.5	
C 3		
R	1.0	
D 1-3		
X		
X	0.5	extending more than 16 beats
D 6	0.125	extending more than 12 beats

4. Determination of structure by melodic similarity

4.1. An example of a computational melodic analysis

The following section gives an example of how a melodic gestalt may be segmented by sequence analysis and how different levels of similarity become significant depending on the melodic context. The example in question originates from a notation of a bridal march played by the fiddler Per Danielsson from Jämtland, Sweden (Fig. 26. Andersson, 1926: 16).

In this example, as the input to the model is score notation, quantization is not

Figure 26.

First section of Jämtlands Brudmarsch (transl. "The county of Jämtland's bridal march") played by Per Danielsson and notated by Nils Andersson (Andersson 1926:16).

Figure 27.

Note representation of input to the computer model (the MODUS model): Tempo, event list with assigned duration and pitch.

necessary[25]. The input to the model can be considered as a list of events (notes or rests) of certain durations (fig. 27), in this example nominal score durations.

The first steps in the analysis are the MPC analysis and the metrical analysis. In this case the MPC analysis is of no particular interest, since there is no variation of intonation (D major scale). The metrical analysis considers the half note level to be the tactus level, primarily due to the tempo and the interonset structure.

The search for melodic similarity, the sequence analysis, is performed within the module phrase and section analysis, which results in a segmentation of the melody. This implies that all possible sequence repetitions of the different types and levels presented in section 3.3 are analyzed. As illustrated in figure 28, even in a relatively simple melody like this one there can be quite a substantial number of possible sequences.

The large number of possible sequences is mainly due to the vague structural

(25) The model is designed for MIDI files originating from performance (raw midi), i.e. unquantized midi files, but it certainly allows for score notation input as well.

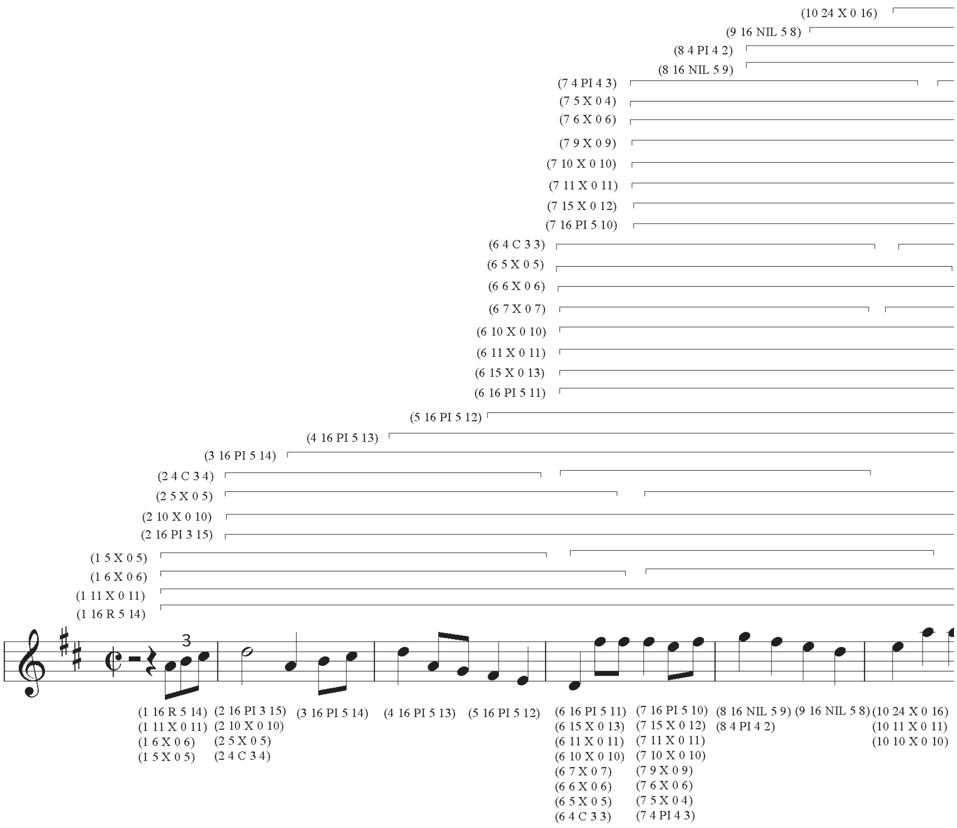

Figure 28.

Listing of the identified sequences in the beginning of Jämtlands brudmarsch. Sequences are designated by the sequence code below the system at the starting point, e.g. (1 16 R 5 14) denoting start-position, length (in beats), sequence-type, specificity level and length (in beats) of similarity. Brackets above the system show the implied segmentation.

indications by global changes of melodic direction and so forth, which can be interpreted in different ways. But even true repetitions may be interpreted in different ways relating to the actual phase of the sequence, and what will be interpreted as the beginning of the sequence.

The assumption is that we will not even recognize the majority of the possible sequences displayed in figure 28, since once we have identified a sequence repetition the recognition of similarities will be confined to the limitations given by the identified sequence. In our example, the PI sequence (2 16 PI 3 15)[26] will, due to its higher sequence type factor than the previous possible sequences (R- and X-sequences) and more beats being similar, receive a higher sequence prominence

(26) Sequences are designated by a sequence code, explained in the caption to fig. 28.

value. However, in spite of a lower sequence type factor than the succeeding possible sequence starting with a true repetition (3 16 PI 5 14), it will override this subsequent sequence since it has higher segment boundary values and no overlapping previous PI sequence.

The result of the sequence evaluation can be viewed in figure 29, which shows the remaining sequences after the evaluation process.

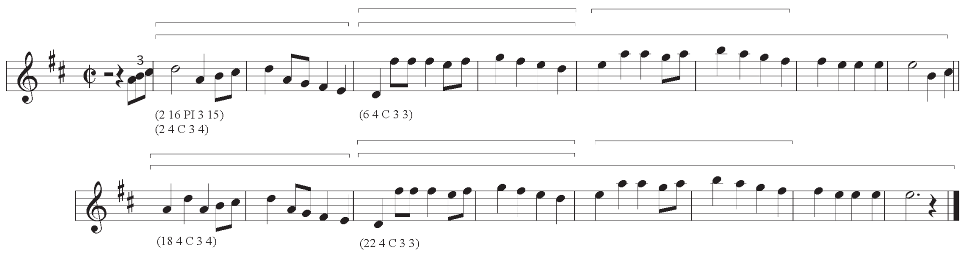

Figure 29.
Remaining sequences after sequence evaluation in the first section of Jämtlands brudmarsch.

Besides the main segmentation determined by more specific similarity (2 16 PI 3 15), the method of analysis also determines a sub-segmentation which divides each of the two major segments into four sub-segments. This segmentation is based on more general pitch contour similarity, a C sequence, which becomes structurally significant because of the congruence with the major segmentation and the lack of competing, more specific similarity. Thus it becomes contextually prominent.

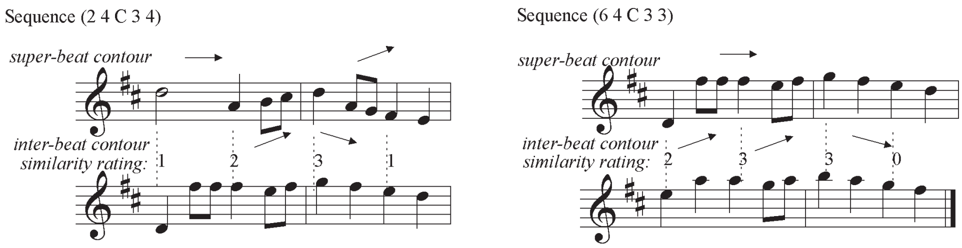

Figure 30.
Structural similarity between two-measure segments in the beginning of Jämtlands brudmarsch. The two systems on the left refer to sequence (2 4 C 3 4) and the two systems on the right to (6 4 C 3 3). The common super-beat contour (pitch change summarized between groups of two beats) is displayed on top of the first system and the common inter-beat contour is displayed between systems. (For explanation of similarity rating, see fig. 19.)

The structural equivalence of this segmentation is supported by expert analyses and listener tests (Ahlbäck, 2004: 348ff and below). However, one may still ask how this similarity is determined since the melodic contour on note-to-note or beat-to-beat levels is actually more different than similar in the (2 4 C 3 4) sequence. The key here is the super-beat level, i.e. the general direction of melodic contour between groups of two beats treated as a register, which can be interpreted as a pattern of same-up, same-up (see fig. 30)[27]. The resulting output segmentation from the computerized sequence analysis is displayed in figure 31.

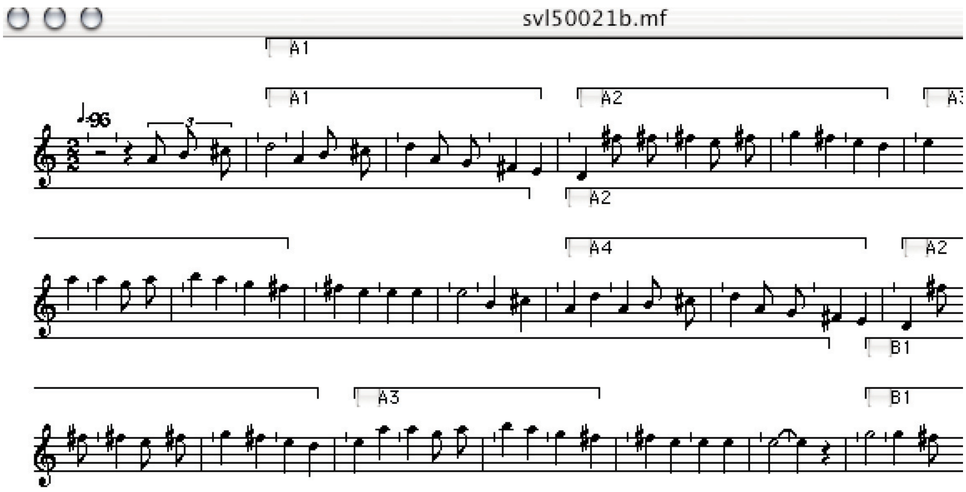

Figure 31.

Output from sequence analysis in Jämtlands brudmarsch. First section. Segments are indicated by brackets at different hierarchical levels. The segments are labeled based on similarity at each hierarchical level, all of the A segments are classified as similar within the same hierarchical level, while the numbers indicate variants within the same class. For example, two segments that are equivalent with regards to PI 5 similarity are given the same number.

4.2. MELODIC ANALYSIS BY SIMILARITY IN THE CONTEXT OF TWO MELODIES

In a comment alongside the notation of *Jämtlands Brudmarsch*, the editor of Svenska Låtar points out that this melody exists in other variants, e.g. an eighteenth century composition "Venus, Minerva" by the famous Swedish poet/song composer Carl Michael Bellman. This song is still well known among Swedes today, who sing it as a birthday song beginning *"Ja må han leva"*.

(27) In particular, the similarity is determined by the registral continuity — same — between the two first beats in relation to the two following, which is broken by the register shift between the second measure and the third measure (see also fig. 18).

The similarity between the bridal march and the composed song is striking when the two melodies are *compared* — even though a minority of the note-to-note transitions and rhythms are identical between the two melodies. The similarity is hence determined by C-level similarity, the basis of which is that the super-beat register changes are identical for the two melodies (fig. 32). (In addition, the local more specific similarities contribute to the total similarity valuation).

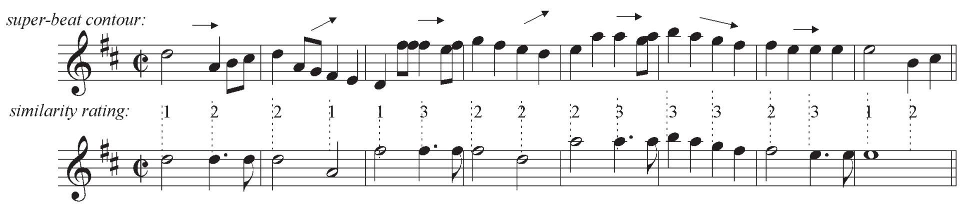

Figure 32.
Comparison between the "Jämtlands brudmarsch" (upper system) and "Venus, Minerva" by C.M. Bellman (lower system). Similarity rating is determined by C-sequence similarity. Common super-beat melodic contour is indicated above the upper system. (For an explanation of the similarity ratings see section 3.3, fig. 20-22).

Thus, when the two melodies are combined (concenated as a single melody), the similarity is recognized as being structurally determinant by the computer model (see fig. 33), which is also mirrored by the categorization of the segments. When the melodies are combined, the model replicates the extension of category of identity, which was exemplified in the variant categorization made by Moberg. The very general similarity between the two melodies is recognized by the model for a number of reasons: there is no competing more specific similarity, the first melody is previously defined as a segment and the two melodies are temporally congruent in the current melodic context.

In order to test if the prediction of the model was valid for this particular example, a small listener test was carried out, involving two groups of ten and eight participants respectively. All of the test participants were students at the Royal College of Music in Stockholm, and had considerable years of formal musical training. Their musical backgrounds differed, but were mainly within popular music. They were all studying to become music teachers.

The first question was whether the *Jämtlands brudmarsch,* when played in isolation, would spontaneously be connected to the extremely well known melody *Venus, Minerva (Ja må han leva)*. The other question was whether the similarity between the melodies would be recognized, and regarded as structurally determinant when they were combined as in the example in figure 32[28]. In order to reduce the

(28) These questions might be regarded as referring to two separate problems that are not necessarily connected, theme recognition and motivic segmentation. However, there is no principal

Figure 33.
Segmentation and categorization of segments by computer model (excerpt) from a combination of "Jämtlands brudmarsch" and "Venus, Minerva. The segment A3 is determined by sequence (18 16 C 3 16) A2-A3, but is also significantly similar on C-level in pairs with A1-A3 and their reversals. N.B. The computer model also provides a second alternative segment start position including the anacrusis of the first segments (see Ahlbäck, 2004).

influence of performance style the melodies were played by a computer program according to the input score (Finale 2004 software, recorder sound), a dead-pan performance. All experimental tasks were presented to the participants in writing.

The first trial involved the group of eight participants. The participants were presented individually with the first section of *Jämtlands brudmarsch* twice, before being asked if they had heard this melody before. If not, they were asked if the melody reminded them of some other melody. Before they were required to answer, the melody was played once more. If the melodies were not yet connected, they were asked if they thought the melody in question had any similarity with *Venus, Minerva (Ja, må han leva)*[29].

---

difference between theme recognition within and between melodies (see also the discussion in section 1.1) Both questions concern formation of a category (of identity) by similarity. While the first question refers only to similarity between two structures, the second also refers to whether a similarity is regarded as structurally determinant, i.e. as having the power to form structures.

(29) It should be noted that most Swedes, and certainly musicians, will be able to sing *Ja, må han leva* at any time, when asked.

The second trial involved the other group of ten participants. The participants were given a notation of the input score of the combined melodies without any notated meter and grouping of the notes (see fig. 27), and were asked to segment the melody and label the segments. No further explanation was given regarding the principle for labelling the segments or, for example, about structural hierarchy. It should be noted, however, that the test participants were music students with everyday experience of segment labelling. The task was illustrated by a pre-trial, in which the test melody was "Twinkle, Twinkle, Little Star".

In the first trial — without the context of the second melody — none of the test participants admitted that they had heard the test melody before. Only three participants out of eight were able to identify the similarity between the melodies after the third listening, when asked to think of a similar melody. However, all of the participants admitted that there was some resemblance to "Ja, må han leva", when asked to connect the two melodies.

The results of the second trial showed a strong concurrence with the predictions of the computer model[30]. All of the participants notated the general structure of the super-phrase repetitions. Eight out of ten participants identified the two melodies as variants, indicated by the same labelling (A-A-A). However, only six of the participants notated the lower phrase level (two measure) consistently. It is not clear whether this result reflected any uncertainty regarding whether or not the task included identifying structural hierarchy.

All together, the results of this test indicate that while even very general similarity may be recognized when it is contextually apparent, it may not be recognized in isolation, even when the reference category is very well established as with *Venus, Minerva.*

(30) That most listeners included the initial pickup in the first super-phrase and that some listeners marked out phrase boundaries at the two-bar level that are shifted one quarter note in relation to the computer model output displayed above may seem to contradict the suggested agreement between the listeners' and the computer model output. However, experiments show that, whether or not a pickup measure is included in a phrase can vary among listeners (Ahlbäck, 2004), while the identification of the segments does not seem to depend on the inclusion of pickups. This is in fact demonstrated by most common labeling of phrases in this trial. The computer model actually predicts all of the segment boundaries marked by the listeners. (see Ahlbäck, 2004: 321ff). In the segment interpretation, based on the sequence analysis, the anacrusis is actually included in the first segments in one of the two alternative solutions.

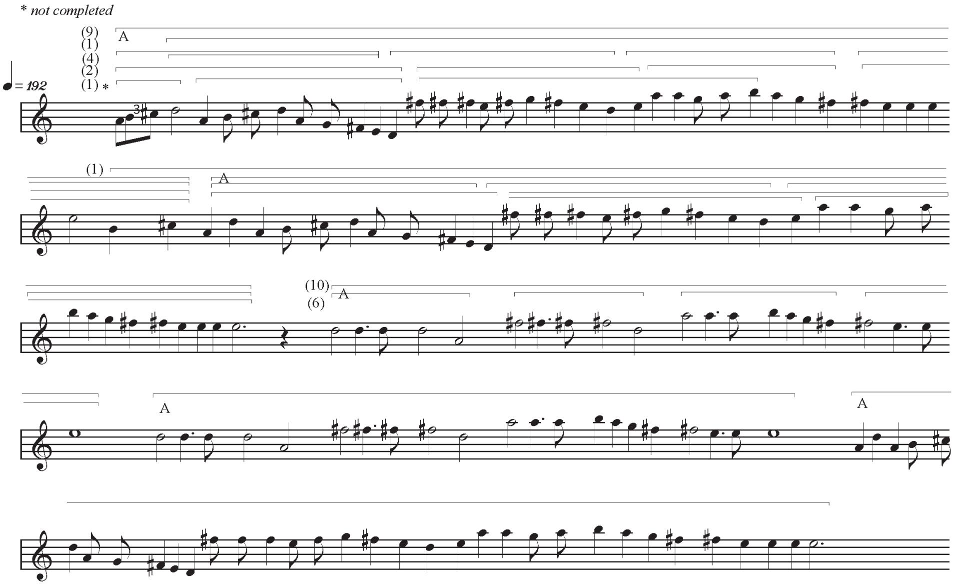

Figure 34.

Summarized result of trial 2 for the 10 participants, displayed in the input score. The numbers within parenthesis denotes the number of test participants who notated the segmentation shown by brackets following the number. Ten test participants out of ten (9 + 1) notated the main repetitions (super-phrase level). Eight out of those were consistently labeled with the letter A (in some cases also with additional numbers). The initial pickup was included in the phrase for all test participants but one. A total of six test participants notated an additional two-bar phrase level, two of which were consistently one quarter note after the beat. Of the six participants who notated two-bar phrases only three labeled these phrases: One participant labeled the phrases A-B-C-D, A-B-C-D in the first melody and two consistently labeled the 2-bar phrases A-A-A-B (with individual numbering) in both melodies. Another test participant started to notate a lower phrase level, but did not finish.

5. Concluding remarks

The method of analysis of melodic similarity within phrase and section analysis of metrical melodies presented in this article has been tested extensively in listener tests and in comparisons with expert analyses, using culturally divergent test melodies (Ahlbäck, 2004). The results of these tests demonstrate that the predictions of segment structure made by the model based on similarity measure are capable of accounting for the greater majority of the segmentations made by people in the tested melodies, on average 80 % of the segments (pairs of boundaries) and 89 % of the individual segment boundaries. (Ahlbäck, 2004)

This supports the notion that it is possible to make such predictions by using gestalt psychologically based measures of structural equivalence, and indicates that these measures may also be used by listeners. This indicates that a measure of melodic similarity based on a one-dimensional representation of melody as a string of notes, without taking into account different levels of melody structure, will be a problematic approach to melodic similarity.

The results of these tests strongly support the notion of melodic similarity as a relative and contextual phenomenon. Perfect similarity is not recognized by participants when it is not congruent with metrical structure (see section 3.2*),* while very general similarity can be sufficient as an indication of structure in a melody like the Bridal March example (see section 4). The analytical model has demonstrated that the dependence of the melodic context for the cognition of structural equivalence can be formally measured, to make predictions of melodic surface structure that are cognitively relevant. The example presented in the last section shows that the level of similarity that is structurally determining varies according to the melodic context. The point is that, when the melody is the measure, the need for cognitive economy, intelligibility and coherence is so great that any type and level of melodic similarity may be structurally determinative.

The current model was originally intended to be a tool for cross-cultural analysis of tonality. Tonality, which can be considered an aspect of deep structure, can be an important factor in melodic segmentation and musical segmentation in general. The most important future development of the model will be the inclusion of tonality aspects in the model.

Acknowledgements

This research was partly supported by the Swedish Research Council, grant nr. 2005-3696

Address for correspondence:
Sven Ahlbäck
Royal College of Music, Stockholm
Box 27711, SE-11591 Stockholm, Sweden
e-mail: sven.ahlback@kmh.se

• References

Ahlbäck, S. (2004). *Melody beyond notes: A study of melody cognition*. PhD thesis, Göteborg University: Publications from the department of Musicology.

Andersson, N. & Andersson, O. (1926). *Svenska låtar*, Stockholm: Gidlunds. *Facsimile 1972.*

Cambouropoulos, E. (1998). "*Towards a general computational theory of musical structure*" Unpublished doctoral dissertation, University of Edinburgh, Scotland.

Cambouropoulos, E. (2001). Melodic cue abstraction, similarity, and category formation: A formal model, *Music Perception*, *18* (347-70).

Cambouropoulos, E. (2005). Musical parallelism and melodic segmentation: A computational approach. (reprint) In W. Brunson & J. Sundberg (ed.), *Music and Music Science*, *Proceedings*, Kungliga Musikhögskolan, Stockholm, Sweden.

Cook, N., (1987). *A guide to musical analysis*, London: J.M.Dent & Sons Ltd.

Conklin, D. & Anagnostopoulou, C. (2001). Representation and discovery of multiple viewpoint patterns. In *Proceedings of the International Computer Music Conference* (ICMC01), Havana, Cuba.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (ed.), *The Psychology of Music* (pp. 473-500). New York: Academic Press.

Deliège, I., (1989). A perceptual approach to contemporary musical forms. *Contemporary Music Review*, vol. 4 (213-30).

Deliège, I., (2001). Introduction, Similarity perception ↔ categorization ↔ cue abstraction. *Music Perception*, *18* (233-43).

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. Orlando: Academic Press.

Fraisse, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (eds.), *Handbook of perception*: Vol. 8 (pp. 203-54). New York: Academic Press.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (ed.), *The psychology of music* (pp. 149-80). New York: Academic Press.

Lamont, A. & Dibben, N., (2001). Motivic structure and the perception of similarity. *Music Perception*, *18* (245-74).

Lartillot, O. (2004). A musical pattern discovery system founded on a modelling of listening strategies. *Computer Music Journal*, *28 (3)*, 53–67.

Lerdahl, F. & Jackendoff, R., (1983). *A generative theory of tonal music*, Cambrige, Mass. The MIT Press.

Ling, J. *Europas musikhistoria: Folkmusiken 1730-1980*. Göteborg 1989.

Middleton, R., (1990). *Studying popular Music*, Philadelphia: Open University Press.

Miller, G. A., (1956). The magic number seven, plus or minus two: some limitations on our capacity for processing information'. *Psychological Review*, *63* (81-97).

Moberg, C-A., (1950). Två kapitel om svensk folkmusik 2. Ro, ro till fiskeskär. *Svensk Tidskrift för Musikforskning 32/1950*, Stockholm: Isaac Marcus Boktryckeri AB.

Nattiez, J. J. (1990). *Music and discourse: Towards a semiology of music*. Princeton: Princeton University Press.

Ruwet, N. (1987). Methods of analysis in musicology. *Music Analysis*, 6:1-2 (11-36). Translation of "Méthodes d'analyse en musicologie". *Revue Belge de Musicologie 20.* (1966) (66-90) Oxford: Basil Blackwell.

Temperley, D. & Bartlette, C. (2002). Parallelism as a factor in metrical structure. *Music Perception*, *20 (2)*, 117-49.

Quantz (1972/1966). "Versuch eine Anweisung die Flöte traversière zu spielen" ("On playing the flute", E.R. Reilly, Translation). New York: Free Press.

Väisänen, A.O. (1954). Kalevalamelodin, En jämförande undersökning In *Svensk Tidskrift för Musikforskning 1954 (39-55)*, Stockholm: Isaac Marcus Boktryckeri AB.

Wertheimer, M. (1923). Laws of organization in perceptual forms. In Ellis, W. (1938). *A source book of Gestalt psychology* (pp. 71-88). London: Routledge & Kegan Paul.

• La similitud melódica como determinante de la estructura melódica

Este estudio presenta una aproximación al análisis de la similitud melódica como determinante de la estructura melódica, que ha sido aplicada con éxito como un sistema computacional de análisis para la predicción de la estructura de segmentación en un ejemplo de melodías monofónicas de diversos orígenes culturales (Ahlbäck 2004).

Aunque el impacto de la similitud en relación con la determinación de la estructura de segmentación de la música es ampliamente reconocido, los métodos basados en la similitud han sido criticados en relación con la dificultad de formalizar criterios y de fijar umbrales para la similitud significativa desde un punto de vista estructural (véase Cook 1987). Se mantiene que a pesar de que la similitud es fundamentalmente un concepto relativo y la categorización de similitud y diferencia está relacionada con un contexto dado, la segmentación basada en la similitud requiere una medición paramétrica del contexto sensitivo de similitud.

El método propuesto de análisis está basado en principios psicológicos tales como la "Gestalt", las limitaciones de la percepción y el proceso cognitivo humano en relación con marcos temporales de atención, la categoría simultánea de aprehensión y los procesos cognitivos de proporciones temporales.

• La similarità melodica quale determinante della struttura melodica

Questo articolo presenta un approccio all'analisi della similarità melodica quale determinante della struttura melodica, applicato con successo come sistema informatico di analisi per la previsione della struttura segmentale in un vasto campione di melodie monodiche di diversa provenienza culturale (Ahlbäck 2004).

Sebbene l'impatto della similarità nella determinazione della struttura segmentale in musica sia generalmente accettato, taluni metodi basati sulla similarità sono stati sottoposti a critiche circa la difficoltà di formalizzare criteri e valori di soglia per una similarità significativa dal punto di vista strutturale (ad es. Cook 1987). In questa sede si sostiene che, poiché la similarità è un concetto fondamentalmente relativo e la categorizzazione di similarità e differenza si riferisce a un contesto dato, la segmentazione basata sulla similarità richiede una misura parametrica della similarità che sia sensibile al contesto.

Il metodo d'analisi proposto si basa su principi psicologici comuni, come i principi della Gestalt, e sulle limitazioni percettive e cognitive dell'uomo circa le cornici temporali d'attenzione, il trattamento simultaneo delle categorie e la cognizione delle proporzioni temporali.

• La similarité mélodique comme déterminant de la structure d'une mélodie

Dans cet article, nous présentons une méthode qui aborde l'analyse de la similarité mélodique comme déterminant de la structure d'une mélodie ; ce système d'analyse

quantitative a été utilisé avec succès pour prédire la structure segmentale d'un échantillon important de mélodies monophoniques d'origine culturelle diverses (Ahlbäck, 2004).

Bien qu'il soit généralement reconnu que la similarité joue un rôle important pour déterminer la structure segmentale d'une musique, on a souvent critiqué les méthodes fondées sur la similarité à cause de la difficulté à formaliser les critères et les valeurs seuils pour une similarité structurelle importante (voir Cook, 1987). Ici, nous montrons que puisque la similarité est un concept relatif fondamental et que la catégorisation de la similarité et de la différence dépend du contexte, la segmentation fondée sur la similarité exige une mesure paramétrique de la similarité sensible au contexte.

La méthode d'analyse proposée ici est fondée sur des principes psychologiques communs tels que ceux de la *gestalt*, des limites perceptuelles et cognitives de l'homme liées à la durée temporelle de l'attention, la gestion simultanée des catégories et la cognition des proportions temporelles.

• Melodische Ähnlichkeit als Determinante der Melodienstruktur

Dieser Aufsatz präsentiert eine Herangehensweise an die Analyse melodischer Ähnlichkeit als Determinante der Melodiestruktur. Diese Herangehensweise wurde erfolgreich als ein rechnerisches Analysesystem für die Vorhersage von Segmentierungsstrukturen in einer großen Auswahl einstimmiger Melodien aus verschiedenen Kulturen eingesetzt (Ahlbäck, 2004). Obwohl der Einfluss von Ähnlichkeit in der Bestimmung von Segmentierungsstrukturen in der Musik allgemein anerkannt ist, wurden ähnlichkeitsbasierte Methoden wegen der Schwierigkeit kritisiert, Kriterien und Schwellenwerte für strukturell signifikante Ähnlichkeit zu formulieren (z.B. Cook, 1987). Da Ähnlichkeit generell ein relatives Konzept ist und Kategorisierungen von Ähnlichkeit und Unterschieden sich auf einen gegebenen Kontext beziehen, wird hier gefordert, dass Segmentierungen, die auf Ähnlichkeiten basieren, ein kontextspezifisches parametrisches Ähnlichkeitsmaß erfordern. Die vorgeschlagene Analysemethode basiert auf bekannten psychologischen Prinzipien wie den Gestaltprinzipien, menschlichen Wahrnehmungs- und Kognitionseinschränkungen bezüglich zeitlicher Aufmerksamkeitsrahmen, simultaner Kategorienverarbeitung und dem Erkennen von zeitlichen Proportionen.