Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Towards Modelling Variation in Music as Foundation for Similarity

Anja Volk,[#1] W. Bas de Haas,[#2] Peter van Kranenburg[*3]
[#]*ICS, Utrecht University, Netherlands*
[*]*Meertens Institute, Amsterdam, Netherlands*
[1,2]{A.Volk,W.B.deHaas}@uu.nl, [3]peter.van.kranenburg@meertens.knaw.nl

ABSTRACT

This paper investigates the concept of variation in music from the perspective of music similarity. Music similarity is a central concept in Music Information Retrieval (MIR), however there exists no comprehensive approach to music similarity yet. As a consequence, MIR faces the challenge on how to relate musical features to the experience of similarity by listeners. Musicologists and studies in music cognition have argued that variation in music leads to the experience of similarity. In this paper we review the concept of variation from three different research strands: MIR, Musicology, and Cognitive Science. We show that all of these disciplines have contributed insights to the study of variation that are important for modelling variation as a foundation for similarity. We introduce research steps that need to be taken to model variation as a base for music similarity estimation within a computational approach.

I. INTRODUCTION

Music similarity is a central concept in Music Information Retrieval (MIR). MIR researches methods that allow users to organise music collections based on similarity, and to retrieve from a collection musical pieces that are similar to a given query. Although the assessment of similarity is considered fundamental for cognitive processes, and a much-investigated concept in Cognitive Science (Goldstone & Son, 2005), there exists no comprehensive approach to similarity in the domain of music. This is a major problem for MIR because the development of content-based search methods for music faces the challenge of relating musical features [1] to listeners' experience of similarity. Both the extraction of musical features from digitized documents that are relevant for human listeners, and the combining of these features into an overall similarity value that approximates the human assessment of music similarity are difficult tasks.

Musicology and experimental studies in music cognition suggest that musical patterns that are repeated and transformed up to a certain extent, but not beyond recognition, are important for the human assessment of music similarity. Cognitive studies have investigated the assessment of similarity by listeners with musical material that contains repeated, yet transformed patterns and concluded that both novices and experts are able to cluster these patterns into groups (Melen & Wachsmann, 2001; Ziv & Eitan, 2007). Repeated yet transformed patterns are recognized because they resemble to an earlier iteration of that pattern. As a result, the corresponding musical passages that contain these patterns (e.g. a thematic section in a sonata, a folk song) are considered more similar to each other than passages that do not contain related patterns (e.g. first vs. second thematic section in a sonata). Musicologists have argued that the repeated, yet transformed patterns help listeners to comprehend the music without having to study the compositional rules (Deliège, 2007; Meyer 2000) because it leads to the experience of similarity within the listening process. Within improvisation, musicians use altered repetitions of musical patterns to enhance the communication of musical ideas. In folk songs, quasi-repeating patterns emerge from the process of oral transmission.

In musicology, the aforementioned phenomena are all subsumed under the concept of *variation*, and musicologists have argued that variation is a universal principle that underlies all music (Deliège, 2007; Nelson 1948). Variation occurs both within and between musical pieces (such as between cover songs). In MIR we seek to model the similarity between musical documents in order to be able to retrieve similar musical pieces in large collections of digitized music. We discern two factors that contribute to the perceived similarity between musical documents: 1) structural similarities in the *data* (digitized musical content, i.e. a digitized score or audio file), and 2) aspects that depend on the specific listener (e.g. time and place of listening, emotional state, personal history, cultural context). Examples of the first factor are sets of variations in classical music (e.g. Bach's Goldberg Variations), occurrences of a theme in a fugue, cover songs in popular music, or repeats of a chorus. Examples of the second factor are a common personal memory attached to different pieces of music, or songs with the same social or ritual function.

In this paper, we address only the first factor – similarities in the structure of the music – which is what we refer to with the term *variation*. Although not often stated explicitly, this is currently the common approach to similarity in MIR because content-based methods set out to extract features from the audio (or digitized notation), and base classification or clustering mainly on those features. This is an understandable point of departure, because the musical data is more easily available than the listening processes and the listening histories of individual listeners. The construction of models of the listening process is very important for MIR (e.g. studies on personalization), but out of the scope of this paper.

Our long-term aim is to develop a computational model of music similarity based on variation. The resulting model will account for those aspects of perceived similarity that are based on structural similarities within the musical data. Hence, if we understand which transformations make up perceptual relevant variations, we can use this information to design musical similarity measures and retrieve pieces that are related by variation. Still, not all variation that is present in

[1] In this article we adopt a typical MIR view on what constitutes a feature. We consider a feature one or more numbers or labels representing a specific characteristic of the music that can be automatically extracted or inferred from any kind of digitally stored musical information.

the music is recognized as such by each listener. This also has both structural and personal aspects. For instance, limited exposure to a certain kind of music during the personal musical history could leave more specific or subtle variation patterns unnoticed, while other variation patterns are easily detectable both for 'novice' and 'expert' listeners, and still other variation patterns might intrinsically be unnoticeable, such as some variational techniques in dodecaphonic music.

For developing computational approaches to similarity based on variation a clear definition of variation is essential. However, the term variation in musicology is not precisely defined; musicologists even have argued that any musical pattern might be derived from any other musical pattern by variation (Meyer, 1973). Hence, there is no clear boundary how to distinguish between variation and non-variation. Therefore, modelling similarity based on variation requires several steps, which we discuss in section V.

In this paper, we first review the literature on variation from three different research strands: MIR, Musicology, and Cognitive Science (sections II-IV). We then introduce in section V research steps that need to be taken to model variation as a base for music similarity estimation.

II. MUSIC SIMILARITY AND VARIATION IN MIR

The massive digitization of music over the last decades has created the opportunity and need for finding new ways of accessing music collections. The success of search engines such as Yahoo and Google has stimulated the development of Music Information Retrieval (MIR) methods that allow users to search in large collections of digitized musical data (Downie et al., 2009). In a long-term vision these retrieval methods should provide access to worldwide music distributions systems. Music similarity is a fundamental principle used in MIR to retrieve music. However, modelling music similarity in MIR is often based on extracting low-level features from music that are difficult to link to human perception (Wang et al., 2005); hence, modelling similarity "remains a huge challenge" (Downie et al., 2009). Modelling the relation between features of the music and listeners' experience of similarity is complex, since musical objects have a much less standardized representation and meaning in comparison to textual objects.

Research in MIR over the last decade has concentrated on extracting low-level features from audio data, which has led to a glass ceiling of performance (Lagrange & Serra, 2010; Downie 2008). The most typical approach to similarity in MIR is to model similarity on the genre level (hence to distinguish between musical documents belonging to different genres, such as Jazz, Classical, Pop). This is based on the assumption that listeners will perceive pieces from the same genre as more similar to each other than across genres. However, both for retrieving and recommending music, music similarity on the genre level provides only very broad categories to the user. Therefore, modelling finer-grained similarities, such as the similarity between cover songs, melodic similarity between folksongs of the same tune family, or the similarity of chord progressions, is of urgent need for MIR. The variation principle has the potential to serve as the ground for finer-grained similarity relations.

Several appealing approaches to symbolic musical similarity have been proposed over the last two decades that go beyond similarity on the genre level. Variation sets from Classical Western music as well as variations of melodies within a tune family have been used as ground-truth data for evaluating similarity measures in symbolic approaches in MIR (e.g. Pickens, 2004; Mardirossian & Chew, 2006; Rizo Valero, 2010; Van Kranenburg et al., 2009). Research into music similarity focuses mainly on melody (e.g. Grachten et al., 2004; Van Kranenburg, 2010), harmony (e.g. Mardirossian & Chew, 2006; De Haas, 2012), and rhythm, (e.g. Toussaint, 2004; Volk, 2008; Volk et al., 2007). All of these methods compare two sequences of symbolic music information, i.e. sequences of notes, chords or inter-onset-intervals, and output a number representing the melodic, harmonic, or rhythmic resemblance, respectively. Although these kinds of similarity measures capture the specific music similarity quite well, they ignore two very prominent aspects of musical similarity. First, most of these methods only examine the sequence of musical events as a whole and ignore repeated information. Hence, comparing a piece with the same piece repeated twice can result in a classification as dissimilar. Second, musical similarity should involve multiple musical dimensions: it is a combination of at least melodic, harmonic, and rhythmic similarity, also factors like tempo, instrumentation and performance have been shown to play an important role (e.g. Foote et al., 2002; Kitahara et al., 2006). Hence, we believe that a general model of musical variation, which captures (approximately) repeated patterns in different musical dimensions, e.g. melody, harmony, rhythm, and which is extensible to other musical dimensions, would be very helpful. We expect that the application of such a model will improve music similarity estimation.

Within computational approaches to music, many pattern matching algorithms have been developed in order to automatically determine motifs and their variants (e.g. Lartillot & Toiviainen, 2007; Cambouropoulos, 2006; Conklin & Anagnostopoulou, 2001; Rolland, 1999; Meredith et al., 2002; Buteau & Mazzola, 2008). An important problem encountered in these algorithms is the combinatory explosion of the results. Hence in general much more patterns are found than those that are considered to be important for a musical piece from the human listener perspective. However, musicological theories on variation of motifs do not provide sufficient insights into what the constraints are for determining salient motivic patterns, which could be used by computational approaches to reach appropriate selectivity. Hence, a better understanding of variation in this respect is of great importance for pattern matching algorithms. Furthermore, pattern matching algorithms have hardly been used in order to determine the overall similarity between two musical pieces, which we will address in section V.

III. VARIATION IN MUSICOLOGY

A. Variation as an omnipresent trait of music

"The principle of variation underlies all music" (Nelson, 1948). Many musicologists have claimed that variation is a universal principle in music. For instance, Von Fischer (1956) called variation a "primary principle" ("Urprinzip") in music, Sisman (2012) states that variations reflect "a technique and process important in nearly all music". Deliège (2007) refers to Arnold Schoenberg, Anton Webern, Simha Arom and Constantinos Brailou, who all consider variation as a central concept in music and discuss its occurrence in classical and folk music. Meyer (1973), who termed the relation between two variation patterns a *conformant relationship,* argues that "there has never been music without conformant relationships". Because of its central role in music, variation and repetition are regarded as characteristic features that distinguish music from language in (Middleton, 1990). From the perspective of developing computational approaches to music similarity in MIR, the occurrence of variation in nearly all music makes the variation principle an attractive candidate for serving as a base for defining similarity measures.

However, despite of the musicological claim that the variation principle is a universal principle in music, there exists no comprehensive theory on variation across the different occurrences of variation: "Many writers have made the variation form the object of their studies, but none of their works includes an extended treatment of the variation as a whole" (Nelson, 1948). For instance, the New Grove entry on *variations* states that variation, as a technique and process, occurs in nearly all music (Sisman, 2012), but describes variation only in Classical Western music. Hence, musicology has claimed that variation in music is universal, yet has not achieved a widely accepted theory of it.

The lack of such a theory is not only a problem for developing models of music similarity in MIR, but for our understanding of music in a broader sense, as argued by Schenker: "Our understanding of musical technique would have advanced much further if only someone had asked: Where, when, and how did music first develop its most striking and distinctive characteristic – repetition?" (Heinrich Schenker, quoted in Kivy, 1993).

B. Variation occurs within and between musical works

Variations occur both within musical pieces, such as in the context of motivic-thematic relationships in Western classical music, and between musical pieces. Examples of the latter are variations between folk songs belonging to a tune family; variation between the different pieces belonging to a variation set in Western classical music; variation patterns in genres of popular music, such as generic chord sequences. Deliège (2007) distinguishes therefore between internal and external similarity relations. In the context of repetition, Margulis (in print) argues that it is still an open question whether "the same cognitive mechanisms and processes underlie responses to these two types".

Similarities based on variation within a piece play an important role in the unfolding of the musical structure over time during the listening process (Zbikowski, 2002). Since these similarities enable listeners to recognize important building blocks of the music and are important for musical memory, they are also important for the experience of similarities *between* musical pieces: what characteristics of a musical piece stay in the mind of listeners such that they will use them to assess the similarity of the piece to another musical piece? These questions are important for developing music similarity measures in Music Information Retrieval that come close to how human listeners perceive similarity.

C. Variation: The lack of a definition

Musicologists have admitted that a precise definition of variation has not been achieved. "All of us know what we mean, or think we know what we mean, when we say that one musical passage is derived from another; yet we may find it hard to come up with a precise definition that will fit all those cases that we consider examples of derivation, and only those cases" (Cone, 1987).

Variation, understood as a relation or transformation between musical patterns, is difficult to distinguish from non-variation: "… 'equivalence' and 'difference' in music are difficult to distinguish precisely; they tend to run into each other, through techniques of transformation and variation, and it can even be argued that, with sufficient analytic sleight of hand, *all* differences can be reduced to being transformations of a single generative source" (Middleton, 1990). Also Meyer refers to the challenge that in theory, one could transform a pattern into any other pattern: "I fully agree with Tovey that 'Nothing is easier than to derive any musical idea whatever from any other musical idea'" (Meyer, 1973).

However, for modelling music similarity based on variation, only those transformations that are perceptually relevant for the listener, are of interest. For instance, though retrogrades (time inversions) of a musical theme can be considered as an instance of variation from a musicological point of view, we would exclude them as an instance of variation for modelling musical similarity perceived by a human listener.

D. Variation studied in specific styles

1) Variation in Classical Western Music

Studies on variation in Classical music focus either on the study of motivic relationships in general (e.g. Meyer, 1973; Zbikowski, 2002; Réti, 1951; Schoenberg, 1967), or specifically on the study of Variation sets (Nelson, 1948; von Fischer, 1955; Sisman, 2012).

Studies on motivic relationships discuss what musical features are important for constituting the similarity relationship between motifs. Schoenberg (1995) provides as a definition of musical motif that it needs to be recognizable throughout the piece despite change and variation. He considers rhythm as most important feature of motifs. With this characterization of motifs he differs from Schenker (1935) and Réti (1951), who consider specific intervallic relationships as most important features of motifs, while rhythm and contour provide secondary features, as Zbikowski (2002) has pointed out. Meyer (1973) considers pitch, duration, and harmony as "primary pattern-forming parameters" of motifs, while dynamics, register, and timbre provide secondary parameters. Meyer (1973, p. 49) even provides a formula on the strength of the perceived relationship between motivic patterns, which considers among others the regularity of the pattern, the temporal distance between the patterns and the similarity of the patterning. Hence, the role of the different musical features in establishing the variation relation between musical patterns is discussed controversially in musicology; a similar discussion on the role of features contributing to similarity based on variation has been carried out in the context of cognitive studies on variation, as we will show in section IV. For modelling similarity based on variation, different weightings of the involved musical features for deriving an overall similarity value therefore need to be evaluated, since the role of single musical features seems not to be fixed.

Moreover, both (Meyer, 1973) and (Réti, 1951) discuss the possibility to distinguish different degrees of similarity. While the study of the variation relations between motifs within a musical piece considers a specific form of variation in which similarities are recognized, Meyer (1973) argues that similarities on a more general level, namely on the style level, go unnoticed: "Once the style has been learned – through the experience of listening, not necessarily through explicit instruction – such similarities are, as it were, taken for granted" (Meyer, 1973, p. 46). Réti (1951) distinguishes different degrees of similarity between motifs along four discrete steps, namely *imitation*, *variation*, *transformation*, and *indirect affinity*. Apart from the fact that Réti uses the concepts of variation and transformation in a rather different manner than the musicological literature we have referred to so far, the characterizations of these different degrees of similarity are very vague (see Réti, 1951, p. 239-240). For instance, imitation described as "literal repetition of shapes, either directly or by inversion, reversion ..." or variation as "changing of shapes in a slight, well traceable manner" does hardly assist in getting a precise notion of what Réti is referring to. While such a distinction between different degrees of similarity in motivic relationships would be very valuable for computational approaches to similarity based on variation, the descriptions given are difficult to formalize.

In the musicological study of so-called variation sets (pieces that are connected by theme and variation relations, e.g. Bach's Goldberg Variations), different types of variation are discussed, resulting in categories of variation sets. The discussion of these categories is interesting from the point of view of evaluating computational models of variation on digitized musical documents. For instance, if a certain piece is known to exhibit constant-harmony variation according to the musicological arguments, it can provide the base for evaluating a model on harmonic similarity. Therefore we give in the following section a brief overview on the discussion of variation sets.

Variation sets emerged in the 16th century, when "themes originating in dance and song" were used as base for variation (Sisman, 2012). According to Sisman, the first variation sets captured two forms of improvisation, namely "the variations in repeated strains of dance music, and varied settings in successive stanzas of a song whose melody can be savoured as a cantus firmus or a springboard to figuration". The different classifications of variation sets that have been proposed are based on determining what parts of the composition remain stable across variations (e.g. constant-harmony variations); however since some of the variation sets belong to a certain musical period, the historic background contributes in some cases also to the classification (e.g. baroque basso ostinato variation vs. 19th-century basso ostinato variation). In the following we discuss three proposed classification systems by von Fischer (1955), Nelson (1948) and Sisman (2012).

Kurt von Fischer (1955) distinguishes the following types of variation:

1. Cantus firmus variations
2. Ostinato variations
3. Constant-harmony variations
4. Melodic-outline variations with constant harmony
5. Fantasy variation
6. Serial variations

For the first four types he lists what elements remain stable and what elements are varied. In the *Cantus firmus variations,* the cantus firmus remains constant, while the other voices, rhythm, harmony and the general form can change. The *Ostinato variations* differ from the Cantus firmus variation in that they do not preserve the entire melody of the cantus firmus, but only some parts of it. In the *Constant-harmony variations,* harmony and often the general form remain constant, while melody and voice leading can change. In the *Melodic-outline variations with constant harmony* important parts of the melody ("melodische Haupttöne"), harmony and form remain constant, while rhythm and tempo can change. For the last two types his distinction between constant and variable elements is less straightforward. For the *Fantasy variation* he claims that some parts of the theme (motifs, melodic fragments) remain constant, while all elements of the themes can be variable. *Serial variation* refers to the variation of 12 tone rows.

The classification in (Nelson, 1948) distinguishes seven distinct kinds of variation sets, each type is bound to a certain time period:

1. Renaissance and baroque variations on secular songs, dances and arias.
2. Renaissance and baroque variations on plain songs and chorales
3. The baroque basso ostinato variation.
4. Ornamental variation of the 18th and 19th centuries.
5. The19th-century character variation
6. The 19th -century basso ostinato variation
7. Free variation of the late 19th and early 20h centuries.

The difference between types 1 and 2 is not only the use of secular vs. liturgical themes, but also that type 2 exhibits a "more serious and complicated style" than type 1. Type 3 variations differ from types 1 and 2 "by virtue of their continuous construction" (i.e. "uninterrupted flow of movement"). Type 4 "exhibits greater simplicity than its prototype in the renaissance and baroque periods", hence "the connection between theme and variations is singularly transparent". Type 5 "contrasts strongly with earlier types in general. … previous variations tend to preserve the expression of the theme throughout a series, the separate members of the character variation frequently alter the expression, or 'character', of the theme profoundly". Furthermore, Nelson argues that in type 5 "we find here, for the first time, an emphasis upon the development of motifs from the theme". In variation type 6 "its component members often depart widely from the expression of the theme; they also present the theme more frequently in upper voices". Type 7 marks "a significant departure from all earlier species in that the bond between variations and theme is now frequently a theme motive rather than the theme in its entirety. This means that the structural and harmonic pattern of the theme is often discarded in favor of a free development". Hence, Nelson (1948) classification reads like a historic development of variation sets through the centuries.

Sisman's (2012) classification highly conforms to Fischer's classification, however she lists two additional categories, namely Formal-outline variations and Characteristic variations. She characterizes Fantasy variations as "departing from any clear structural similarity" with the theme while the variations are especially grounded in melodic motifs. Hence, this type corresponds to Fischer's type 6 and Nelsons's type 7. In the formal-outline variations, "aspects of the theme's form and phrase structure are the only features to remain constant". Furthermore, Sisman (2012) emphasizes that Characteristic variation is "not the same as character variation", but that "individual numbers take on the character of different dance pieces, national styles or programmatic associations", while formally this can involve variations according to other variation types (such as Constant-harmony variations, Melodic-outline variations etc.).

In general, the different classification systems of variation sets proposed in Musicology do not provide clear-cut definitions of the involved types. From the perspective of modelling variation, one might argue that variation sets provide rather specific types of variation in Western classical music, which are possibly difficult to generalize. Nevertheless, the discussion of the types (alongside with concrete musical examples) can assist to find evaluating methods for computational models of variation, since annotated data sets that indicate where in the musical documents variation patterns occur, are rare (an exception is Volk & Van Kranenburg, to appear). Variation sets have been used in MIR to test models of similarity (e.g. Pickens, 2004; Rizo Valero, 2010) or to test the use of Schenkerian reductions for recognizing variations (Marsden, 2010). Hence, for computational modelling of variation, using these variation sets allows for comparison to other computational approaches in MIR. Apart from the classification systems for variation sets introduced in Musicology, single composers have been discussed regarding their specific contribution to variation sets in music history (e.g. Bach, Haydn, Brahms), which can further assist in establishing ground-truth data on variation for computational models.

2. Variation in Folk Music and Oral Traditions

Constantinos Brailou has argued (according to Deliège, 2007), that "the study of … variations … is the most important question in the field of musical folklore". Van der Merwe (1989) lists among the features that distinguish European *folk* from *art* music, repetition and variation.

An important characteristic of folk music is that this music has evolved through the process of oral transmission (Karpeles, 1968). During this process, variation to the musical material is introduced. Variation can be seen as a consequence of the limitations of human musical memory, such that while reproducing a heard musical piece, variation is introduced. However, variations can also be considered as the product of creative processes: "variation which springs from the creative impulse of the individual or the group" as argued by Karpeles, (1968), hence variants can be distinguished to be produced either by a person or a larger musical context, as (Middleton, 1990, p. 136) argues: "an individual (an idiolect) and those associated with a context".

For capturing the variations introduced through oral transmission, both theoretical studies and computational approaches have been developed. For instance, Wiora (1941) discussed the different types of changes that occur to a given folk melody in the course of oral transmission. Bayard (1950) introduced the concept of tune family, denoting a group of melodies that are supposed to have a common 'ancestor' in the line of oral transmission. Bronson (1950) discusses characteristic features of tune families in a collection of British-American folk songs and has been one of the first to propose a computational approach for organising a collection of folksongs. He used a punch-card system to sort melodies according to features such as the final cadence and the mid cadence. Cowdery (1984) argues that melodies from the same tune family are composed from the same 'pool of motifs'. Early efforts in musicology to establish classification systems for folk songs for ordering folk melodies according to similarity, such as by Krohn, Bartok, and Kodaly (summarized in Suchoff, 1981), used characteristics such as number of phrases, number of syllables in each phrase, pitches of the cadence tones of the phrases etc. These methods have been further developed in recent computational approaches to similarity between folksongs (e.g. Sagrillo, 1999; Van Kranenburg, 2010).

The modelling of variation in folk song melodies is valuable for MIR because the resulting models are supposed to capture kinds of melodic variation that are tightly bound to the characteristics of the human processes of remembering and reproducing melodies.

3. Variation in Popular Music

Middleton (1990) argues that for popular music, "a high level of repetition may be a specific mark" since it enables "an inclusive rather than exclusive audience". Referring to variation in folk music, where a common pool of musical material is used in oral tradition, he argues that the typical composition methods of popular music production might lead to "similar structural patterns" in popular music. Indeed, the empirical study in (Frieler & Riedemann, 2011) on melodic improvisation based upon a standard harmonic progression seems to back up this claim to a certain extent, since in some cases participants would improvise unconsciously a known popular hit.

However, in comparison to studies in folk music, to our knowledge not many theoretical studies exist on the description of the specific variation patterns occurring in popular music. Van der Merwe (1989) discusses how early beginnings of popular music emerged from folk music, along with typical variation patterns (such as standard blues chord progressions). Burns (1987) argues that the variation principle in popular music is used to produce *hooks* by generating novel, yet familiar patterns: "this article will be concerned with the definition and classification of the structural elements of music as specifically exemplified in pop records, and with the analysis of how songwriters, performers and record producers manipulate these structural elements through use of repetition, variation and modulation to produce hooks." For instance, harmonic hooks are variants of chord patterns, containing a radical change but preserving basic chord patterns that became genre conventions. He gives a (rather high-level) overview on the different types of hooks (e.g. rhythmic, melodic, harmonic) in popular music.

In the context of MIR, popular music has been studied extensively, while the high level of repetition within pieces of popular music is used, for instance, to automatically detect segments based on within-piece similarities. In comparison to this, the modelling of variation between pieces has gained much less attention so far. The study of prototypical harmonic patterns, such as in blues (e.g. De Haas, 2012; Steedman, 1996) might be a promising starting point for modelling variation in popular music, since these have been discussed in the musicological literature, providing ground-truth information for the evaluation of computational models.

IV. VARIATION IN COGNITIVE SCIENCE

A. Variation studied in Music Cognition

Margulis (2012) argues that in the context of music cognition, variation has not gained sufficient interest yet: "Although music's repetitiveness has been a perennial topic of theoretical and philosophical interest, we know surprisingly little about the psychological processes underlying it." Yet, variations have been investigated in a number of listening experiments in the context of categorization in music, music similarity assessments and the study of musical transformation. Hence, patterns in the musical piece that are considered to be variations of each other are expected to be classified into the same group and to receive higher similarity ratings than patterns that are not considered as variation patterns. In this section we provide a detailed overview on the research questions, methods and results.

Welker (1982) tested whether listeners are able to abstract themes from melodic variations. A set of transformations of a melody was generated by systematic application of five transformational rules. After listening to these transformations (without presenting the original melody), participants had to draw the melodic contour best describing the central tendency of the presented transformations. Participants were able to abstract the original melody while listening to the set of transformations, while no difference between novices and experts has been found.

Melen and Wachsmann (2001) studied the categorization of musical motifs by infants from 6 to 10 months for a piece by Franz Schubert. The infants were able to form categories of musical motifs. However, little information about the characteristics that served as the basis for the categorization has been concluded. One hypothesis is that infants are able to categorize melodies as long as the melodic contour remains unchanged, hence they can abstract from changes in intervals and absolute pitch height. Koniari, Predazzer, and Melen (2001) investigated categorization processes in music perception by 10- to 11 year- old-children. Using two pieces by Diabelli and Schubert, children categorized motifs into two groups stemming from different thematic sections and evaluated the similarity of motifs on a scale between 1-5. The authors concluded, that perceived similarity is mainly influenced by musical surface features, such as melodic line, register or dynamics and also from elements related to the underlying harmonic structure, such as the harmonic cadences.

Lamont and Dibben (2001) conclude from a study on the perception of motivic variation in Beethoven and Schoenberg that perceived similarity rather depends on surface features of the music, not on motivic relations. Here, surface features are defined according to music theoretic notions of features of motifs (see Meyer, 1973; Réti, 1951), such as changes of texture, orchestration, register and pace. They are opposed to deep features of motifs, such as the derivation and fragmentation of the original pitch and rhythm information.

Ziv and Eitan (2007) investigate whether listeners' thematic and motivic categorization may differ considerably from the categorizations suggested by music theorists and musicologists. While music theorists consider primarily pitch and pitch-class relationships and secondarily rhythmic-metric structure as defining the motivic and thematic identity, hence as the musical features that remain unchanged across variations, listeners might attend more to surface features such as register, instrumentation, dynamics, tempo, textural density or melodic contour in order to categorize motifs. Using the same stimuli as in Lamont and Dibben (2001), Ziv and Eitan (2007) asked participants to categorize extracts of the music as belonging to one of the two principal themes. Furthermore, listeners rated on a scale ranging from 1 to 11 the degree to which each of the excerpts belongs to one of the thematic sections. Using this approach, they tested the hypothesis that categorization will emphasize deeper-level structural features, opposed to the results concluded in Lamont and Dibben (2001). Ziv and Eitan (2007) show that listeners' categorization agrees fairly well with music theoretic notions of variation in Beethoven's piece, but diverge from those of Schoenberg's piece. Furthermore, Ziv and Eitan (2007) compiled a list of characteristic musical features for the two themes in Beethoven's piece (e.g. phrase structure, initial and terminal melodic intervals, conspicuous voice-leading, melodic schemas, rhythmic figures, rhythmic density) as well as for Schoenberg's piece (e.g. aspects of row-structure). Using Spearman rank order correlations, they compared diverse combinations of rankings of these features with rankings derived from listeners' mean scores of ratings as to how well an excerpt belongs to a thematic section. As the results show, among the features that best correspond to listeners' categorization are aspects of texture, rhythm, melodic contour and dynamics. The authors conclude that surface features seem to provide the thematic categorization. Another important finding of this study is that musical experts and novices did not differ in their categorization.

Deliège (2001) studied the formation of thematic categories in adult human listeners using a violin sonata by J.S. Bach. Subjects listened to the first part of the piece; then they listened to items either from the part already heard (*heard items*), from the second part not heard (*unheard items*) and items that were slightly changed in pitch and rhythm (*modified items*) resulting in stylistic incongruities. Afterwards they had to decide whether they had heard the item before. Deliège thus tested the hypothesis that unheard items would be erroneously be judged as having been heard before due to their similarity to items present in the part already heard, while the stylistic incongruities in the modified items should be easily detected as items that do not belong to the piece. The results confirm her hypothesis, while it is interesting to note that pitch changes in the modified versions alone led to a decreased rate of items correctly identified as not having been heard in comparison to pitch and rhythm changes.

Volk & Van Kranenburg (in print) have investigated in an annotation study the relation between musical features, perceived similarity and human categorization by musical experts. In the study an annotation data set for 360 folk song melodies in 26 tune families was created. The analysis of the annotation data set has revealed that the importance of single musical features for assessing similarity varies both between and within tune families. In general, the recurrence of short characteristic motifs is most relevant for the perception of similarity between songs belonging to the same tune family. Global melodic features often used for the description of melodies (such as melodic contour) play a less important role.

In summary, cognitive studies have shown that both novices and experts are able to perceive similarities based on variation. This confirms that variation is not only an important compositional technique (see Nelson, 1948: "the student who wishes a comprehensive training in composition"), but backs up the claim that variation allows the listeners to comprehend music. These findings provide also strong arguments for using variation as a base for similarity in music.

However, most research has been carried out with Western classical music. Moreover, it remains less clear what musical features allow listeners to perceive similarity based on variation. The distinction between "surface" and "deep" features, which is a major point of discussion in cognitive studies, corresponds to the distinction between "surface" and "deep" similarity discussed in Cognitive Science (Vosniadou & Ortony, 1989). While such a distinction would be useful to predict, for instance, different similarity assessments between musical experts and novices, it is still an open debate how to exactly determine surface and deep features in music. For modelling similarity based on variation, the empirical studies provide important insights that variation is not only accessible to expert listeners. However, time-intensive listening studies are restricted to only a small number of musical pieces that have been tested, while a computational approach will allow modelling variation within a data-rich approach.

B. Similarity in Cognitive Science

While studies in music cognition have investigated similarity in music based on variation, general studies on similarity in Cognitive Science have hardly considered the domain of music. Yet, similarity is considered fundamental in Cognitive Science, since it plays a crucial role for mental processes such as learning, problem solving, memory, prediction and categorization (Goldstone & Son, 2005). Cognitive Science strives to find general principles of similarity underlying many domains and has developed formal models for similarity, such as geometric models (e.g. multidimensional scaling models), featural models (e.g. Tversky's Contrast Model based on weighting common and distinctive features), alignment-based models (based on determining how features align with each other) and transformational models (similarity defined as transformational distance). Cognitive Science regards similarity as a very flexible concept: the similarity between two objects perceived by humans is not a stable entity. Therefore, research in Cognitive Science has concluded that there is not "one kind of similarity", but many "kinds of similarity" (Medin et al., 1993; Smith, 1989), such as surface and deep similarity (e.g. Vosniadou & Ortony, 1989), global and dimensional similarity, holistic and analytic similarity, attributional and relational similarity. These distinctions mainly draw on a classification of the involved features. However, it is still an open question in how far the general findings on similarity concluded in Cognitive Science apply to the domain of music. The formal models for combining different features into an overall similarity value, such as alignment-based and transformational models developed in Cognitive Science for other domains than music, need to be evaluated regarding their usefulness in music. This will contribute to evaluating similarity models across different domains. Alignment-based models have been developed in Cognitive Science as a response to earlier similarity models which were based on matching features, taking into account that the comparison of objects is based on determining how elements correspond to each other (Goldstone & Son, 2005). Transformational models define similarity in terms of transformational distance (Goldstone & Son, 2005). Since variation between two musical patterns can be described in terms of a musical transformation, transformational models seem to provide a natural approach to music. Moreover, since time is an important feature of music, alignment-based models that align correspondences between musical pieces in time are promising candidates for modelling similarity in music based on variation.

V. MODELLING VARIATION

In this section we provide an outline over modelling similarity based on variation with computational approaches, which is at the core of the recently started MUSIVA-project[2] (Modelling Musical Similarity over Time through the Variation principle) at Utrecht University.

For developing computational approaches to similarity based on variation, we first need a formal definition of the concept of variation. None of the three disciplines, MIR, Musicology, and Cognitive Science, have yet come up with a definition of what exactly constitutes variation. Since in our context of MIR, we aim to model music similarity as experienced by listeners, we view variation from the perspective of the listener and not from the perspective of the composer or music creator. Hence, a formal definition needs to incorporate some notion of perceptually relevant (approximate) repetition. A variation can be viewed as a repeated pattern that has been transformed. Hence, a definition of variation should place boundaries on what kind of transformations are recognised as a variation. Also the amount of transformation is important and needs to be accounted for.

Modelling similarity between musical pieces (or parts of pieces) based on variation requires two general steps. First, computational approaches to detect variation patterns that are

[2] http://www.cs.uu.nl/research/projects/vidi-volk/

perceptually relevant need to be developed (modelling variation). Second, methods to derive an overall similarity measure between the two pieces (or parts of pieces) based on the variation patterns detected in the first step need to be developed (modelling similarity).

For modelling variation in the first step, we take an important finding from Margulis (2012) on variation perception into account: the phrase structure of pieces is crucial for the perception of variation patterns by listeners. Literal repetitions of a pattern are not detected if they are not conform with the phrase structure. Hence, the perceptual relevance of a specific variation does not only depend on the kind of transformation involved, but also on the musical context within which the variation patterns occur. Therefore, we will develop segmentation methods in order to partition pieces into perceptually meaningful units, based both on cognitively salient boundaries and repetition borders, such that the global structure of a musical piece is described as a patchwork of local contexts. We will then take the phrase positions of variation patterns into account in order to determine whether they might be perceived as variations of each other or not. This will contribute to solving the issue of combinatorial explosion in pattern matching algorithms, as described in section II. Taking phrase positions into account as an indicator of the perceived relevance of variation patterns can contribute to constraining the number of patterns.

According to the musicological and music cognition literature, many different musical features contribute to the perceived similarity of variation patterns. We have developed computational approaches to melody (Van Kranenburg, 2010), harmony, (De Haas, 2012) and rhythm (Volk, 2008) and will investigate their potential to model aspects of variation. Moreover, we plan to conduct listening tests on variation with respect to the position of variation patterns within phrases.

For modelling the overall similarity between musical pieces (or parts of pieces) based on variation in the second step, we face the challenge that music cognition studies presented in listening tests in most cases short excerpts from the music containing variation patterns, hence research on similarity in cognitive studies has concentrated on short local contexts, not on overall similarity. However, we have a unique annotated corpus of folk songs (Volk & Van Kranenburg, in print), for which we have information both on overall human similarity assignment (in terms of the membership of a melody to a tune family), and on features contributing to the overall similarity (rhythm, melodic contour, motifs). Moreover, for 360 melodies, domain experts annotated occurrences of what they considered important melodic patterns for the overall similarity assessment. Hence, this annotated corpus provides valuable ground truth on deriving an overall melodic similarity value based on variation patterns and we will use it as a first test set. We have used the annotated corpus in a first attempt to determine the overall similarity between folk songs based on sequences of melodic patterns in (Van Kranenburg et al., 2012). Since the automatic detection of perceptual meaningful repeating patterns is yet unsolved, in (Van Kranenburg et al., 2012) we explicitly defined a set of melodic patterns for this corpus, based on the annotations of melodic patterns by the experts. We defined 15 abstract melodic patterns, or motif classes, such as a broken chord, a big leap, a series of relatively long notes, etc. It appears that solely based on the occurrences of these motif classes, melodies from the same tune family can be retrieved from a collection of c. 5000 melodies. Since concrete occurrences of the motif classes show variation, this result confirms the importance of the variation principle for establishing melodic similarity.

Another distinctive high-level musical feature that is subject to variation is harmony, for instance in Bebob by the addition of secondary dominants to popular songs. We propose to do a repeated pattern analysis of a chord sequence based on adapted suffix tree algorithms from computational biology (Gusfield, 1997, Chapters 6-9). The suffix tree data structure allows for the creation of fast algorithms that can detect repetitions in a sequence of symbols. Additionally, De Haas (2012) proposes several approaches to harmonic similarity between two sequences of chords labels that take into account the function of the chord within its local context, e.g. the surrounding chords, and within its global context, e.g. the key of the piece. We expect that combining these two approaches will contribute to a general model of variation. After all, repetitions can be identified by a suffix tree based repetition search, and the transformations between these repetitions can be quantified by context-aware harmonic similarity measures, both within and between pieces. Hence, we expect that such a model will enhance general harmonic similarity measures because possible asynchronies between the number of repetitions in two matched chord sequences will not affect the global similarity estimation.

Furthermore, we consider the time sequence of patterns as important for modelling an overall similarity value based on variation patterns: the musical structure of a piece unfolds in time during the listening process which influences the experience of music similarity. For addressing the unfolding in time we will further develop an approach to the unfolding of rhythm over time based on the procedural approach to rhythm in Volk (2005) and our successful application of rhythmic similarity using Inner Metric Analysis (Mazzola, 2002; Volk, 2008) both on salient local cells (Chew et al., 2005) and on the global structure of a piece (Volk et a., 2007). In a next step, the unfolding over time of other relevant features will be modelled.

For combining different features into an overall similarity value, both alignment-based models and transformational models have been discussed in Cognitive Science as having several advantages over earlier formal models of similarity, such as Tversky's featural model (Goldstone & Son, 2005). We will evaluate these models for the domain of music.

A specific challenge for the computational modeling is the question on what digitized musical data we can use to evaluate our computational approaches to variation. As a first step we will take variation sets into account that have been discussed in Musicology as a very specific form of variation, since these discussions provide some information on what type of variations (melodic, harmonic) occurs in the pieces. There exist some small data bases on variation sets that have been used in MIR for evaluating similarity models (e.g. Mardirossian & Chew, 2006; Pickens, 2004), which provide a starting point for the evaluation of computational models of variation.

VI. CONCLUDING REMARKS

In order to realize the two main steps as described in the previous section for modelling similarity based on variation, namely 1) to develop computational approaches to detect variation patterns that are perceptually relevant and 2) to derive an overall similarity value between pieces based on the variation patterns detected, we propose the following research agenda:

- Development of segmentation methods: modelling of the global structure of a musical piece as a patchwork of local contexts.
- Development of computational models for harmony, rhythm, and melody that detect variation patterns.
- Modelling of the interaction between local variation patterns and their local contexts as defined by segmentation methods in order to determine the salience of the local variation patterns.
- Modelling of overall similarity based on variation patterns by integrating segments into their global context within the piece. Similarity measures based on the interaction of the local and global contexts will be derived. For modelling the contribution of different features to the integral experience of similarity, similarity models such as alignment and transformational models will be tested.

In general, a formal cognitively motivated model of variation is necessary to improve the current standard in music similarity. Such a model of variation needs to account for approximate repetitions in sequences of musical events that are recognised in various musical dimensions, e.g. melody, harmony, rhythm, etc. Hence, such a model combines local information from (approximately) repeated segments with global information about the structure (or form) of a piece to obtain a final similarity assessment.

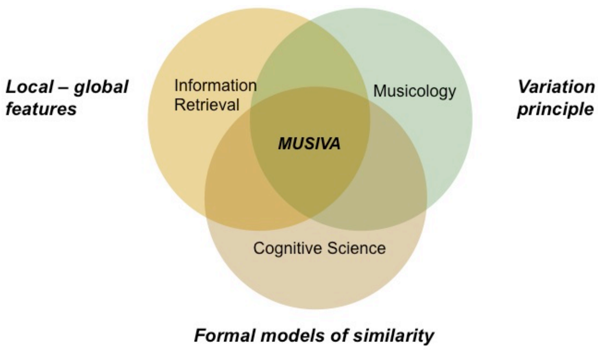

Figure 1: Disciplines involved in modelling similarity based on variation in *MUSIVA*

Moreover, modelling similarity based on the variation principle requires input from Musicology, Cognitive Science and Music Information Retrieval (see Figure 1), but will also deliver important contributions to the three disciplines involved:

Information Retrieval: Modelling similarity based on a model of variation addresses high-level processing in establishing similarity in music and hence will deliver an essential requirement for building cognitively plausible search algorithms in Music Information Retrieval. Modelling the different levels in the musical organization with respect to *local and global information* and the interaction of features of music addressing *different time spans* will allow the development of a similarity concept that considers adequately the complexity of music.

Musicology: The current musicological discourse on the variation principle is mainly based on small numbers of selected musical examples and has not led to a general concept of variation. The computational modelling allows to formalize the concept and to explicitly test it on a large collection of music.

Cognitive Science: While research in music cognition has strongly focussed on the *experienced listener* of Western classical music (Peretz, 2006), research on music similarity contributes an excellent topic on basic music skills. Similarity is used as a default method to reason about a domain, even if we do not have specific knowledge about it (Goldstone & Son, 2005). Thus, understanding music similarity will demonstrate that accessing music is not reserved to the highly trained specialist. Moreover, a model of music similarity will contribute an underrepresented domain to similarity research in Cognitive Science and hence contribute new aspects to the search for general principles of similarity across different domains.

ACKNOWLEDGMENTS

We thank Frans Wiering and Marcelo Rodriguez-Lopez for providing detailed comments and discussions of an earlier draft of this paper. Anja Volk and W. Bas de Haas are supported by the Netherlands Organization for Scientific Research, NWO-VIDI grant 276-35-001. Peter van Kranenburg is supported by the Royal Netherlands Academy of Arts and Sciences, through the Computational Humanities funding program.

REFERENCES

Bayard, S. (1950). Prolegomena to a study of the principal melodic families of british-american folk song. *Journal of American Folklore*, 63(247), 1–44.

Bronson, B. H. (1949). Mechanical help in the study of folk song. *The Journal of American Folklore,* 62(244), 81–86.

Bronson, B. H. (1950), Some observations about melodic variation in british-american folk tunes. *Journal of the American Musicological Society*, 3, 120–134.

Burns, G. (1987), A typology of 'hooks' in popular music, *Popular music* 6(1), 1-20.

Buteau, C. and Mazzola, G. (2008), Motivic Analysis According To Rudolph Réti: Formalization By A Topological Model, *Journal of Mathematics and Music,* 2(3), 117-134.

Cambouropoulos, E. (2006), Musical parallelism and melodic segmentation: A computational approach, *Music Perception*, 23(3), 249-267.

Chew, E., Volk, A. and Lee, C.Y., (2005), Dance music classification csing Inner Metric Analysis, in: *The Next Wave in Computing, Optimization and Decision Technologies*, Springer OR/CS Interfaces Series, 29, 355-370.

Cone, E. T. (1987), On derivation: syntax and rhetoric, *Music Analysis*, Vol. 6, 237-255.

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

Cowdery, J. (1984), A fresh look at the concept of tune family. Ethnomusicology, 28(3): 495–504.

De Haas, W.B. (2012), *Music information retrieval based on tonal harmony*. PhD thesis. Utrecht University.

Deliège, I. (2001), Prototype effects in music listening: An empirical approach to the notion of imprint, *Music Perception*, 18 (3), 371-407.

Deliège, I. (2007), Similarity relations in listening to music: How do they come into play? *Musicae Scientiae*, Discussion Forum 4A, 9-37.

Downie. J.S. (2008), The music information retrieval evaluation exchange (2005–2007): a window into music information retrieval research. *Acoustical Science and Technology*, 29(4), 247–255.

Downie, J.S., Byrd, D., Crawford, T. (2009), Ten Years of ISMIR: Reflections on challenges and opportunities, *Proceedings of the International Society for Music Information Retrieval Conference,* 13-18.

von Fischer, K. (1955), *Die Variation,* Das Musikwerk 11, Köln, Arno Volk-Verlag.

Foote, J., Cooper, M., Naam, U. (2002), Audio retrieval by rhythmic similarity, *Proceedings of the International Society for Music Information Retrieval Conference*.

Frieler, K. & Riedemann, F. (2011), Is independent (re)creation likely to happen in pop music, Musicae Scientiae, 15, 17-28.

Goldstone, R.L. & Son, J. (2005), Similarity, *The Cambridge Handbook of Thinking and Reasoning*, 13-36, Cambridge U Press.

Gusfield, D. (1997), *Algorithms on strings, trees, and sequences: computer science and computational biology,* Cambridge University Press.

Grachten, M., Arcos, J.L., Lopez de Mantaras, R. (2004), Melodic similarity: Looking for a good abstraction level, *Proceedings of the International Society for Music Information Retrieval Conference*.

Karpeles, M. (1968), The distinction between folk and popular music, *Journal of the International Folk Music Council*, 20, 9-12.

Kitahara, T., Goto, M., Komatani, K., Ogata, T., Hiroshi, O., (2006), Musical instrument recognizer "Instrogram" and its application to music retrieval based on instrumentation similarity, 8th IEEE International Symposium on Multimedia ISM06, pp. 265-274.

Kivy, P. (1993), *The fine art of repetition: Essays in the philosophy of music*. Cambridge: Cambridge University Press.

Koniari, D. Predazzer, S. & Melen, M. (2001), Categorization and schematization processes used in music perception by 10- to 11- year old children, *Music Perception*, 18, 297–324.

Lagrange, M. & Serra, J. (2010), Unsupervised accuracy improvement for cover song detection using spectral connectivity network, *Proceedings of the International Society for Music Information Retrieval Conference 2010.*

Lamont, A. & Dibben, N. (2001), Motivic structure and the perception of similarity, *Music Perception,* 18, 245–274.

Lartillot, O. and Toiviainen, P. (2007), Motivic matching strategies for automated pattern extraction, *Musicae Scientiae*, Discussion Forum 4A, 218-314.

Mardirossian, A., Chew, E. (2006), Music summarization via key distributions: Analyses of similarity assessment across variations, *Proceedings of the International Society for Music Information Retrieval Conference*.

Margulis, E.H. (2012), Musical repetition detection across multiple exposures, *Music Perception,* 29 (4*),* 377-385.

Margulis, E.H. (in print), Aesthetic responses to repetition in unfamiliar music, *Empirical Studies of the Arts*.

Marsden, A. (2010): Recognition of variations using automatic Schenkerian reduction, *Proceedings of the International Society for Music Information Retrieval Conference*, 501–506.

Mazzola, G. (2002) *The topos of music*. Birkhäuser.

Medin, D.L., Goldstone, R.L., Gentner, D. (1993), Respects for similarity, *Psychological Review*, 100(2), 254–278.

Melen, M. & J. Wachsmann, J. (2001), Categorization of musical motifs in infancy, *Music Perception,* 18, 325–346.

Meredith, D., Lemström, K. and Wiggins, G. (2002), Algorithms for discovering repeated patterns in multidimensional representations of polyphonic music, *Journal of New Music Research*, 31(4), 321-345.

Meyer, L.B. (1973), *Explaining music,* U of California Press.

Meyer, L.B. (2000), *The spheres of music*, U of Chicago Press.

Middleton, R. (1990), *Studying popular music*, Ballmoor, Bucking- ham Open University Press.

Nelson, R. U. (1948), *The technique of variation*, U of California Press.

Nettl, B. (1983), *The study of ethnomusicology: Thirty-one issues and concepts*. Urbana-Champaign: University of Illinois Press.

Peretz, I. (2006), The nature of music from a biological perspective, *Cognition,* 100, 1-32.

Pickens, J. (2004): *Harmonic modeling for polyphonic music retrieval*, PhD thesis, University of Massachusetts Amherst.

Réti, R (1951), *The thematic process in music*, Macmillan, New York.

Rizo Valero, D. (2010), Symbolic music comparison with tree data structures. PhD-thesis, Universidad de Alicante, 2010.

Rolland, P.Y. (1999), Discovering patterns in musical sequences, *Journal of New Music Research*, 28(4), 334-350.

Sagrillo, D. (1999), *Melodiegestalten im luxemburgischen Volkslied: Zur Anwendung computergestützter Verfahren bei der Klassifikation von Volksliedabschnitten*, Bonn: Holos-Verlag.

Schenker, H. (1935), *Der freie Satz*, Neue musikalische Theorien and Phantasien, 3. Vienna: Universal-Edition A.G.

Schoenberg, A. (1995), *The musical idea and the logic, technique, and art of its presentation*, ed. and transl. P. Carpenter and S. Neff, New York: Columbia University Press.

Schoenberg, A. (1967), *Fundamentals of musical composition*, London: Faber & Faber, 1967.

Sisman. E. (accessed May 1, 2012), "Variations." In *Grove Music Online*. *Oxford Music Online*, http://www.oxfordmusiconline.com/subscriber/article/grove/music/29050

Smith, L.B. (1989), From global similarities to kinds of similarities: the construction of dimensions in development, *Similarity and analogical reasoning* (eds: S. Vosniadou, A. Ortony), Cambridge U Press, 146-178.

Steedman, M., (1996), The blues and the abstract truth: Music and mental models. *Mental Models in Cognitive Science: Essays in Honour of Phil Johnson-Laird*, pp. 305–318, Psychology Press.

Suchoff, B. (1981), *The hungarian folksong*. Albany: State University of New York Press.

Toussaint, G.T. (2004), A comparison of rhythmic similarity measures. *Procedings of the International Society for Music Information Retrieval Conference*, 242—245.

Tversky, A. (1977), Features of similarity, *Psychological Review*, 84, 327-352.

Van der Merwe, P. (1989), *Origins of the popular style*, Clarendon Press, Oxford.

Van Kranenburg, P. (2010), *A Computational approach to content-based retrieval of folk song melodies*. Ph.D. thesis. Utrecht University.

Van Kranenburg, P., Volk, A., Wiering, F., Veltkamp, R.C, (2009) Musical models for folk-song melody alignment, *Proceedings of the International Society for Music Information Retrieval Conference*, 507–512.

Van Kranenburg, P., Volk, A., Wiering (2012), On identifying folk song melodies employing recurring motifs, *Proceedings of the International Conference of Music Perception and Cognition*.

Volk, A. (2005), Modelling a Processive Perspective on Meter in Music, *Proceedings of the International Computer Music Conference*, Barcelona, Spain, 251-254.

Volk, A., Garbers, J.,Van Kranenburg, P., Wiering, F., Veltkamp, R.C., Grijp, L.P. (2007), Applying rhythmic similarity based on inner metric analysis to folksong research, *Proceedings of the International Society for Music Information Retrieval Conference*, 293-296.

Volk, A. (2008), Persistence and change: local and global components of metre induction using Inner Metric Analysis, *Journal of Mathematics and Computation in Music*, 2, 99-115.

Volk, A. and Van Kranenburg, P. (in print), Melodic similarity among folk songs: An annotation study on similarity-based categorization in music, *Musicae Scientiae*.

Vosniadou, S. and A. Ortony, A. (1989): Similarity and analogical reasoning: a synthesis, *Similarity and analogical reasoning* (eds: S. Vosniadou, A. Ortony), Cambridge U Press, 1-15.

Wang, J.C., Lee, H.S., Wang, H.M., Jeng, S.K. (2005), Learning the similarity of audio music in bag-of-frames representation from tagged music data*, Proceedings of the International Society for Music Information Retrieval Conference*, 85–90.

Welker, R.L. (1982), Abstraction of Themes From Melodic Variations*, Journal of Experimental Psychology*: *Human perception and Performance*, 8, 435-447.

Wiora, W. (1941). Systematik der musikalischen Erscheinungen des Umsingens. *Jahrbuch für Volksliedforschung*, 7, 128–195.

Zbikowski, L.M. (2002), *Conceptualizing music*, Oxford University Press.

Ziv, N. & Eitan, Z. (2007), Themes as prototypes: Similarity judgments and categorization tasks in musical contexts, *Musicae Scientiae*, Discussion Forum 4A, 99–133.