

# Melody Beyond Notes

*A Study of Melody Cognition*

Sven Ahlbäck

GÖTEBORGS UNIVERSITET

Humanistiska fakulteten

## **Abstract**

### **Melody beyond notes - a study of melody cognition**

*Keywords:* Melody, Cognition, Melodic segmentation, Melodic Parallelism, Pitch Structure, Meter, Rhythm, Grouping, Swedish Folk Music, Music Theory, Computer-aided analysis

This thesis is a music theoretical approach to cognition of surface structure in monophonic melodies. It can briefly be described as a study into what extent we may acquire a common experience of melodic structure, such as *phrase* structure, only from listening to a melody. More precisely, this work concerns the question as to whether a cognitively based method of analysis can provide analyses of melodic surface structures in different styles that will concur with listeners' conceptions better than chance.

In order to investigate this question a general model of melody cognition was developed, relying primarily on a few general cognitive principles. The model was designed to be general in the sense that it should apply to any style for which the concept of melody is relevant. This model provided the framework for a computer-aided method of analysis, which performs analysis of different aspects of melodic surface structure based on information of relative pitch and temporal information only. These aspects involve: Categorical perception of pitch and duration at basic levels, such as context-sensitive quantization and melodic pitch categorization; Analysis of metrical and non-metrical temporal structures, e.g. heterometric structures; phrase and section structure, including analysis of structural implications of melodic similarity, structural hierarchy and symmetry. This development has required new theoretical concepts and methods to be created, e.g. regarding the relationship between rhythm and meter, some of which are presented for the first time in this thesis.

In order to evaluate the performance of the model a series of listener tests were performed, which together with corpora of musical notations from different styles, constituted the reference material of the study. This material has included Scandinavian folk music styles and Western classical music, but also examples of Eastern European folk music, Middle East and Indian Classical music, Jazz and Western popular song.

The results of these tests indicated that melody can be conceived differently by people even within a limited cultural sphere. But the results also suggested that this variability is possible to model by a rule-based method of analysis, since the predictions given by the model generally were well above chance level. It is herein suggested that variability in grouping conception to a considerable degree can be accounted for in terms of start- and end-oriented grouping preference. Moreover, the results also indicate that important aspects of even culturally foreign music can be conveyed, also in limited melodic stimuli.

Generally, the results support the assumption of a general cognitive framework for melodic surface structure. This might be interpreted as to indicate that, metaphorically speaking; melody may indeed be a universal 'language', but one which we all understand in our own way.

© Sven Ahlbäck 2004

Skrifter från Institutionen för musikvetenskap, Göteborgs universitet, nr 77, 2004

Tryckt av Intellecta Docusys, Västra Frölunda 2004

ISBN 91 85974 73-0

ISSN 1650-9285

# Contents

<b>PREFACE</b> .....	<b>7</b>
<b>1 INTRODUCTION</b> .....	<b>9</b>
1.1 THE FUNDAMENTAL QUESTION .....	9
1.2 THE OBJECT OF THE PRESENT STUDY .....	13
1.3 OUTLINE OF THE PRESENT STUDY .....	15
1.3.1 <i>General perspectives</i> .....	15
1.3.2 <i>The current model of melody cognition</i> .....	16
1.3.2.1 The process of melody cognition .....	16
1.3.2.2 Dimensions of melodic surface structure .....	18
1.3.2.3 The application of the model – the MODUS implementation .....	19
1.3.2.4 Outline of this thesis .....	22
1.3.2.5 The computer implementation – The MODUS environment .....	23
<b>2 RELATED WORKS</b> .....	<b>25</b>
2.1 INTRODUCTION .....	25
2.2 COOPER & MEYER: THE RHYTHMIC STRUCTURE OF MUSIC .....	26
2.3 LERDAHL & JACKENDOFF : A GENERATIVE THEORY OF MUSICAL STRUCTURE .....	27
2.4 NARMOUR: THE IMPLICATION-REALIZATION MODEL .....	29
2.5 RUWET AND NATTIEZ: PARADIGMATIC ANALYSIS .....	30
2.6 LEVY: “THE WORLD OF THE GORRLAUS SLÄTTS” .....	31
2.7 CAMBOUROPOULOS: “TOWARDS A GENERAL COMPUTATIONAL THEORY OF MUSICAL STRUCTURE” .....	32
2.8 TEMPERLEY: THE COGNITION OF BASIC MUSICAL STRUCTURES .....	34
<b>3 THE GENERAL MODEL</b> .....	<b>37</b>
3.1 FUNDAMENTAL ASSUMPTIONS .....	37
3.2 PRINCIPLES OF MELODIC STRUCTURE .....	39
3.3 DIMENSIONS OF MELODIC SURFACE STRUCTURE .....	40
<b>4 ANALYSIS OF MELODIC PITCH CATEGORIES</b> .....	<b>45</b>
4.1 THE CONCEPT OF MELODIC PITCH CATEGORIES .....	45
4.2 THE SIGNIFICANCE OF MELODIC PITCH CATEGORIES TO THE ANALYSIS OF MELODIC STRUCTURE .....	49
4.3 BASIC INFORMATION NEEDS FOR ANALYSIS OF MELODIC PITCH CATEGORIES .....	49
4.4 BASIC ANALYTICAL CONCEPTS .....	49
4.4.1 <i>Basic category discrimination</i> .....	49
4.4.2 <i>Structural independency - Chromaticism</i> .....	50
4.4.3 <i>Ideal sets of melodic pitch categories</i> .....	51
4.4.4 <i>Division of the octave – perfect fifth/ fourth and equidistant tuning</i> .....	52
4.4.5 <i>Additional analytical considerations</i> .....	53
4.4.6 <i>Individual differences in categorization – limitations of the method</i> .....	54
4.5 METHOD OF ANALYSIS IN OUTLINE .....	54

4.5.1	<i>Overview of the procedure</i> .....	54
4.5.2	<i>Interval implications</i> .....	55
4.6	<b>EVALUATION OF THE ANALYSIS OF MELODIC PITCH CATEGORIES</b> .....	64
4.6.1	<i>Means of Evaluation</i> .....	64
4.6.2	<i>The significance of the MPC analysis at the level of melodic surface structure</i> .....	65
4.6.3	<i>Chromaticism and extra-MPC alterations</i> .....	69
4.6.4	<i>Evaluation of the success of the Analysis of Melodic Pitch Categories</i> .....	72

## 5 METRICAL ANALYSIS .....73

5.1	<b>METER – GENERAL CONCEPTS</b> .....	73
5.1.1	<i>Meter and Gestural Rhythm</i> .....	73
5.1.2	<i>Pulse</i> .....	76
5.1.3	<i>Metrical complexity</i> .....	77
5.1.4	<i>Accentuation and the impulse quality of meter</i> .....	81
5.2	<b>METRICAL ANALYSIS – METHOD DESIGN</b> .....	83
5.2.1	<i>Outline</i> .....	83
5.2.2	<i>Metrical Analysis I - Metrical Quantization</i> .....	83
5.2.2.1	General concept.....	83
5.2.2.2	Method design.....	84
5.2.2.3	Examples of the performance of the quantization method.....	86
5.2.3	<i>Metrical Analysis II – Beat Atom Analysis</i> .....	94
5.2.3.1	General concept.....	94
5.2.3.2	Method design.....	95
5.2.4	<i>Metrical Analysis III – Complex Metrical Analysis</i> .....	99
5.2.4.1	Overview.....	99
5.2.4.2	Analysis of low-level grouping – general principles.....	101
5.2.4.3	Principles of low-level grouping.....	106
5.2.4.4	Implementation principles.....	109
5.2.4.5	Segmentation indication by sequences – method design.....	110
5.2.4.6	Segmentation indication by discontinuity/change – method design.....	121
5.2.4.7	Tracking periodical grouping – composite pulse and time.....	145
5.2.4.8	Final analysis of primary beat-grouping.....	161
5.2.5	<i>Means of evaluation</i> .....	173
5.2.6	<i>Evaluation of Metrical Quantization Analysis</i> .....	174
5.2.7	<i>Beat atom analysis - Basic analysis of pulse levels by onset information</i> .....	174
5.2.7.1	General discussion.....	174
5.2.7.2	The performance of the beat analysis regarding obscured beat onsets by pre-onsets and syncopation in Swedish folk music transcribed from live performances.....	176
5.2.7.3	Syncopation/obscured beat-atom in 20 <sup>th</sup> century popular music – “Fascinatin’ rhythm” by G. Gershwin.....	178
5.2.7.4	Irregular syncopation relating to a metrical organization provided by accompaniment – Bolero by M. Ravel.....	179
5.2.7.5	Phase displaced pulse – Example from Norwegian ‘Halling’ music.....	180
5.2.7.6	Uneven metrical complexity.....	186
5.2.8	<i>Complex metrical analysis – metrical grid analysis</i> .....	188
5.2.8.1	General description.....	188
5.2.8.2	Experiment 1. Sensitivity to grouping by pitch sequencing (1996-97).....	188
5.2.8.3	Experiment 2. Sensitivity to grouping by pitch discontinuity and pitch sequencing at metrical levels.....	192
5.2.8.4	Evaluation of complex metrical analysis – metrical grid analysis and asymmetrical beat group analysis through comparison with culturally informed notations.....	230

## 6 METRICAL PHRASE AND SECTION ANALYSIS.....250

6.1	WHAT ARE WE LOOKING FOR? GENERAL CONCEPTS.....	250
6.1.1	<i>The syntactic and hierarhic structure of melody.....</i>	250
6.1.2	<i>Start- and end-oriented grouping.....</i>	251
6.1.3	<i>The concept of sequence.....</i>	253
6.1.4	<i>Implications of metrical structure for melodic structure on phrase and section level.....</i>	254
6.1.5	<i>The means of melodic structure in the view of melodic similarity.....</i>	256
6.1.5.1	Pitch structure.....	256
6.1.5.2	Rhythmic structure.....	258
6.1.5.3	The relationship between pitch and rhythmic structure as structural determinants.....	260
6.1.6	<i>Hierarchy of melodic segments.....</i>	261
6.1.7	<i>General grouping principles.....</i>	261
6.2	THE DESIGN OF THE METRICAL PHRASE AND SECTION ANALYSIS.....	264
6.2.1	Overview.....	264
6.2.2	<i>The relationship to beat structure – superimposed-meter-alert.....</i>	264
6.2.3	<i>The sequence analysis – Segmentation by melodic parallelism.....</i>	266
6.2.3.1	Typology of melodic similarity.....	266
6.2.3.2	Classes of sequences.....	266
6.2.3.3	General classification of pitch similarity – NIL –sequences.....	268
6.2.4	<i>Evaluation of sequence structure.....</i>	284
6.2.4.1	Outline – general design.....	284
6.2.4.2	Quantifying structural prominence of sequences.....	288
6.2.4.3	Quantifying structural hierarchy.....	295
6.2.4.4	Segmentation by symmetrical implication.....	297
6.2.4.5	Successive evaluation of sequence structure - overview.....	302
6.2.5	<i>Local phrase boundary analysis – Segmentation by discontinuity, periodicity and symmetry.....</i>	315
6.2.5.1	Outline.....	315
6.2.5.2	Method of local phrase boundary analysis.....	316
6.2.6	<i>Final analysis – start- and end-oriented interpretation.....</i>	321
6.3	EVALUATION OF METRICAL PHRASE AND SECTION ANALYSIS.....	327
6.3.1	<i>The means of evaluation.....</i>	327
6.3.1.1	General problems.....	327
6.3.1.2	Comparisons with expert analyses and score information.....	327
6.3.2	<i>Listener tests.....</i>	328
6.3.2.1	How to evaluate the success of the model in relation to listener segmentations?.....	329
6.3.2.2	Outline.....	330
6.3.3	<i>Phrase structure in Swedish instrumental folk tunes.....</i>	330
6.3.3.1	General problems.....	330
6.3.3.2	Structural ambiguity by asymmetry within a symmetrical general structure.....	332
6.3.3.3	Structural ambiguity by disguised phrase starts.....	340
6.3.3.4	Structural ambiguity by melodic variation within sequences.....	343
6.3.3.5	A quantitative test of the model in relation to melodic segmentations by E. Övergaard.....	349
6.3.4	<i>Phrase structure in Norwegian gangar melodies.....</i>	354
6.3.4.1	General problems.....	354
6.3.4.2	Example 1. Sordölen. A chained sequence structure.....	355
6.3.4.3	Example 2. Norafjells. Asymmetrical hierarchy.....	360
6.3.5	<i>Phrase structure challenging the metrical structure – examples from Baroque music and 20<sup>th</sup> century Western popular music.....</i>	374
6.3.5.1	The fortpinning principle - Asymmetry in between global and local symmetry.....	374
6.3.5.2	20 <sup>th</sup> century Western popular music – separation of meter and phrase structure.....	386
6.3.6	<i>Sequenced and non-sequenced phrase structure within asymmetrical metrical structure.....</i>	397
6.3.6.1	Balkan dance music with asymmetrical beat structure – Transformation of phrase identity within a symmetrical framework.....	397

6.3.6.2	Non-sequenced asymmetrical segment structure within a symmetrical general framework –south Indian classical music.....	403
6.3.7	<i>Evaluation by listener tests</i> .....	409
6.3.7.1	Purpose of the tests .....	409
6.3.7.2	General methodology.....	409
6.3.7.3	Results.....	411
6.3.7.4	Summary of results of listeners tests.....	448
6.3.8	<i>Comparison with other models</i> .....	449
6.3.8.1	Evaluation of Grouper and LBDM models .....	449
6.3.8.2	Comparisons of results.....	452
6.3.9	<i>Conclusions</i> .....	460
6.3.9.1	General success of the model.....	460
6.3.9.2	Interpretation of the results in relation to the basic assumptions for the model.....	461
6.3.9.3	Concluding remarks and future development.....	468
<b>7</b>	<b>NON-METRICAL FIGURE AND PHRASE ANALYSIS.....</b>	<b>472</b>
7.1	WHAT IS NON-METRICAL STRUCTURE? GENERAL CONCEPTS .....	472
7.1.1	<i>Outline</i> .....	472
7.1.2	<i>Non-metrical and quasi-metrical structure</i> .....	472
7.1.2.1	Determining temporal relationships in a non-metrical context.....	476
7.1.2.2	Levels of non-metrical grouping structure.....	479
7.2	METHOD OF ANALYSIS OF NON-METRICAL MELODIES.....	484
7.2.1	<i>General design of the model</i> .....	484
7.2.2	<i>Non-metrical quantization</i> .....	487
7.2.2.1	From musical sound to midi.....	487
7.2.2.2	Analysis of melodic pitch categories (MPC).....	488
7.2.2.3	Categorical quantization of event durations.....	489
7.2.2.4	Preliminary phrase analysis.....	493
7.2.2.5	Primary grouping analysis at figure level.....	495
7.2.2.6	Phrase analysis at central phrase level.....	499
7.2.2.7	Phrase analysis at higher phrase levels.....	500
7.3	ANALYSIS OF QUASI-METRICAL MELODIC STRUCTURE.....	503
7.3.1	<i>General problems</i> .....	503
7.3.2	<i>Analysis of two melodies with quasi-metrical structure</i> .....	504
7.3.2.1	From Audio-to-midi-conversion to preliminary phrase analysis.....	504
7.3.2.2	Primary grouping analysis at figure level .....	509
7.3.2.3	Phrase analysis.....	513
7.4	EVALUATION OF NON-METRICAL ANALYSIS.....	516
7.4.1	<i>Means of evaluation</i> .....	516
7.4.2	<i>Discussion</i> .....	517
<b>APPENDIX: LIST OF ASSUMPTIONS .....</b>		<b>520</b>
ASSUMPTIONS IN CHAPTER 4 .....		520
ASSUMPTIONS IN CHAPTER 5 .....		521
ASSUMPTIONS IN CHAPTER 6 .....		524
ASSUMPTIONS IN CHAPTER 7 .....		527
<b>REFERENCES .....</b>		<b>530</b>
SOURCES OF MUSIC EXAMPLES.....		540
RECORDINGS.....		542
SOFTWARE AND TECHNICAL EQUIPMENT.....		543

# Preface

---

*“Three things belong to composing, first of all melody; then again melody; then finally, for the third time, melody” (Salomon Jadassohn (1831-1902), Book of instrumentation 1889)*

This was going to be something quite different from what it turned out to be... In the late 1980s, the professor in musicology at Göteborg University, Jan Ling, encouraged me to go further with my studies in tonality in Swedish folk music as a postgraduate student. During the inspiring seminars at the Department of Musicology both my supervisor Jan Ling and my fellow students questioned if the aspects of tonality I was studying, such as microtonal alterations, actually were perceivable to people in general; and moreover, to what extent the structural analyses I presented were influenced by my own intuitions.

This led me to take a quite different approach to the subject. I started to explore the literature within the field of music perception and cognition, performed tests to investigate how listeners perceived tonality in the music I was interested in. During this work I received invaluable and inspiring guidance from professor Alf Gabrielsson at the department of Psychology at Uppsala University. Furthermore, I started to develop a formalized method of analysis of tonality, in order to rule out possible influences of musical intuitions in the analytical process. Considering the highly complex matter of analysis of tonality I soon found out that a computer implementation of the analytical method would be a perfect tool for this purpose.

Thanks to the inspiring help from professor Johan Sundberg and PhD Anders Friberg at the department of Speech, Music and Hearing at the Royal Institute of Technology (KTH) in Stockholm I got in touch with a student at KTH, Sven Emtell, who accepted to collaborate with me in the development of a computerized method of analysis of tonality, as a part of his examination in computer science. Since the results of the listener tests and the structural analyses indicated that melodic surface structure was influential in tonality cognition, the suggested method of analysis involved analysis of phrase structure in melodies. The plan was to develop the computer model during three months in the summer of 1992. However, this turned out to be a considerably more complex task than we thought and Sven Emtell was kind enough to continue the development of the model over a period of two years; still, the performance of the model was not sufficient for even the first stage of the analytical process.

At that time, I continued the development of the model with invaluable help from Kalle Mäkilä, a computer programmer but also fiddle student of mine. He inspired me to learn the LISP programming language which made it possible for me to continue the development by myself.

It turned out that the original question about the tonality structure in older-style Swedish folk music was never reached. The elusive nature of melodic structure has consistently forced me to reject previous premises and required my full attention. The focus has become the cognition of melodic structure, relating to the general question of how music is communicated. This has furthermore forced me to cross the boundary of the original stylistic

perspective, since music is a universal feature of humans. A method of analysis based on cognitive principles must hence view the specific traits of a style from a general perspective.

During this work it has become clear to me that music as a means of communication and expression, is what really interests me with music – which both in spite of and thanks to its elusive, ambiguous and mysterious qualities can express what is essentially human. This study concerns only one aspect of the cognition of music, the cognition of melodic surface structure from the perspective of music theory.

There are many, who have contributed significantly to this work. I would first of all like to express my deepest gratitude to my supervisors, prof. Jan Ling, prof. Olle Edström and prof. Alf Björnberg at the Department of Musicology at Göteborg University and prof. Alf Gabriellson at the Department of Psychology at Uppsala University who have all supported me with invaluable comments, through the long period during which the work has been done. Furthermore, without the support from prof. Johan Sundberg and PhD Anders Friberg of the Music Acoustics Group within the Department of Speech, Music and Hearing at Royal Institute for Technology in Stockholm, the development of the model would not have been possible. Prof. Carol Krumhansl, Department of Psychology at Cornell University, has made valuable comments on the direction of the work at one point in the process and docent Märta Ramsten at the Center for Swedish Folk Music and Jazz research, has supplemented research material.

Profound contributions have also been made by Sven Emtell and Kalle Mäkilä, who both, besides their direct involvement in the development of the model – which required significantly more work than was originally scheduled – have been kind enough to introduce me to the basics of computer programming.

I would also like to thank my colleague at the Royal College of Music in Stockholm, Bill Brunson, who has contributed greatly by trying to make my English comprehensible – with unlimited patience and respect for the difficulties involved when writing in a foreign language. All errors as regards language are my own.

Furthermore I am much obliged to my fellow colleagues at the Royal College of Music, in particular Rector, Prof. Gunilla von Bahr, who have supported me personally during the completion of the thesis. My colleagues at the Department of Folk Music have supported me by their patience with the inconveniencies which my involvement in this work have caused.

I would also express my deepest gratitude to my students at KMH, over the last 10 years, whom have been patient with me not being able to give them my full attention at all times. Moreover, they have contributed significantly by taking part in endless listening tests generally involving musically very boring examples. In this connection I would also like to mention my friends and colleagues, in particular Ole Hjorth, Ellika Frisell, Mikael Marin, Karin Rehnqvist and Shivakumar K., who like my family and close relatives have contributed greatly by accepting to be subjects in tests, but also by supporting me during the process. It is the musical minds of all these people who have made this work possible. In particular, my father and mother and my parents in law, have constantly supported me throughout the work.

But most of all this work owes to my wife, friend and colleague Susanne Rosenberg, who has supported me in every possible respect over the years. This is for you, with gratitude.

*Stockholm, January 14<sup>th</sup>, 2004*

Sven Ahlbäck

# 1 Introduction

---

## 1.1 The fundamental question

*“Music is the universal language of mankind.”*  
 Henry Wadsworth Longfellow (1807-82) *Ontre-Mer*

*“It is the accent of languages that determines the melody of each nation”*  
 Jean-Jacques Rousseau (1712-78) *Dictionnaire de musique*, 1767

In a recent televised discussion about the new policy of the *Swedish Radio* to allow popular songs with English lyrics in *Svensktoppen*<sup>1</sup>, the company representative argued that since “*music is a universal language*”, there is no such thing as specifically Swedish music, and consequently the company has no responsibilities to maintain this illusive category of music.

Not much later, during a public discussion of future directions of research in music, one of the more distinguished Swedish musicologists and a colleague of mine stated: “*I don’t believe in any musical universals*”.

The extent to which music communicates across socio-cultural barriers, and conversely, to what degree the intelligibility of music is limited to a socio-cultural context, is the subject of a never-ending discussion within the different sub-disciplines of musicology and related disciplines. It surely relates to the more general question of human behavior and cognition as being the product of the environment with which the human beings interact or a product of the biologically determined capabilities of humans.

Although many researchers in music today seem to agree that different aspects of musical behavior and cognition need to be understood both in the socio-cultural context and in the context of the biological and psychological capabilities of the human being, it is not hard to find a general disagreement between, on the one hand, researchers schooled in the humanities and social sciences and on the other hand researchers within the natural sciences, such as e.g. cognitive psychologists and biomusicologists. Among the traditional musicologists, one finds often the music historians favoring the concept of music as a product of culture, while music theorists often seem to favor a more scientifically influenced view of music as a product of human cognition.

In his influential book “*How Musical is Man*”, the British ethnomusicologist John Blacking provides a radical general criticism of the study of music removed from its socio-cultural context (Blacking 1973/1976)<sup>2</sup>. His chief argument is:

*“Functional analyses of musical structure cannot be detached from structural analyses of its social function: the function of tones in relation to each other cannot be explained adequately as part of a*

---

<sup>1</sup> the “*Swedish Top List*”

<sup>2</sup> More specifically, he argues against the elitist and exclusive view of musicality which he attributes to Western society and for the application of ethnomusicological method in the study of Western music.

## Chapter 1

*closed system without reference to the structures of the sociocultural system of which the musical system is a part and to the biological system to which all music makers belong” (Blacking 1973:30-31)*

Furthermore, Blacking argues against the possibilities of understanding musical structure without the preconceptions given by an experience of its cultural context:

*“Music can express social attitudes and cognitive processes, but it is useful and effective only when it is heard by the prepared and receptive ears of people who have shared, or can share in some way the cultural and individual experiences of its creators.” (Blacking 1973:54)*

This argument is expressed even more radically concerning the creation of music in his study of the music making among the *Venda* people in South Africa:

*“...In order to create new Venda music, you must be a Venda, sharing Venda social and cultural life from early childhood. [...] I am convinced that a trained musician could not compose music that is absolutely new and specifically Venda, and acceptable as such to Venda audiences, unless he had been brought up in Venda society. Because the composition of Venda music depends so much on being a Venda, and its structure is correspondingly related to that condition of being, it follows that an analysis of the sound cannot be conceived apart from its social and cultural context.” (Blacking 1973:98)<sup>3</sup>*

Blacking takes thus a view that contrasts radically to the popular view of music as a universal language of mankind. He supports his notion with examples of how a culturally uninformed analysis of Venda music might lead to misconceptions about structural features, its historical development and its musical content.

However, Blacking has a problem: If an understanding of musical structure is not possible across the boundaries of socio-cultural barriers, what is the purpose of the study of ethnomusicology when there are no grounds for comparisons between different musical systems? And if so why do people listen to music from foreign cultures, and moreover, sometimes devote their lives to be performers of music of foreign cultures, which they consequently lack the possibilities to understand? If significance of musical content is limited to cultural preconceptions – if it has meaning only in its context – how is it possible to appreciate and why even bother about music from a foreign culture? And, how were then the rapid changes of musical behavior and taste that have been taking place in e.g. the Far East Asia during the last hundred years possible?<sup>4</sup>

Blacking admits this problem, thus modifying his point of standing by admitting that:

*“Music can transcend time and culture. Music that was exciting to the contemporaries of Mozart and Beethoven is still exciting, although we do not share their culture and society. [...] Many of us*

---

<sup>3</sup> He further concludes that: “Music is not a language that describes the way society seems to be, but a metaphorical expression of feelings associated with the way society really is. It is a reflection of and response to social forces, and particularly to the consequences of the division of labor in society.” (Blacking 1973:104)

<sup>4</sup> Blacking’s argument that “music cannot change society, but at its best confirms situations that already exists” implies that music cannot influence society. But this implies that the dissidents of the former communist regimes were completely wrong when regarding Western music as an eye-opener towards other aspects of western cultures. And furthermore, are the lifestyle and values connected with e.g. reggae or hip hop musical styles prerequisites for the appreciation of the music or can the music sometimes be the factor that influence the adaptation to a change of lifestyle?

*are thrilled by koto music from Japan, sitar music from India, Chopi xylophone music, and so on. I do not say that we receive the music in exactly the same way as the players (and I have already suggested that even the members of a single society do not receive their own music in the same ways), but our own experiences suggest that there are some possibilities of cross-cultural communication. I am convinced that the explanation for this is to be found in the fact that at the level of deep structures in music there are elements that are common to the human psyche although they may not appear in the surface structures.* “(Blacking 1973:109)

The British musicologist Nicholas Cook takes a similar position in “Music, Imagination and Culture” (Cook 1990), where he thoroughly discusses the meaning of understanding music from different perspectives. More specifically he addresses the question of requirements of a true understanding of musical form. On the one hand he refers to e.g. the American music theorist Leonard B. Meyer, citing his claim that deeper experience of music requires knowledge of style-dependent norms:

*“An American must learn to understand Japanese music just as he must learn to speak the spoken language of Japan (1956:62)” (Meyer 1956:62)*

On the other hand, he refers to Blacking, who obviously has developed his revised position cited above, argues that:

*“It is sometimes said that an Englishman cannot possibly understand African, Indian, and other non-English music. This seems to me as wrong-headed as the view of many white settlers in Africa, who claimed that blacks could not possibly appreciate and perform properly Handel’s Messiah, English part-songs, or Lutheran hymns. Of course music is not a universal language, and musical traditions are probably the most esoteric of all cultural products. But the experience of ethnomusicologists, and the growing popularity of non-European musics in Europe and America and of Western music in the Third World, suggest that the cultural barriers are somewhat illusory, externally imposed, and concerned more with verbal rationalizations and explanations of music and its association with specific social events, than with the music itself. ... When the words and labels of a cultural tradition are put aside and ‘form in tonal motion’ is allowed to speak for itself, there is a good chance that English, Africans and Indians will experience similar feelings. (Blacking 1987: 129-30, cited after Cook 1990:150)*

Cook himself rather takes Blackings position, implying that within the limitations given by the time and society that musical communication at some level is possible:

*“It is almost inevitable that a Westerner will misinterpret the music of a foreign culture when he listens with pleasure to it, just as he will probably misinterpret the past music of his own culture, in that he will not understand it in total conformity with the manner in which the musicians who produced it understood it or expected it to be understood. [...] We listen to it not in order to understand it in the manner which Indian and Vanda musicians understand it, but to enjoy it. And we can do this without going to live in India or Africa, just as we can listen to Machaut’s music, and enjoy it, without going to live in fourteenth-century France.” (Cook 1990:151)*

What Cook suggests is thus that there are different levels at which music can be understood. He provides evidence that what he designates as “musicological”, i.e. culturally informed or deeper understanding of music which is the kind of intrinsic understanding of a

## Chapter 1

musical style to which Meyer refers, is not always accessible to listeners in general within a given culture (e.g. Cook 1990:50-68).

But if Blacking and Cook are right in their assumptions, that music at some level is understandable across cultural boundaries, what determines this possibility and to what extent is this possible?

Blacking refers to universal cognitive, biological and social restraints on musical systems, in particular structural principles akin to gestalt psychological principles:

*“There seem to be universal structural principles in music, such as the use of mirror forms [...], theme and variation/ repetition, and binary form. It is always possible that these may arise from experience of social relations or of the natural world: an unconscious concern for mirror forms may spring from the regular experience of mirror forms in nature, such as observation of the two “halves” of the body.” (Blacking 1973:112)*

During the past 30 years, there have been an increasing number of cross-cultural studies using methodology from cognitive psychology<sup>5</sup>. These studies, generally comparing the performance of a group of Western listeners with a group of non-Western listeners for different experimental tasks involving music of different cultural origin, suggests that general cognitive principles apply to music of different cultures. In the words of Krumhansl et al (2000):

*“The results [...] showed considerable agreement between the groups, and suggested that the inexperienced listeners [i.e. listeners not acquainted with a particular style] were able to adapt quite rapidly to different musical systems. Moreover, they showed that similar underlying perceptual principles appear to operate, and that listeners are sensitive to statistical information in novel styles which gives important information about basic underlying structures.” (Krumhansl et al 2000:14)*

Studies like these have been criticized from methodological perspectives (see e.g. Cook 1990:149, Cook 1987b). There are also other studies performed by ethnomusicologists the results of which indicate that basic aspects of music cognition are essentially culture-specific (see e.g. Merriam 1964). However, in spite of the methodological problems, there is converging evidence that there are principles of music cognition that have relevance for quite distant musical styles and which cannot be ignored.

To what extent similar cognitively-based structural principles are relevant for different musical styles and for different individuals is precisely the question in focus in the present study.

This is fundamentally a question about musical communication. If we believe that musical communication is possible, to what extent can music communicate across the boundaries between different individuals and across the boundaries given by the socio-cultural contexts of the individuals? Are the similarities as determined by general similarity in human experience and society and by the biologically determined cognitive capabilities of man greater than the differences or vice versa?

In order to study this question we need an analytical approach, which is not primarily style-dependent. Blacking emphasizes this point, implying that a theory that is not general

---

<sup>5</sup> e.g. Kessler, Hansen and Shepard 1984; Castellano, Bharucha and Krumhansl 1984; Vaughn 1993; Krumhansl 1995b, Krumhansl, Toivonen, Eerola, Toivainen, Järvinen and Louhivuori 1999 and 2000

enough to apply for any culture or society is automatically inadequate as a theoretical framework for a specific style.<sup>6</sup>

However, since Cook's argument that any analytical approach is made in a socio-cultural context from which the researcher cannot be excluded (Cook 1990:3) is obviously true, any model of analysis will be to some extent culture-specific. This further implies that an analytical model has to consider the cognitive and cultural generality of the analytical problem to which it refers.

To summarize, there seem to be good reasons to believe that music at some level can communicate across socio-cultural and individual boundaries, but that at the same time musical communication at other levels is restricted by those boundaries. The extent to which we can cross these boundaries by adaptation to musical structure in one particular field of musical communication is the subject of the present study.

What fascinates me most about music is the diversity of and richness in musical expressions that originate from the particularities of different musical styles. The non-universality of musical expressions on both individual and cultural level is really something that enriches the world of musics. But in order to understand and communicate such style-specific treats of music, including the understanding of a particular style by its own terms, a general theory of how humans cognize musical structures is essential.

## **1.2 The object of the present study**

The present study concerns only one very limited aspect of musical experience, namely the cognition of surface-structure in melody. Since the concept of melody in general is not relevant for all musical cultures or musical styles, the generality of the study is limited to musical styles in which melody can be regarded as a relevant concept.

Melody is, in this context, regarded as the conception of a gestalt involving pitch change over time, which is conceived as a significant exponent of musical expression with structural identity. Melody is here regarded as a fundamentally monophonic phenomenon, which means that a specifically melodic conception of pitch change over time implies the conception of pitch change between individual, more or less stable, pitch categories. Further, melody is used in the sense of a complex structure, involving the conception of at least one sub-structural level.

This implies that a phenomenal structure involving pitch change over time can be regarded a melody only when experienced as a significant musical gestalt, when it is regarded as comprehensible whole at some level. This further implies that substructures and individual events in the melody must, to some degree, be conceived as related to each other. The converse would be a series of pitch changes over time conceived just as a series of individual tones or events and nothing more, without any intelligible, conceivable inner structure, which could be designated an experience of a chaotic temporal pitch structure.

---

<sup>6</sup> Speaking of "The language of music" (Cooke 1959); "Cooke cannot be faulted for choosing a particular area of music, but, because his theory is not general enough to apply to any culture or society, it is automatically inadequate for European music." (Blacking 1973:72). See also e.g. Nettl 1964:98ff.

This further implies that one sound structure may be conceived as a melody by one individual and not a melody by another. The phenomenal definition of melody thus refers to phenomenal structures that may be conceived as melodies.

The particular question, which the present study attempts to address, is if it is possible to construct a relatively style-independent general theory of the cognition of melodic structure in monophonic melodies from cognitively based assumptions of the relationship between phenomenal structure and conceived structure by which it is possible to make relevant predictions of peoples conceptions of melodic surface structure.

Such a theory implies the development of a rule-based and formalized model, which is able to make testable predictions of cognitive implications of phenomenal structure.

Moreover, the study attempts to evaluate the influence of different dimensions of melodic surface structure by regarding only influence of relative pitch and absolute and relative duration in the evaluation of the theoretical model.

**The general hypothesis is that it is possible to make predictions about people's conceptions of melodic surface structures by the application of a general model based on cognitively-based assumptions of the relationship between phenomenal structure and conceived structure which concur with peoples conceptions better than chance using only information of relative pitch and relative and absolute duration of events.**

The validity of the hypothesis is evaluated by the development of such a model from generally gestalt psychologically based assumptions of melody cognition, transformed into a formalized computer implementation of the model. This is followed by subsequent testing of the application of the model on music of different cultural and stylistic origin for which the concept of melody as defined above is assumed to be relevant. The evaluation of the relative success of the model is then performed by comparisons of the results obtained by the application of the computer model with (1) results obtained from psychological experiments of melody cognitions; (2) with analyses melodic structure made by music analysts; (3) with conceptions of melody structure as manifested in musical notations; and (4) by using extra-musical cues as exponents of melody cognition, in particular song lyrics and dance patterns.

The aim is not to evaluate whether the hypothesis is generally valid, i.e. holds for all melodies (in the sense of the above definition), but rather to test the plausibility of the hypothesis in a given reference material.

The study thus has an essentially music theoretical approach, but borders on disciplines such as cognitive psychology and ethnomusicology. It is not a study within the field of computer science or artificial intelligence, even though the computer implementation of the model is an essential part of the evaluation of the model.

Since the study aims at evaluating to which extent the same cognitive rules are relevant for the cognition of melodic surface structure in music of different cultural and historical origin, it follows that that rules based on style- and culture-specific codes are generally excluded from the model. This implies that I have generally neglected, for instance, melodic structuring based on polyphony, such as e.g. grouping indications based on implications of harmony. This makes the model evidently less relevant in a style-specific perspective of Western music. It is, however, a prerequisite for the purpose of the study since music with and without harmonic structure is within the scope of the study.

For the same reason, implication of melodic structuring by tonality is generally excluded. Even if there are studies indicating that similar tonal structuring, such as consonance, operate in different musical systems (e.g. Krumhansl et al 2000), the diversity of tonal systems in different musical cultures is profound. For this reason, and the possibility to evaluate the influence of tonality on the cognition of melody structure and of melodic structure on the cognition of tonality, it needs to be excluded from the model.<sup>7</sup>

A common criticism regarding models of structural analysis of surface structure is that surface structure is created by the realization of deep structures, rather than the opposite way around<sup>8</sup>. But it is hard to neglect the fact that, if musical communication is possible through listening to musical sound, there must be a way to extract the deeper structures through the surface structure from which the music emerges. It is, as discussed above, not always clear that deeper structures connected which are important in the production of music in a certain culture are actually conceived by listeners even within a cultural context.

This study is limited to the study of basic aspects of surface structure in melody, in complete awareness of this being a limited scope regarding the concept of melody and in complete trust that what is really important in music cannot be expressed by words or analyses.

The ultimate goal of this work is, however, to contribute to the understanding of musical communication.

## 1.3 Outline of the present study

### 1.3.1 General perspectives

Music is, in this study, regarded from the perspective of the listener. It is by no means the only perspective which is relevant from the point of view of the objective of this study. It excludes e.g. the cognition of melody in terms of music production, such as the cognitive processes involved in music making, the kinesthetic aspect of motion involved in both music making and the cognition of music and the relationship between cognition and musical functions<sup>9</sup>.

As a musician, knowing from my own experience the large impact of doing on thinking and the close relationship between experience of movement and experience of music, I do not deny the relevance and benefit of such perspectives. However, I believe that the perspective of the listener is fundamental for the *general* understanding of musical communication across cultures and styles – and hence, for the cognition of musical structure – since the specifically

---

<sup>7</sup> However, this does not imply that all aspects of the cognition of pitch structure in melody is excluded (see below).

<sup>8</sup> For a summary, see Cook 1990:47ff

<sup>9</sup> An extensive literature within ethnomusicology has explored the relationships between music and other means of musical expression. Most notably regarding much African music, there seems to be a general agreement that any attempt to separate cognition of musical sound from cognition of musical movement, both in terms of interaction with music and in terms of production of music, is foreign to the concept of music within the culture (see e.g. Merriam 1964, Blacking 1973, Nketia 1974 Baily 1985). Such a perspective is actually relevant in the study of many musics, to study the cognition of e.g. dance music or religious chant, without regarding its functions is evidently to exclude the some of the most important aspects of the cognition of the music.

musical domain is the communication by means of sound. Experiencing music by listening is thus thought to be a greater category, which includes sub-categories such as integrated music and dance experience. Similarly, the cognition of the production of music will be a sub-category to listening to music as long as there are listeners who are not performers and when people hear sounds not intended to be music as music.

Since we are here dealing with musical communication in a general perspective, the listener's perspective seem to be the most general, as long as we regard communication by sound to be central to music.

Throughout this study I am also making frequent use of the concepts of cognition and perception and derivations thereof, sometimes in a way that may be regarded as misuse of these concepts from a psychological point of view. The use of these concepts in this study is, however, related to the general theory of melody cognition as presented in section 1.3.2 in this chapter. I use cognition with reference to mental representations or constructs, regardless of whether this involves any awareness in the sense that it can be expressed verbally or not. Perception, in the sense it is used throughout this study, refers to assimilation, which does not involve mental representation.

Moreover, the current model involves to some extent the introduction of new music theoretical concepts, and further involves the redefinition of some common music theoretical terms. This may seem confusing and inconvenient, but is a consequence of the general approach taken in this study.

## 1.3.2 The current model of melody cognition

### 1.3.2.1 The process of melody cognition

The general model of melody cognition, which is the core of the present work, is presented in detail in Chapter 3. It is based on a limited number of assumptions of the psychological foundations for the cognition of melody.

The core of these assumptions is the concept of *the perceptual present* (Michon 1978, Fraise 1982), understood as the basic temporal scope of attention for the immediate perception of melodic structure at primary levels, and what I have called *the categorical present*, the category capacity of short-term memory (Miller 1956). In addition, I have postulated that the perception and correspondent cognition of durational proportions is restricted to the terms given by what I have called *the categorical proportion rule*, with reference to psychological research (e.g. Fraise 1956, Povel 1981). Further, I have proposed that the cognition of temporal structures in melody may be hierarchical, that higher temporal scopes of attention may be formed by categorical reduction of lower-level scopes of attention, for which I have proposed the term *parallel nows* (which denotes a plural of now).

Besides these basically contextual assumptions, I am proposing a set of general grouping principles, defined in terms of gestalt psychological rules, which are active in melodic structuring. This model includes the relative prominence and interaction of these principles on different levels of melody cognition.

The core of this model is the classification of the different grouping principles by the role they play in the cognition of grouping structure in melody. It postulates that groups are primarily formed by the categorization into same and different, i.e. grouping by similarity and discontinuity, designated *primary grouping principles*. It further postulates that groups can be

formed by implication based on properties of grouping determined by primary grouping principles. These *secondary grouping principles* which denotes grouping by good continuation and symmetry. The third class, *tertiary grouping principles*, are defined as being active in the perceptual

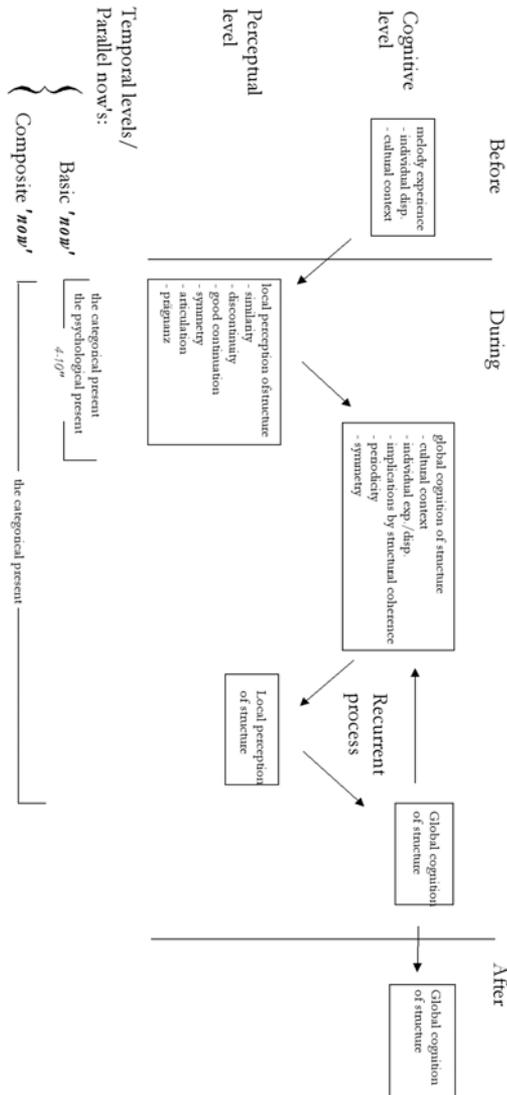


Figure 1-1. Outline of the general model of the melody cognition process, specifically regarding retrieval of melodic surface structure. “During” refers to the time during which melody sounds (or during which a melody is conceived from memory or notation.)

and cognitive selection of grouping. To this category belongs grouping by perceptual prominence/foreground and grouping by perceptual integrity/*prägnanz*.

The strengths and function of the different grouping principles are assumed to be different at different levels of melody cognition and for different kinds of structures. From the above general assumptions, I have formulated the model, which is outlined in the graphical description below: The assumption that several levels of melody structure may be conceived to interact implies that this model regards melody structure as fundamentally syntactical, that sub-structures relate to each other to form a meaningful whole. It thus implies that the cognitive process results in a mental segmentation of the melody, which may be more or less coherent.

### 1.3.2.2 Dimensions of melodic surface structure

The general model describes melodic surface structure in a limited number of dimensions. The fundamental dimension is melody structure in terms of grouping of events at different temporal levels. These are designated primary grouping levels, phrase levels and section levels. Grouping at primary level and lower phrase levels belong to what are traditionally regarded as rhythmic levels, while higher grouping levels belong to what is traditionally regarded as form.

Besides grouping structure, the model describes fundamental categorization of pitch structure in melody and metrical structure of melody. In the proposed model it is assumed that structural coherence, in particular periodicity, can imply grouping. This assumption is even more general in the model, implying that when periodicity is conceived to be a general property of the melodic structure, it becomes regulative in the sense that conceived grouping structure must be congruent with the metrical structure. This implies that metrical structure, whenever present, is conceived as the fundamental temporal grid of the music to which temporal elements and grouping is related.

Consequently, grouping structure in music with metrical structure and music without metrical structure are regarded as categorically different. Hence the model is different for metrical and non-metrical structure. The cognition of meter is assumed to be related to absolute duration in the sense that pulse perception is most salient within a certain time-span (Parncutt 1994).

Since the analysis is based on information of relative pitch and absolute duration, I have assumed that a number of properties of these parameters are cognitively significant in melody cognition in a cross-cultural perspective. Therefore, the model assumes the following main dimensions of pitch and duration perception and cognition to be generally valid in relation to melody cognition:

- *Categorical perception of pitch change in melody;*
- *Local memory of pitch in melody;*
- *Ability to perceive and conceive pitch change in melody as a continuity, as melodic contour;*
- *Ability to conceive pitch change between groups of tones, such as change of register and change of pitch set;*
- *Categorical perception of event durations (interonset intervals/ offset-to-onset intervals)*
- *Ability to conceive duration changes between groups of tones categorically, as rhythmical figures;*
- *Ability to perceive and conceive periodicity at different levels;*
- *Ability to perceive and conceive hierarchical temporal relationships and proportions.*

The model presupposes that perception and cognition of the above temporal and pitch dimensions are generally spontaneous and categorical and that the significance of different

dimensions is contextual. It is further assumed that the perception and cognition of different dimensions is basically parallel.

Parallel temporal scopes

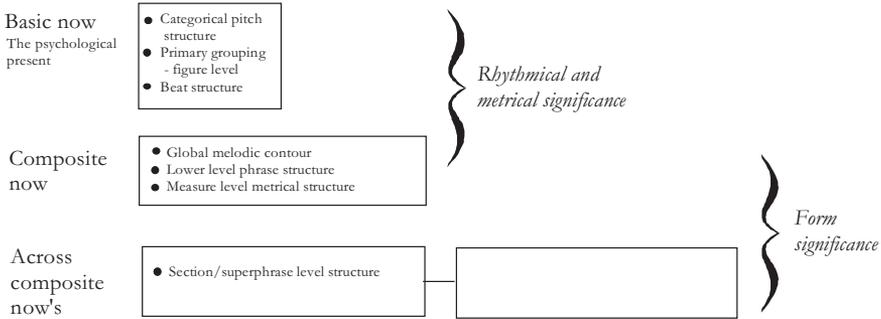


Figure 1-2. Structural dimensions and levels in relations to the proposed categories of temporal scopes. The levels of a structural dimension that can be retrieved within a temporal scope are listed within the correspondent boxes. The overlapping division between levels with rhythmic and form significance are enclosed by brackets.

### 1.3.2.3 The application of the model – the MODUS implementation

From the above general model, a method of analysis is developed, which is described in this thesis and implemented in a computer application of the method (the MODUS model) by which the method and ultimately the general model is evaluated.

Even though the model presumes that melody cognition at all levels is a parallel process, the method of analysis is a stepwise process, due to both practical circumstances and the need for evaluating the method step-wise. In this sense, the method of analysis in-directly models the processes that are assumed in the general model.

The method of analysis is outlined in the figure below:

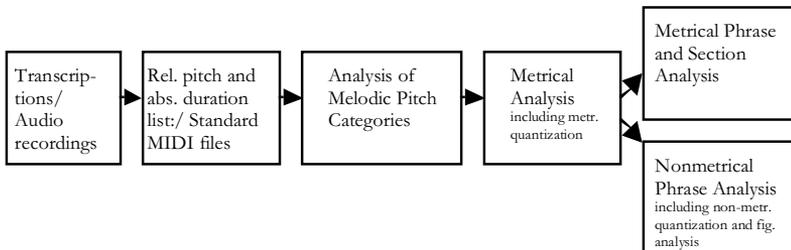


Figure 1-3. Outline of the proposed method of analysis of melodic surface structure/the MODUS implementation of the method.

The current method regards the analysis of monophonic melodies. Input to the model can be compared to the information content in a piano roll<sup>10</sup>, i.e. a list of events<sup>11</sup> with the two parameters pitch and absolute duration. In the computer model this information is retrieved from a Standard MIDI file. This means that the input can originate either from a notation of melody transcribed into a notation software<sup>12</sup> and converted to MIDI format or from the conversion of an audio recording to MIDI format.

The first step in the analysis regards analysis of *melodic pitch categories* (MPC). The Melodic Pitch categories designate the basic level of pitch change in a melody of structural significance (c.f. scale degree), without involving any aspects of tonality in the analysis. Thus it forms the basis for the analysis of significant pitch change in subsequent steps of the analysis.

The second step is the *analysis of metrical structure*. This analysis is essentially a bottom-up process, which initially involves a *metrical quantization analysis* that implies an equalization of durations from the hypothesis that all durations can be generated from a common duration value and will be cognized in terms of this standard value (c.f. score notation). This level of analysis thus regards tempo fluctuations as insignificant for the cognition of melody structure.

Further, the metrical analysis then attempts to identify a central pulse/tactus level based on the interonset structure of the melody, through a process called *beat-atom analysis*. When such a central pulse structure is not found, the analysis tries to find periodicities at any metrical levels through a *complex metrical analysis*, which also involves metrical cues given by pitch structure. If metrical structure is found on measure level but not on central pulse level, composite metrical structure is analyzed. This analysis also allows for compound meter at measure level from a common denominator duration at the level of metrical division. Only when no metrical level is identified, the method regards the melodic structure to be essentially non-metrical.

The fourth step in the analytical process is the analysis of higher levels of grouping structure. This is, for melodies with a metrical structure, performed in the *metrical phrase and section analysis*. This can in brief terms be described as a segmentation of the melody by analysis of melodic parallelism and structural discontinuity. The output of this process is a hierarchical description of melodic phrase structure involving categorization of phrases by phrase content and by syntactical relationships.

For non-metrical melodic structure (as defined by a nil result from the application of the metrical analysis) the method applies initially a *non-metrical quantization*. This is, in essence, a categorization of durations based on the sequence of durations in the melody. This is followed by a *primary grouping analysis at figure level* which involves a preliminary segmentation on phrase level. Finally, the non-metrical phrase analysis is applied to the melody resulting in an output parallel to the metrical phrase and section analysis.

The application of the method of analysis is illustrated in the figure below, which displays the performance of the different steps of the analysis on three different melodic surface structures.

---

<sup>10</sup> The analogy is borrowed from Temperley 2001

<sup>11</sup> In the current model, events may have or not have pitch. For a description of the melody representation in the implementation, see Emtell 1992 and Ahlbäck and Emtell 1993 and below

<sup>12</sup> Since the MIDI interpretation was developed using Standard MIDI Files generated by Finale 2.6.3 for Macintosh (© Coda Music Software 1987-1989) the MIDI interpretation is limited with regards to MIDI files generated by other software.

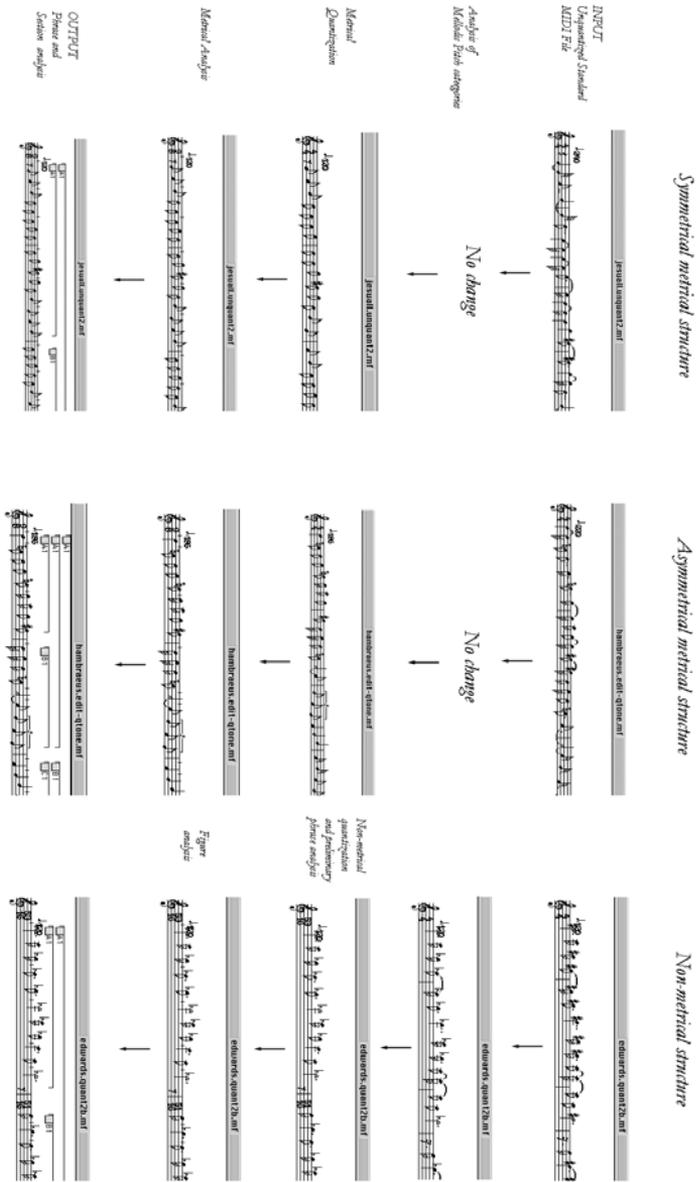


Figure 1-4. The 'MODUS' method of analysis in outline, illustrated by the application of the method of analysis on three different types of melodic surface structures. Input obtained by audio-to-midi-conversion.

### 1.3.2.4 Outline of this thesis

A brief description of the contents of each chapter is provided below:

The chapters that describe the method of analysis all have the same basic disposition. They begin with a description of the theoretical assumptions that form the basis for the particular method of analysis. This part is followed by a section in which the design of the method of analysis in question is described. In a third section, the particular method of analysis is evaluated and the results of the evaluation are discussed. Examples of the performance of the method of analysis are given in each chapter and conclusions are provided at the end of each chapter .

Chapter 2. *Related works*. This chapter contains an overview of earlier works which are related to the current model, in particular some theoretical models with a general approach. Specific attention is shown to two recent theoretical models, which involve computer implementations.

Chapter 3. *The General Model of Melody Cognition*. This chapter describes the foundations of the proposed general model, which provide the framework for the method of analysis described in subsequent chapters.

Chapter 4. *Analysis of Melodic Pitch Categories*. This chapter describes the method of analysis of melodic pitch categories. It involves initially a definition of the concept of melodic pitch category and the theoretical foundations of the method. It is followed by a description of the method design in theoretical terms and finally some examples of the performance of the method.

Chapter 5. *Metrical Analysis*. The first part of this chapter discusses general concepts of metrical structure. This is followed by a section which in order describes the *metrical quantization analysis*, the *beat-atom analysis* and the *complex metrical analysis* that together form the metrical analysis. Thereafter, the performance of the method of analysis is evaluated by examples of analyses of melodies from different styles, by comparisons of results from listener tests and comparisons with notations of metrical structure in different repertoires.

Chapter 6. *Metrical Phrase and Section Analysis*. This chapter concerns segmentation above primary grouping level in melodies with metrical structure. The cognitive foundations for grouping structure in metrical melodies are given in the first part of the chapter, followed by the description of the design of the method. This is focused on the three main elements in the phrase and section analysis: the sequence analysis, segmentation by local phrase boundaries and analysis of syntactic relationships in phrase and section structure by categorization of phrase identity. The method of metrical phrase and section analysis is evaluated by comparison with expert analyses, notations of melody structure and listener tests.

Chapter 7. *Non-metrical Structural Analysis*. The initial section in this chapter discusses the fundamental concepts regarding non-metrical rhythmic structure. This discussion involves the delineation of non-metrical and metrical structure, involving the introduction of the sub-category quasi-metrical structure. The description of the method design involves *non-metrical quantization*, *preliminary phrase analysis*, *figure level grouping analysis* and *phrase analysis* on different levels. The evaluation section involves only a minor study of non-metrical structure in one particular musical style.

### **1.3.2.5 The computer implementation – The MODUS environment**

The computer implementation has been developed in Macintosh environment and is written in the language Macintosh Common LISP versions 2.0-4.3. The graphical interface and the internal system of music representation were developed by Sven Emtell in collaboration with the author between 1992 and 1994, including a first version of the Method of Analysis developed by the author (Emtell 1992, Ahlbäck & Emtell 1993).

The development of the implementation was continued, in collaboration with Kalle Mäkilä between 1994 and 1996, and from 1996 on by myself. With the exception of the internal music representation and the graphical interface, all code of relevance for the current thesis has been written by me.

A description of the internal musical representation and the graphical interface can be found in "*Computer Aided Music Analysis*" (Emtell 1992).

*Chapter 2*

# 2 Related works

---

## 2.1 Introduction

The literature in the field of melody cognition is extensive to the degree that it is impossible to include all the work with relevance for this thesis in an overview of related research.

One might rightfully ask why yet another general model of melodic surface structure has to be created? I will try to explain the reasons behind this in the following. Therefore, the statement will concentrate on weaknesses of the reviewed theories rather than focusing on the benefits of the proposals.

I have chosen to focus on a limited number of works which are related to the current one in the sense that they propose cognitively-based theoretical models of melodic structure.

Further, I have concentrated on models that claim to be general to some extent, which fundamentally follows from a cognitive approach. Special attention is given to two quite recent proposals that include a computer implementation of the proposed theory, “*The General Computational Theory of Musical Structure*” by Emiliós Cambouropoulos (Cambouropoulos 1998) and parts of a computational theory proposed by David Temperley, presented in “*The Cognition of Basic Musical Structures*” (Temperley 2001).

These two proposals are of particular relevance since they are formalized, rule-based models and include computer implementations of the proposed theory. Computer models involving machine learning, i.e. self-organizing maps or neural networks, are not referred since they generally do not impose a fixed set of rules on different stylistic material. Such models have proved to be successful in adapting to different styles (see e.g. Krumhansl et al 2000, Hötkher et al 2002a). For the purpose of the current work, to study the degree to which the same set of cognitively based rules can apply to different musical styles, style adaptation by machine learning is not directly relevant.

Besides the two proposals mentioned above, five earlier music theoretical approaches will be reviewed in this chapter. The first of these, which involves a general theoretical approach to musical surface structure, was developed by Leonard B. Meyer and Grosvenor Cooper in “*The Rhythmic Structure of Music*” (Cooper & Meyer 1960). It is significant in this context since it was, to my knowledge, the first theory which attempted to understand melodic surface structure from concepts of gestalt psychology.

The second is *Paradigmatic Analysis* initially developed by Nicolas Ruwet (Ruwet 1966/1987) and later extended by Jean-Jacques Nattiez (Nattiez 1972, 1990). This theory proposes a approach to music analysis which is of particular interest in the current context since it attempts to analyze the syntactic structure of a musical piece from properties of surface structure, in terms of equivalence.

The third theory is probably the theoretical proposal that has been most thoroughly tested by cognitive psychologists. It was developed by Fred Lerdahl & Ray Jackendoff and presented in “*A Generative Theory of Tonal Music*” (Lerdahl & Jackendoff 1983). It proposes a

general theory of musical structure for Western tonal music, but since it is based on assumptions drawn from linguistics and cognitive psychology, it claims style-independent generality at certain analytical levels. (Lerdahl & Jackendoff 1983:278-301)

The fourth theoretical proposal reviewed here is “*The Implication-Realization model*” developed by Eugene Narmour (Narmour 1990). Like the Lerdahl & Jackendoff model, it has been tested experimentally, implemented in computer models, and is interesting in the current context since it is based on assumptions of the cognition of musical surface structure.

The fifth theoretical proposal will be reviewed primarily in the context of the evaluation of melodic phrase and section analysis. It is a style-specific theory of a particular repertoire within Norwegian folk music developed by Morten Levy (Levy 1983/1989) presented in the thesis “*The World of the Gorrlaus Slåtts*”, as such a study within the field of ethnomusicology. It does, however, propose a general theoretical model of melody structure based on concepts of linguistics and semiotics.

I will in the following briefly describe the main features of these theories with reference to the herein proposed model.

I have excluded a number of theories and formalized methods of analysis from this review of related works, although some of those will be referred to in the chapters on method. Most importantly from the perspective of this work, I have excluded models which only addresses specific aspects of the current study. In this context some of the formalized models of metrical structure should be mentioned, such as the models of metrical structure developed by Povel & Essens (1985), Lonquet-Higgins and Lee (1982), Lee (1991), Desain & Honing (1992) Large & Kohlen (1994) and Parncutt (1994).<sup>13</sup>

Even if studies in the field of musical performance have many implications for the present proposal, such studies will only be referred to in the context of the model description. For example, “*The Quantitative Rule System for Musical Expression*” developed by Frydén, Sundberg and Friberg (see e.g. Friberg 1995) can be regarded as mirroring the present work from a point of musical expression. In spite of the significance of experimental studies in cognitive psychology for the current proposal, they are not reviewed in this chapter, since they do not generally involve predictive music theoretical models based on structural features of melody.

## 2.2 Cooper & Meyer: The Rhythmic Structure of Music

This work is of great influence for the subsequent research in and theoretical models of melodic surface structure. For instance, Lerdahl & Jackendoff refer consistently to Cooper & Meyer in their book.

Cooper & Meyer propose a theory of rhythmic structure of music, including meter and form, which is based on theoretical concepts borrowed mainly from prosodic theory and Gestalt psychology. One important result of the classification of rhythmic grouping in terms

---

<sup>13</sup> See summaries in e.g. Clarke 1999 and Temperley 2001.

of traditional prosody was a description of grouping in two dimensions, accent pattern and duration pattern. They exemplified the influence of different gestalt principles on grouping and accent structure and further demonstrated how the same grouping principles may be regarded to interact on multiple hierarchical levels in Western classical music.

The limited stylistic perspective is evidently rather problematic in relation to the rather broad approach that is implied by the title. Moreover, if one assumes that the rhythmic conception of musical structure is governed by general psychological principles, it is problematic to assume that they are basically style-related; Hence they must, to some extent, be regarded as valid also for other musical styles. (See also comment about the theory of Lerdahl & Jackendoff below.)

The greatest problem with this theory in the current context is that it is not formalized in a sense that it can be used as a theoretical foundation of a quantifiable model such as the current. It does not in fact explain how accents are derived from musical structure in explicit terms; Nor does it make explicit assumptions of the strength of different grouping factors. Hence, their analytical approach has been criticized from this point of view; Clarke, in his review of research of rhythm and timing in music, commented that “*they adopted a very broad approach that did little to focus on the concept*” (Clarke 1999:478).

Even if I am rejecting the general concept of this work for the purpose of the current study, one cannot deny the importance of this work for focusing on different theoretical aspects of rhythm, meter and form structure. And especially one concept presented in their work will be used also in the current model: the division between end-accented and beginning-accented temporal organization. (Cf. Chapter 6, section 6.1.4 *Start- and end-oriented grouping*).

## 2.3 Lerdahl & Jackendoff : A Generative Theory of Musical Structure.

This most influential theory makes use of concepts derived from linguistic theory (i.e. the school of generative-transformational grammar) and Gestalt psychology. It proposes a set of principles to account for segmental structure in chiefly Western tonal music in terms of formalized Well-Formedness and Preference rules. Basically, it attempts to derive a complete deep structural analysis of a musical piece from hierarchical analysis of surface structure (including multi-part musical structure) primarily in a bottom-up manner, but also involving a top-down perspective to some extent (*e.g. prolongational reduction*). Thus, it involves aspects of melodic structure that are not addressed in the current model, such as harmonic and melodic tension and relaxation.

Although it is primarily designed to reflect the intuitions of a listener within the idiom of Western classical music, the theoretical approach is essentially cognitive. This implies that the general principles can be assumed to apply for other musical styles as well. Lerdahl & Jackendoff do actually address this question discussing the levels at which their theory can be regarded to be style-independent (Lerdahl & Jackendoff 1983:278-301).

In the early 1980's I developed a model of temporal organization in music (see e.g. Ahlbäck 1983/1986/1996, 1986, 1989) which described the temporal cognition of music in basically two dimensions, gestalt and periodicity cognition, termed rhythm/grouping structure and metrical structure. These were regarded as two categorically different aspects of temporal

experience/organization.<sup>14</sup> It was not until in the early 1990's that I was happy to discover that a view similar to mine had already been proposed by Lerdahl & Jackendoff. The theoretical distinction between grouping and meter made by Lerdahl & Jackendoff seems today to be widely accepted and is appointed a significant contribution to the understanding of temporal structure in music (Clarke 1999:478).

Of specific interest in the context of the current model is the grouping preference rules proposed by Lerdahl & Jackendoff. These rules predict how different structural grouping principles interact in the formation of grouping structure. These rules have been subject to empirical tests (e.g. Deliège 1987), which have indicated their perceptual validity.

The model proposed by Lerdahl & Jackendoff has many implications for the current model that will be referred to in the context of the methods. There are, however, certain general problems regarding their model which makes it problematic as a general framework for the current proposal:

- Their model assumes musical structure to be completely coherent at all levels as a result of the application of the tree model of syntactic relationships, thus implying syntactic relationships at deep structural levels which may not always be cognitively relevant and hence neglecting other structural principles. (See e.g. Cook 1990:68)
- Their model does not, in any explicit manner, handle grouping implication by similarity (musical parallelism) and its relationship to grouping by discontinuity, which is of crucial importance for the cognition of grouping structure. (See e.g. Lerdahl & Jackendoff 1983:53.)
- Their model does not handle interaction between grouping and meter in a complete sense. In particular, it does not address how grouping can influence metrical conception and vice versa. This results, for instance, in a set of rules for metrical structure which cannot account for metrical conception in music without durational contrast or phenomenal accentuation.
- The general preference approach is problematic from a general cognitive point of view, since it implies a selection process governed by reflection upon different alternative interpretations.
- The "*Generative Theory of Tonal Music*" is, in spite of the general approach, developed with reference to the Western classical repertoire. The inclusion of style-specific traits defined in cognitive terms makes it problematic in relation to other musical styles; If one, for instance, assumes meter conception to be dependent upon perfect periodicity in one style, this has to hold for other styles as well.<sup>15</sup>

---

<sup>14</sup> This categorization of temporal experience been criticized by e.g. Blom 1989:38 "This example demonstrates that a division between cognition of temporal gestalts (rhythms/grouping) and periodicity (meter) is inadequate." (my translation).

<sup>15</sup> This problem is addressed by Lerdahl & Jackendoff, who (referring to the existence of aperiodicity at tactus level in e.g. Balkan music) suggest that "These differences in rules represent what one must learn about an idiom to become an experienced listener" (Lerdahl & Jackendoff 1983:99). If this is true, the proposed rules (MWFR 3 and 4) must be a special case of a more general principle, which is not described by L & J.

## 2.4 Narmour: The Implication-Realization Model

The great benefit of the implication-realization model developed by Narmour (Narmour 1990, 1992) is that it takes the cognitive influence of expectation into account. Thus it regards the cognition of musical structure as a dynamic process in time.

It is especially interesting with reference to the current model since it is centered on the analysis of melody.

The model claims to be a general model of human conception of melodic structure to a degree that is quite radical;

*“The theory will analyze (and thus partly explain) all melodies ever written or to be written regardless of stylistic origin” (Narmour 1992:7)”*

*“...the theoretical constants invoked herein are context free and thus apply to all styles of melody” (Narmour 1990:1)*

Narmour attempts to achieve this by taking a similar perspective as in the current model, making some fundamental assumptions about the nature of melody cognition. He assumes that expectations are governed by two principles that are assumed to be universal: (1) Similarity implicates continued similarity (non-closure); (2) Difference implicates further difference. Similarity is determined by a syntactic, parametric scale which is assumed to be a style-independent constant.

The general concept of this theory is that bottom-up processes, which are assumed to be style-independent, innate and universal, interact with top-down style-dependent processes in the cognition of melody structure. It is assumed that higher order structure is retrieved basically by combinations of lower level structures but that top-down processing influences this process.

The core of the theory is that postulated structural implications of continuation on a note-to-note level, derived from Gestalt principles, creates non-closure when they are realized and closure when they are not realized through the succeeding note (Narmour 1990:3ff). Like the theory of Lerdahl & Jackendoff, this theory has also been subject to evaluation by experimental studies, which have to a certain degree supported the generality of these rules (e.g. Krumhansl 1995, 1997; Krumhansl et al 2000; Thompson et al 1997, 1998; Schellenberg 1996,1997).<sup>16</sup>

The very general approach of this theoretical model is obviously relevant from the point of view of the current study. Further, the so called *primitives* of low-level structure can be said to mirror the local discontinuity rules which are proposed in the current model. There are, however, some important problems with the approach, that make it unsuitable as a fundament for the current model.

- Narmour’s model implies causality and structural coherence on a note-to-note level. This implies in turn that melodic structure is generated by an act of creation that is controlled at a note-to-note level, hence predictable on that level. Thus it rejects the cognitive process of deducing the melodic structure from a series of pitch events in time. Hence,

---

<sup>16</sup> The cross-cultural tests of the model performed by Krumhansl et al have also extended the model, adding rules and are highly significant in present context.

the process of deducing higher levels from lower levels becomes highly complex, which does not seem to be valid from a cognitive perspective. (Cf. Narmour 1999:445-457).

- The problem of assumed causality, which originates from the concept of implication-realization, is inherent in the fundamental assumption by which the implication rules are generated. It states that similarity implicates continuous similarity and that difference implicates further difference, which is demonstrated to be invalid regarded as general principles in the current work (see Chapter 6, *Metrical phrase and section analysis* sections 6.3).
- The model does not quantify the level of similarity between groups of events at which similarity is structurally significant with regards to context.
- The model does not explain explicitly in a formal manner how rhythmical and metrical relationships are involved in the cognition of melody.<sup>17</sup>

In summary; the general concept of implication-realization, however useful in many respects, neglects certain aspects of melody cognition which make it inadequate as a foundation of the current model.

## 2.5 Ruwet and Nattiez: Paradigmatic Analysis

*Paradigmatic Analysis* was developed in the 1960s by Ruwet (Ruwet 1966, 1987) with inspiration from structural linguistics method applied to poetry. As is evident from the present overview, the influence of structural linguistics has been profound in the development of general methods of music analysis<sup>18</sup>. It is based on the assumption musical structure is determined by the cognition of equivalent units, defined by repetition, which includes varied repetition and transformation. Anything, which can be regarded as repeated in a broad sense, is defined as a unit, and this is true for all structural levels.

Middleton (1990:183) describes the benefits of *Paradigmatic Analysis* thus; “*In principle [Ruwet] can segment a piece without reference to its meaning, purely on the basis of the internal grammar of its expression plane*”. But having said that he also addresses the problems generally attributed to the method;

*“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. [...] there is no attempt, as there is in the somewhat analogous methods of distributional linguistics, to delineate the make-up of a system – though segmentation and unit relationships, that would constitute the beginnings of a description of a code.” (Middleton 1990:183)*

Thus, since *Paradigmatic Analysis* is not a formalized model that makes precise interpretations of phenomenal melodic structure, it cannot make the fundament of a general model such as the one proposed in this thesis. Moreover, it neglects factors of continuity and discontinuity (addressed in e.g. the aforementioned theories by Lerdahl & Jackendoff and Narmour) as means of structural determination. In doing so, it implies an evidently style-

---

<sup>17</sup> See also comment by Cambouropoulos (Cambouropoulos 1998:13) and Cross (Cross 1995:502)

<sup>18</sup> See also Middleton 1990:177

dependent assumption, namely that all musical structure is founded on repetition and transformation principles. That this assumption cannot be regarded as general is evident from the results presented e.g. in the current thesis.<sup>19</sup>

But in spite of these flaws, the important contribution of the model is to indicate the possibility and need for a style-independent top-down segmentation, which is not in focus in any of the aforementioned models.

## 2.6 Levy: “The World of the Gorrlaus Slåtts”

In an extensive doctoral thesis, the Danish ethnomusicologist Morten Levy proposed a general theory of musical structure, based on the analysis of a tune-family within the repertoire of Norwegian fiddle music. As such, this theory makes general assumptions about musical structure from a very limited musical material, which is obviously questionable. However, the particular repertoire has an extremely complex surface structure and Levy attempts to approach it in a very general way, building his theory from concepts of mainly linguistics and semiotics.<sup>20</sup>

The method of Levy is directed to the retrieval of deep structures from surface structure. The nature of the repertoire in focus gives rise to very different structural entities than what is addressed in the models that are focused on the structure of Western Classical music. The form principle defined by Levy is designated ‘cyclic form’, which is a well-known term within ethnomusicological theory. Further, the analysis of pertinent deep structures involves reduction of surface structure not assuming syntactical coherence and allows for variance of surface structures within an identified deep structural entity.

Most central to the method is the reliance on performance cues to the analysis of e.g. accent structure. But, in other respects, the analysis is based entirely on Levy’s intuitions, and though he offers a formalized method for reduction of pertinent events, it is not used consistently throughout the work. The analysis of pitch structures, which are essential to the retrieval of deep structures, is not formalized in a way that it can be repeated from the description in Levy’s thesis.

Moreover, it does not attempt to address questions of generality in relation to other musical styles, which makes it inadequate as a foundation for the current model. Still, this model raises many important questions about the generality of other models by adopting a totally different perspective.<sup>21</sup>

---

<sup>19</sup> See e.g. the example segmentation of “Raag Suddha Danyasi” (chapter 6, *Metrical Phrase and Section Analysis*, section 6.3.6.1)

<sup>20</sup> Similar attempts to apply theoretical concepts from linguistics and symbolic logic from ethnomusicological perspectives are offered by e.g. Powers (Powers 1980) and Seeger (Seeger 1980)

<sup>21</sup> See further discussion in Chapter 6, *Metrical Phrase and Section Analysis*, section 6.3.4.3

## 2.7 Cambouropoulos: “Towards a General Computational Theory of Musical Structure”

In his thesis Cambouropoulos proposes a general theory of musical structure, which claims to be of general significance regarding analysis of musical surface structure:

*“The General Computational Theory of Musical Structure (GCTMS) may be employed to obtain a structural description (or set of descriptions) of a musical surface. This theory is independent of any specific musical style or idiom, and can be applied to any musical surface” (Cambouropoulos 1998:2)*

This model thus has the same claim of generality as e.g. the model proposed by Narmour. Cambouropoulos describes the output analyses of the model as related to

*“...the intuitive ‘understanding’ a listener has when repeatedly exposed to a specific musical work (the listener need not to be familiar with the particular style of idiom the work belongs to).”*

This view of the listening process as a learning process is also adopted in the current model.

Cambouropoulos model is essentially a method of computer-assisted analysis of musical surface structure and as such it takes both the approach of computer science and cognitive music theory. It is not explicitly based on a general model of cognition of musical surface structure but is indirectly implying such a model. The method is outlined in figure 1.

Since Cambouropoulos method of analysis is, by far, the most similar method of analysis to the one proposed in this thesis I will describe it somewhat more explicitly than the above models.

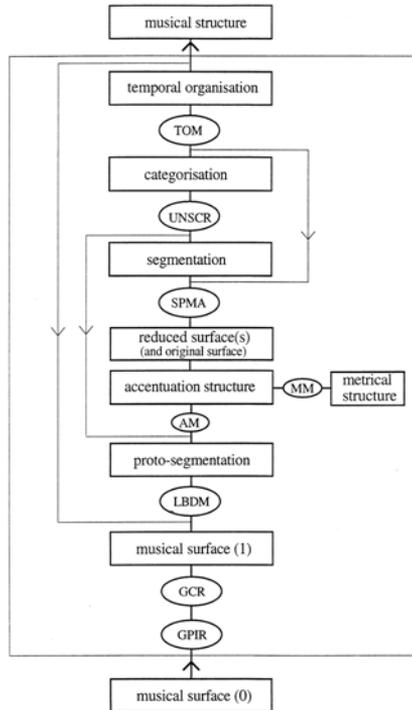
The fundamental concepts of the GCTMS model are basically drawn from cognitive psychology and logic and are described as *“a set of principles which are assumed to be part of the way in which a human makes sense of the world”*. (Cambouropoulos 1998:26). Those principles are: the *identity-difference* principle; the *economy* principle (the preference for reduction of information); the *informativeness* principle (the preference for abstractions that enables a human to achieve desired goals); and the *exposure priming effect* (that salience relates to exposure).

As can be seen in the overview (figure 1) displayed above, the GCTMS model involves a model of General Pitch Interval Representation, which basically matches the sequential intervals of a melody against an interval representation schema derived from frequency of occurrence of pitches within an interval set and a set of given scale structures

Even if it is generally style-independent, it does not allow for interval differences below semitone level. Most importantly, it requires manual input, hence a pre-analysis, of the scale structure.

The model also involves a proposal for a general chord representation algorithm, a local boundary detection model, which derives grouping in a manner similar to the model proposed by Lerdahl & Jackendoff. Further, it involves a metrical analysis module in which metrical accent structure is based on duration and pitch discontinuity. Unique to this model, in relation to most other earlier models, is the analysis of melodic surface structure by similarity. This module is based on detection of note-to-note similarity, which is reduced basically by the principle of frequency of occurrence of similar patterns to select pertinent patterns of structural significance. Based on a weighed analysis of influence of local boundary detection

and grouping obtained by similarity analysis, a lower level phrase structure is obtained. These phrases are categorized by the *unscramble algorithm*, which from a similarity rating labels the phrases according to the principle of least category overlap.



TOM: Temporal Organisation Model  
 UNSCR: UNSCRAMBLE Algorithm (Categorisation Model)  
 SPMA: Sequential Pattern-Matching Algorithm & Selection Function  
 MM: Metrical Matching  
 AM: Accentuation Model (event salience)  
 LBDM: Local Boundary Detection Model  
 GCR: General Chord Representation  
 GPIR: General Pitch Interval Representation

Figure 2-1. Overview of the General Computational Theory of Musical Structure (Cambouropoulos 1998:29)

The GCTMS

model is attractive in many respects. It is based on a number of well-defined formalized rules, derived from cognitively based assumptions of the cognition of musical surface structure. It is especially interesting in the current context since it is mainly addressing problems regarding melodic structure. It offers a rather sufficient number of levels of analysis. Further, it addresses many important issues in the field of cognitively based music theory and a number of interesting approaches to different problems within the field.

There are, however, some general problems with regards to the current context.

- It is not based on an explicit model of cognition of musical surface structure. Hence, the interrelationships between different analytical levels and modules of the analytical model are not addressed in an explicit way.
- Some parts of the computer model were not implemented, raising questions about the performance of the model and the plausibility of the results.
- Hierarchical structure is derived entirely through a bottom-up process, which is not described clearly in the thesis. This proposed process will ultimately lead to inconsistent results at higher hierarchical levels, according to the results of the method of analysis proposed in the current thesis. Thus the output of the model seems to be in essence non-hierarchical.
- Cambouropoulos explicitly assumes 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.”* (Cambouropoulos 1998:9). This assumption makes it impossible to analyze structures, which are determined by categorical similarity (as will be shown in subsequent chapters). This is especially problematic regarding the determination of hierarchical levels, but with regard to the metrical analysis.

Thus, in spite of its many benefits, the model cannot function as the fundament of the current model.

## 2.8 Temperley: The Cognition of Basic Musical Structures

Like the work by Cambouropoulos, the book “The Cognition of Basic Musical Structures” (Temperley 2001) can be regarded as a major contribution in the field of cognitively based music theory, specifically regarding computer-assisted research. They share both the computational approach to musical structure and the focus on musical surface structure, although the models obviously have been developed independently of one another.

Temperley’s work is, however, explicitly based on the *“A Generative Theory of Tonal Music”* by Lerdahl & Jackendoff. In brief terms, it is an implementation of the preference and well-formedness rules proposed in this theory into a computer model. However, this implementation involves a number of contributions to the theoretical model by Lerdahl & Jackendoff, and in some cases, different interpretations and analytical concepts. Moreover, it involves evaluations of the success of the model with respect to earlier proposed models. Even if it is, like Lerdahl & Jackendoffs’ theory, focused on what Temperley designates the common-practice repertoire belonging to the Western classical music idiom, the application of the model to different musical styles is of particular interest in the current context. This includes e.g. Rock music and African traditional music.

It has a considerably wider scope than the present study and provides computer implementations of contrapuntal structure, harmonic structure and key structure, in addition to computer models of metrical structure, melodic phrase structure and pitch spelling which applies also to the current proposal.

A general difference is that this model allows polyphonic input, which is not in question in the current work.<sup>22</sup>

Regarding the parts of Temperley's work which are related to the current work, the proposed models are relatively undeveloped. The metrical analysis regards basically only interonset information, not taking pitch information into account. This may be a useful approach regarding polyphonic music, but in monophonic music it is not sufficient, which is demonstrated in Chapter 5, *Metrical Analysis*. In a more recent article (Temperley & Bartlette 2002), this question is addressed, specifically concerning melodic parallelism. This addition to the model seems perhaps, to be a fruitful, but still rather undeveloped approach which does not impose any general theory that can account for structural implication of melodic parallelism.

The metrical analysis implementation becomes highly style-dependent, since it assumes only perfect periodicity at tactus level.

The proposed model of analysis of melodic structure seems also to be rather undeveloped. It does actually not make use of e.g. melodic parallelism and does not, in fact, imply any analysis of structural implication by pitch structure. Temperley explicitly states that;

*“As observed in chapter 2, recognizing motivic parallelism – repeated melodic patterns – is a major and difficult problem which is beyond our scope”.*

Moreover, it requires a prior metrical analysis that in the tests of the model reported by Temperley were provided by manual input. This is a problematic circumstance because metrical parallelism is an important rule within the model. The results of the application of current model will be viewed in relation to results from a comparative study involving Temperley's model of melodic analysis in Chapter 6 *“Metrical Phrase and Section Analysis”*, Section 6.3.8

To summarize:

- Temperley's approach does not involve a specific theory which can form as the basis of the current model.
- The computer-assisted analytical methods proposed by Temperley are too limited to form the basis of the herein proposed method of analysis.

---

<sup>22</sup> The current model can, however, be extended to address polyphonic structure.

*Chapter 3*

# 3 The General Model

---

## 3.1 Fundamental assumptions

The general model of cognition of melodic surface structure proposed herein is based on a limited number of fundamental assumptions. These can be divided into two categories; assumptions concerning general cognitive capacities and processes and assumptions specifically concerning the determination of melodic structure. These assumptions can be regarded the axioms on which the general model and the method of analysis rely.

The core of the general assumptions is the concept of *the perceptual (or psychological) present* (Michon 1978, Fraisse 1982, Clarke 1999) which, in this context, is to be understood as the basic temporal scope of attention for the immediate perception of melodic structure at primary levels, and what I have termed *the categorical present*, the category capacity of short term memory (Miller 1956).

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 time-span, which is here termed *the perceptual present*, is generally estimated to between 3-8 seconds (Clarke 1999:476). Most frequently, a value around 5 seconds is mentioned (Fraisse 1982). Considering this evidence, it seems reasonable to assume this time-span to be the fundamental scope of attention in melody cognition. Asserting this, it follows that the cognition of melodic structure must be based on perception of structure at this level.

*The perceptual present* is generally associated with capacity of short term memory; For instance Clarke (1999:476) suggests that "...it looks very much as though the perceptual present should be understood as the temporal view of the contents of working memory (Baddeley, 1986) with all the properties that have been described in this area." In a classic article Miller estimated the number of categories that can be simultaneously addressed, drawing from test of how accurately people can assign numbers to the magnitudes of various aspects of a stimulus ("*The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information*" Miller 1956). His conclusion, based on results of experiments on pitch, loudness, taste intensities, visual positioning, color perception etc., interpreted in terms of information processing, is that within '*the span of immediate memory*', we are able to separate  $7 \pm 2$  categories of a unidimensional stimuli: "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. This process – the structuring – is herein assumed to be governed by three general principles; (1) The principle of *reduction of information*, i.e. that we are consistently and spontaneously trying to reduce the amount of information that needs to be processed; (2) The principle of *coherence*, i.e. that we seek for general principles by which the

local information can be generated; (3) The principle of *significance*, i.e. that we consistently are looking for significance, meaning or a message in the structure.

The general means of structuring are assumed to be reduction of information by the interrelated processes of *categorization* and *filtering*; the first referring to the reduction of information content by similarity and discrepancy; the second referring to the reduction of information content by the selection of pertinent information.

Besides the limitations of categorization capacity given by *the categorical present*, I also assume that there are specific limitations in our capacity of estimation that can be described in categorical terms. One such limitation, which is significant in the current context, is the capacity to estimate temporal proportions. The assumed human capacity in this respect is depicted by *the categorical proportion rule*, which states that estimation of unidimensional temporal relationships is categorical and is maximized to conception of four-division. In other words: the maximum division we can immediately conceive or recognize, without regarding the division as composed by subcategories, is restricted to division in four, with regards to temporal relationships between atomic units. The evidence from the psychological literature points rather to triple division as the upper limit (see e.g. Fraisse 1956, Povel 1981), but the form principles identified in the development of the current work (see Chapter 6, section 6.2.4.4) have lead to the assumption that also '*fourfoldness*' can be conceived.<sup>23</sup> This rule is thought to apply to all levels of temporal relationships both in hierarchical and sequential terms.

It is further evident that cognitive capacity is restricted by biologically determined limitations of the auditory system. It is herein suggested that limitations regarding pitch- and interonset-interval discrimination along these dimensions influence salience and categorization of intervals. For example, since subjective rhythmization and categorical durational discrimination seems to be difficult for interonset intervals below 100 ms (Fraisse 1982, Monohan and Hirsch 1990) there is reason to believe that there is a lower limit at which durations are structurally significant. On the basis of studies of pulse salience (Parncutt 1994) and considerations based on the categorical present, it is assumed in this model, that there is an absolute range within which *referential durational intervals* generally occur. This ranges from about 200 to 1800 ms, with a peak around 500-700 ms. (c.f. London 2002:546). Correspondingly pitch differences below 30 cents are regarded structurally insignificant.

The capacity of temporal attendance and categorization depicted by the perceptual and categorical presents is, however, assumed to be extendable through the cognitive process of reduction of temporal and categorical presents into units. This is the theoretical basis for the assumption of *hierarchival structuring*, reflected e.g. in the assumption that higher temporal scopes of attention may be formed by categorical reduction of lower-level scopes of attention, for which I have proposed the term *parallel now's*. There is, however, also in this case assumed absolute limits that are based mainly on evidence derived from properties of different musical styles. For example, I assume that regulative periodical patterns cannot consist of more than 48 sub-elements at the central regulative level, based on the evidence of extensive regulative periodical cycles, e.g. cycles of forty beats in Indian music. (see Chapter 5, *Metrical analysis*).

---

<sup>23</sup> The argument is briefly, that a fourfold division of equal units, can be perceived as a whole since it do not form a discrete duple division and is below the level of the categorical present. Thus it cannot be regarded a continuum.

Cognition of structural hierarchy is assumed not to be limited to scope of temporal attendance, but to the rule of categorical capacity (*the categorical present*). This implies that the  $7 \pm 2$  rule of simultaneous categorization is assumed to apply also to simultaneous hierarchical levels at a given structural dimension. The complexity of structural hierarchy is assumed to be related to the level of structural coherence and temporal symmetry according to the following rule; the more coherent and symmetrical a structure is, the more structural levels can be cognized; And conversely, the more chaotic a structure is the less hierarchical levels can be conceived. On the other hand cognition of structural hierarchy is also assumed to be depending on the level of structural integrity of substructures. In general terms, this implies that the salience of a structural level relates to the number of unique and discrete substructures it can be divided into; the greater the number of discrete substructures the less salient will the structural level become. This is in the current model regarded a structural quality, for which I have suggested the term "*Latticity*", referring to the degree to which a structure can be divided into discrete substructures. (see Chapter 6, *Metrical phrase and section analysis*)

## 3.2 Principles of melodic structure

Besides these basically contextual assumptions I am proposing a model of how general grouping principles, defined in terms of gestalt psychological laws, are active in melodic structuring. This model includes the relative prominence and interaction of these principles on different levels of melody cognition.

The core of this model is the classification of the different grouping principles by the role they play in the cognition of grouping structure in melody. It postulates that groups are primarily formed by the categorization into same and different, i.e. grouping by similarity and discontinuity, designated *primary grouping principles*. It further postulates that groups can be formed by implication based on properties of grouping determined by primary grouping principles. These *secondary grouping principles* refers to grouping by good continuation and symmetry. The third class, *ternary grouping principles*, are defined as involved in the perceptual and cognitive selection of grouping. To this category belongs grouping by perceptual prominence/foreground and grouping by perceptual integrity/*prägnanz*.

### Primary grouping principles

*Grouping by similarity, continuity and proximity – 'sameness':*

*Similar and close events tend to be grouped together.*

*Grouping by discontinuity, dissimilarity and distance – 'difference'. Change / difference indicates group boundaries.*

### Secondary grouping principles

*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.*

*Grouping by symmetry. Symmetrical grouping is more prominent than asymmetrical grouping and can be implicative in the case of grouping by hierarchical symmetry.*

### Tertiary grouping principles

*Grouping by perceptual prominence / foreground, i.e. articulation, impulse and gravity: Grouping between most articulated events is prominent.*

*Grouping by integrity/ 'prägnanz' / contrast:*

*Discrete grouping in the sense that the content of the group contrasts to the group boundaries is prominent.*

The strengths and function of the different grouping principles are assumed to be different at different levels of melody cognition and for different kinds of structures.

From the above general assumptions I have formulated the model, which is outlined in Chapter 1, *Introduction*, section 1.3.2.

This model implies that a listener enters the process of cognition of a melody with preconceptions about the structure of a melody mainly originating from prior experience of melody, as determined by socio-cultural and individual limitations. The expectations that arise from these preconceptions are tested against the spontaneous structuring of the musical sound at the perceptual level, within the limits of the perceptual present, and determined by the above listed grouping principles.

This process leads to a global conception of the melody structure at all structural levels, governed by spontaneous reduction of information density by generalization. This conception gives rise to expectations of melody structure that may include implicated structure based on structural coherence, such as periodicity and symmetry.

These expectations are in turn tested against new information retrieved from perception at the local level, which gives rise to new conceptions about the global structure of the melody as well as the local structure. Global structures are thus formed by generalization of lower level structures at different temporal scopes of attendance. These can according to the model be conceived as interacting, forming a structural hierarchy. Most important, this continuous revision of the global and local conception of the melody by means of new information retrieval works retrospectively, thus revising previous conceptions about local structure and global structure.

The cognitive process is assumed to typically continue also after the sound of the melody has stopped.

The assumption that several levels of melody structure may be conceived to interact implies that this model regards melodic structure as fundamentally syntactical, that sub-structures relate to each other to form a meaningful whole. It thus implies, that the cognitive process results in a mental segmentation of the melody, which may be more or less coherent. The level of structural hierarchy and coherence is, however, related to the above noted structural qualities of a certain level.

A fundamental assumption inherent in the model is the view of melody cognition as a dynamic learning process. Repeated exposure to the same melodic stimuli is considered to result in a continuation of the cognitive process, allowing for further revision of the mental image of the melodic structure guided by the aforementioned principles of reduction of information, coherence and significance.

### **3.3 Dimensions of melodic surface structure**

The general model describes melodic surface structure in a limited number of dimensions. One fundamental level is as have been described above melodic structure in terms of grouping of events at different temporal levels. These are designated primary grouping levels, phrase levels and section levels.

Grouping at primary level and lower phrase levels belong to what are traditionally regarded as rhythmic levels, while higher grouping levels belong to what is traditionally regarded as form. The model does not, however, imply any distinct division between rhythmic levels and form levels, although it is possible to derive such a distinction from the model based on the assumed maximum scope of a parallel now and the minimum complexity of a structure defined by melodic content. There is psychological evidence indicating that such a distinction can be made (see e.g. Clarke 1999:476). The model does not either imply any distinct division between phrase and section levels, even though such a division is also possible to make from the assumptions of complexity of melody structures.

Besides grouping structure there can be yet another dimension of temporal structure in melody, namely meter. The two dimensions are in the current model assumed to be representing two distinct cognitive categories: temporal *Gestalt conception* (Rhythmic grouping/gestural rhythm) and conception of regulative periodicity (*meter*). (Ahlbäck 1983/1986/1989/1996).

Meter can according to this view be regarded an aspect of the gestalt psychological concept of streaming, conceived as a stream of temporal points. It is herein suggested that *meter* and *gestalt cognition* can be conceived independently of one another, i.e. meter without grouping and grouping without meter. (see further Chapter 5, *Metrical structure*, section 5.1). While the conception of *meter* without grouping in its purest form is assumed to be relatively rare due to spontaneous grouping of temporal events, conception of grouping without *meter* is assumed to be a more common feature in the cognition of melodic structure in different musical styles.

When grouping and *meter* coexist, which seems to be most common, it is assumed in this model that grouping and *meter* interacts; that *meter* at pulse level forms the basic temporal grid to which grouping is related; and that regularity in grouping implies metrical conception. (e.g. Ahlbäck 1996, c.f. Cambouropoulos 1998:65)

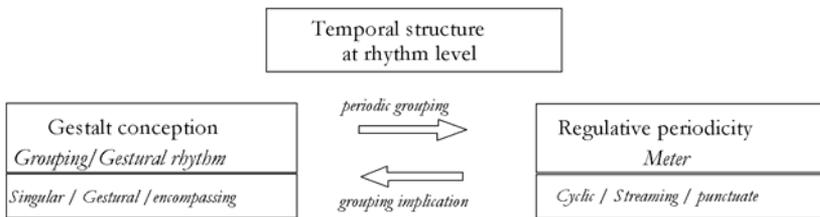


Figure 3-1. General model of temporal structure at rhythm level in music. Typical qualities of grouping and meter are noted. The arrows indicate that periodicity can be derived from grouping and grouping can be implicated by periodicity.

It is further suggested that conception of meter and grouping is spontaneous and primarily derived from perception of structural properties within a perceptual present. This implies that when phenomenal structural periodicity exists, listeners are assumed to recognize

it and derive a regulative temporal grid (*meter*) from this structure. Similarly, when phenomenal temporal or pitch differentiation exists, listeners are assumed to conceive a grouping structure.

Though grouping can apply to all levels of musical surface structure, grouping within the scopes of temporal attendance (the parallel now's of rhythmic significance) is here assumed to involve conception of temporal relationships of implicative power. (c.f. Clarke 1999:476)

This model implies that when periodicity is conceived to be a general property of the melodic structure, it becomes regulative in the sense that conceived grouping structure must be congruent with the metrical structure. This implies that metrical structure, whenever present, is conceived as the fundamental temporal grid of the music to which temporal elements and grouping is related. Consequently, grouping structure in music with metrical structure and music without metrical structure are regarded as fundamentally different.

The influence of the pitch dimension on melodic structure in relation to temporal relationships, is generally assumed to increase with the length, in both categorical and temporal terms. This is based on the assumption that pitch can address more unique events in time than durational relationships. This is in turn due to the two-or more-dimensionality of pitch.

In the current model, influence of the tonality dimension of pitch is disregarded, for several reasons. Most important, it is herein suggested that the interrelationship between melodic surface structure and tonality cognition is asymmetrical; this means that tonality is influenced by melodic surface structure on local level while the opposite is not true on a style-independent level.<sup>24</sup> This is because tonality cognition, since it involves the cognition of hierarchical relationships, is essentially a cognitive phenomenon, derived from the context - a quality of the context. Thus tonality does not explicitly and primarily imply structure on a style-independent level, as is e.g. interonset relationships or pitch distance.

Once a tonality is recognized, it can be structurally determinative, as is evident from the role of e.g. harmony in Western classical and popular music. The reason why the model disregards tonality, as a determinant of melodic structure, is that it is developed to be a vehicle for analysis of tonality. To address the influence of melodic surface structure on tonality cognition, the model cannot include tonality as a determinant of melodic surface structure. This is, however, not a basis for rejecting the generality of the proposed general model here suggested, since the influence of tonality is addressed.

The fundamental dimensions of pitch that are regarded in the model are pitch height and primitives of pitch consonance, basically octave equivalence and in certain contexts also fifth consonance. Fundamental to the model is the assumption is that pitch change in melody is generally categorized as changes between points along the pitch height dimension. Thus it does not represent continuous pitch change as a unidimensional stream, but as continuous change between pitch categories. Further, it is here suggested that categorization of pitch in melody involves three dimensions (when omitting the tonality dimension): (1) Categorization of pitch, based on sequential pitch change in melody, properties of preferred pitch category sets and primitives of pitch consonance; (2) Categorization of pitch based on local pitch memory; and (3) Categorization of pitch based on octave equivalence.

---

<sup>24</sup> On a style-dependent (intra-opus) level this is not true, however, since a listener can use expectations of tonality implications on lower-level grouping.

The first of these categorizations is derived from the assumption that the pitch change pattern within the melody determines the categorization. This category, here termed melodic pitch category, is regarded the most significant structural level of pitch change in melody, of greater structural significance than categorization based on perfect pitch similarity and octave equivalence. (See chapter 3, *Analysis of Melodic Pitch Categories*).

Thus, the model assumes the following main dimensions of pitch and duration perception and cognition to be generally valid in relation to melody cognition;

- *Categorical perception of pitch change in melody*
- *Local memory of pitch in melody*
- *Ability to perceive and conceive pitch change in melody as a continuity, as melodic contour*
- *Ability to conceive pitch change between groups of tones, such as change of register and change of pitch set*
- *Categorical perception of interonset intervals<sup>25</sup> (IOI/OOI)*
- *Ability to conceive duration changes between groups of tones categorically, as rhythmical figures*
- *Ability to perceive and conceive periodicity at different levels.*
- *Ability to perceive and conceive hierarchical temporal relationships and proportions*

The model presupposes that perception and cognition of the above temporal and pitch dimensions are generally spontaneous and categorical and that the significance of different dimensions is contextual. It is further assumed that the perception and cognition of different dimensions is basically parallel. The relationship between parallel temporal scopes is outlined in Chapter 1, *Introduction*, section 1.3.2.

---

<sup>25</sup> The interonset interval (IOI) is the time-span between one onset and the next onset. OOI refers to the time-span of an offset to the next onset, rest.

*Chapter 4*

# 4 Analysis of Melodic Pitch Categories

---

## 4.1 The concept of melodic pitch categories

*“How is it that when we hear a melody, for example, one consisting of six tones, and then the same melody transposed to a new key, that we recognize it as the same, even though the sum of the elements is different?” (Wertheimer 1938, quoted after Krumhansl 1990:139)*

This chapter concerns the structurally significant levels of pitch change in melody. This is obviously of crucial significance for the determination of melodic surface structure, which is why pitch categorization is the initial step in the method of analysis. This question is generally addressed in terms of *pitch spelling* (see e.g. Temperley 2001, Cambouropoulos 1998, 2003) and is often a more-or-less developed feature within music notation software. In the following it will be argued, however, that this is not primarily a matter of notation practice, but a matter of melody cognition.

There is converging evidence that pitch in melody is, in general, perceived categorically in several respects (Krumhansl 1990:271 ff). The most fundamental of these has been already noted in the previous chapter, namely that pitch change in melody is generally conceived as change between more or less stable pitch categories along the pitch height dimension<sup>26</sup>. Whether this holds for all tonal Gestalts that can be labeled melodies and all individuals can be questioned<sup>27</sup>. It is, however, evident that melodies in the vast majority of different musical cultures around the world exhibit what can be perceived as stepwise motion between relatively stable pitches<sup>28</sup> and verbal concepts that indicate that this treat of the phenomenal pitch structure is a cognitive reality (see Dowling & Harwood 1986, Krumhansl 1990). Common Western notation as well as other notation systems supports this notion, since pitch is represented categorically.

It is herein suggested that fundamental pitch categorization in melody basically emerges from conception of melodic structure, through melodic motion along the pitch height dimension. Categorical pitch perception is, in this sense, a property of melodic cognition. This suggestion implies that melodic pitch categorization in this basic sense does not emerge as a property of tonality cognition, which is mostly an inherent assumption in e.g. applied Western music theory or in pitch spelling algorithms. Rather, fundamental pitch categorization in

---

<sup>26</sup> In some cultures (see Kubik 1970) this dimension is associated with and expressed in terms of size or age.

<sup>27</sup> Phenomenal pitch structures characterized by gradual pitch change without any temporal differentiation might be conceived as a continuity – a *glissando* – not evoking any conception of categorical pitch shift. (see e.g. Sachs & Kunst 1962, Malm 1967)

<sup>28</sup> relative pitch stability both in the sense of fundamental frequency and in the sense of perceived pitch



training. For the other transformations, the result was above chance level for both categories of listeners. What is interesting here is that the D transformation implies a category overlap in relation to the previous transformations, since the categorical difference between the first and third tone is distorted.

My interpretation<sup>29</sup> of these and other results from experiments by Dowling et al (e.g. Bartlett & Dowling 1980:4) is that structurally determining melodic pitch categories allow for pitch variation within the categories, i.e. can involve subcategories/variants.

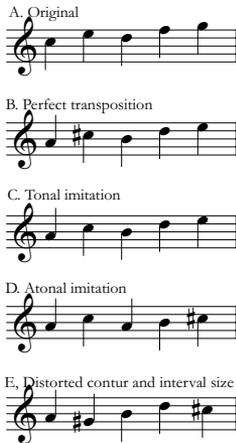


Figure 4-3. Stimulus material in discrimination test. After Dowling & Harwood 1986

I use the term melodic pitch category (MPC) to designate the central level of structurally pitch categories within perception, conception and performance of melodic structure.<sup>30</sup> A melody played in “major” or “minor” mode still remains the essentially same melody – a major-minor variation – regarding melodic structure in certain cultural contexts, in spite of the actual pitch variation. The melodic pitch categories are the ‘bricks’ of pitch, which we use to form, to experience and to cognize a melody. In practice, terms like the English terms ‘scale step’ and ‘scale degree’ or German term ‘stufe’ are sometimes used in this signification.

---

<sup>29</sup> Which is different from Dowling’s, who attributed the significant difference in the results between the three first pair and the fourth to the influence of tonality – which is interesting, since the D may very well be conceived in a tonal context.

<sup>30</sup> In Swedish I use the term ‘tonplats’, which literally means a ‘tone place’. This expression conveys an important aspect of the conception, referring to the place in a given tonal space. In English, one could possibly use the term scale degree, which focuses on the categorical aspect of the conception. However, it seems to imply the existence of a scale with reference to a tonic, which does not have to be present for a melodic pitch category to be conceived. Another possibility is perhaps the term ‘pitch level’, but the use of this term seem somewhat unclear, and seems to be frequently used (see e.g. the article of pitch in New Grove’s 1979) to designate the chromatic steps in a diatonic context, which rather applies to the concept of relative pitch.

Conceptually, the categorization of pitch variation in melodies can be regarded as a reduction of the amount of information processing needed, cognitive economy (Rosch1978); It is a definition of places of pitches in a virtual pitch space, which makes it possible for us to recognize melodic structure in terms of location, i.e. whether we have been on that pitch location before or not. This capacity is herein assumed to be limited to the capacity of categorical present ( $7\pm 2$ ). Some degree of virtual pitch stability thus has to be present. And, it can be argued that the concept in itself is dependent on pitch stability over time, although a stability in the sense of low note, high note, intermediate note is sufficient to create such a pitch space/room.

Thus, MPC, is not identical to pitch. The conception of MPCs is a result of melodic structure while the conception/perception of pitch in general is not bound to any specific musical structure. The common Western notation includes implicit MPC categorization through the graphical placement of pitches belonging to the same category on the same note line or gap and through correspondent naming of notes. Thus, according to this view, the notes c and c# belong to the same MPC. The naming of a note as c# or db can be understood as a designation of a note to an MPC. Solmization might analogously be regarded as a system of designating MPCs. The Indian sa-re-ga-ma notation and Kodaly's relative solmization are typical examples. In contrast to common Western notation solmization often also demonstrates the relationship and order between the different MPCs within a tonal context. That means that it is possible to distinguish between the strictly categorical dimension of a MPC and its position in a tonal context, which could be labeled scale degree. This part of the analysis regards only the category conception.

The MPCs can thus be understood as the fundamental categories by which we conceive melodies. Even if the system of categorization is internalized and presupposed when we begin listening to a new melody (Valkare 1987, Krumhansl 1990), the ability to adapt to melodies from a different cultural context without previous cultural learning (Krumhansl 1990 240ff) indicates that the categorization to some extent has to be explicit in the music itself.<sup>31</sup> This can be exemplified by the common ability to adapt to foreign MPC systems. After some time, people used to a heptatonic system, do not usually feel that there is anything missing in a pentatonic melody and vice versa, which indicates that the system of categorization is made explicit in the melody and that we have an ability to adapt to different categorical systems.

In traditional musicological and ethnomusicological terminology, it is customary to designate the number of MPCs per octave in terms of pentatonic, heptatonic etc. scales. However, I will avoid the term scale here, since it traditionally implies one pitch member of each category and further, that the tones can be used in immediate succession, can be played as a scale. Instead I will use the term MPC set. Hence, a pentatonic MPC set consist of five MPCs per octave.

In the following I am using this assumption to extract MPCs from the MIDI notes of Standard MIDI files. The method of analysis proposed can be viewed as a method of

---

<sup>31</sup> The most obvious indication of this is the pace by which melodies created within the western tonality have become popular over the world during the previous century. Even if it can be assumed that people can experience the same melodies very differently depending on different cultural background, it would probably not be possible to recreate, appreciate, adapt musically to a music if it 's structure was totally incomprehensible. See discussion in Chapter 1, *Introduction*.

predicting the conception of a MPC set based on the phenomenal properties of the pitch set and sequential pitch change within the melody.

## **4.2 The significance of melodic pitch categories to the analysis of melodic structure**

This part of the analysis is as have been mentioned above, the first and basic part of the analysis of melodic structure, since it concerns a fundamental aspect of melody as defined above (see Chapter 1, section 1.2 The object of the present study), melodic contour. One way of understanding melodic structure is to regard it as a significant pattern of pitch shifts between MPCs.

## **4.3 Basic information needs for analysis of Melodic Pitch Categories**

Information input in the current method of analysis is an event list, which contains pitch information at the granularity of quartertone level. The computer implementation uses the sequence of MIDI notes stored in Standard MIDI Files. There is also information about mode and key present in Standard MIDI Files, but this information is not used in the analysis of several reasons.

First, a premise of this analysis is that conception of MPCs can exist without the experience of any tonal center/tonic. Thus the MPC level is structurally primary while tonality perception is structurally secondary. This does not imply that melodic pitch categorization cannot be dependent on and, in some sense, secondary to the conception of tonality or that a certain melodic pitch categorization cannot be regarded a premise in a musical style. But it states that experience of MPCs does not primarily depend on tonality conception, which makes it possible to perform the analysis of MPC without prior analysis of tonality.

Second, a future aim of this analysis is to obtain an analysis of tonality based on the structural analysis. Further, an important part of the test material of interest in this study, does not contain interpretation of tonality in the sense that it can be used as a basis for further analysis. Rather, notation of tonality, e.g. mode and key, is in many cases arbitrary or based on notation conventions and not on actual experience of tonality.

The MIDI note numbering thus can be understood as "piano roll"-information (c.f. Temperley 2001). It simply tells which key on a keyboard to press and how long the corresponding tone should sound.

## **4.4 Basic analytical concepts**

### **4.4.1 Basic category discrimination**

How do the MPCs emerge in the melody line? How do we separate the categories from the different pitch variants?

What defines a category, a 'pitch brick', is that it can be combined with other categories 'pitch bricks' to form a melodic structure. From this notion, we can conclude that structural independency and stability is what we are looking for.

Another dimension follows from the properties of pitch perception. Pitch perception is not in itself categorical. We can experience a gradual change, raise or lowering of pitch. Pitch categorization hence means that we are dividing the pitch continuum, from which follows that category granularity is most evidently determined by the smallest pitch shifts in the melodic structure. These shifts between neighboring categories are in Western music terminology designated steps.<sup>32</sup>

From this we can make the first basic assumptions regarding what determines MPCs in a melody.<sup>33</sup>

- (1) **Pitches which follows, in immediate succession belong to different melodic pitch categories (which the exception of chromaticism noted below)**
- (2) **Neighboring pitches which do not follow in immediate succession in a melody can (under certain circumstances, see below) be experienced as pitch variants of the same category**

These two assumptions are the two basic axioms of the analysis. The melodic minor scale may serve as an example the two categorical levels implied in these assumptions, since the sixth and seventh scale degrees of a melodic minor scale – the MPC level – have two pitch variants each – two category members/variants.

## 4.4.2 Structural independency - Chromaticism

The extent to which a certain pitch is conceived as a MPC is, according to the above, a function of its structural independency or stability, which can be interpreted as the extent to which it can be combined with other pitches; the greater the number of pitch intervals it can form the higher the stability/ independency. This interpretation opens the possibility for levels of stability/independency and hence a conception of hierarchy of structural stability of pitches.

Hierarchy of structural stability is a possible interpretation of the phenomenon of chromaticism typical to Western music. In MPC terms, chromaticism can be defined as a pitch shift in which two category members/pitch variants follow in immediate succession. According to the basic assumption above, this creates structural stability/independency defining the two pitches as belonging to different categories. Still, it does not ruin the structural integrity of the MPC category.

Conception of hierarchy of structural stability requires categorical differentiation of structural stability; otherwise, we would confuse categories with subcategories<sup>34</sup>. Thus, it is

---

<sup>32</sup> I use the term pitch here even if it is, in reality, a question of pitch shift – change of pitch as a quality of tones – which eventually determines the conception of pitches or more correct 'tones of a certain pitch quality'. From the assumptions of (1) local pitch memory, i.e. the ability to recognize the reoccurrence of a tone of a certain pitch within the limits of our pitch discrimination abilities; (2) the assumption that we spontaneously quantize pitch shifts within tones; (3) and the quality of pitch is a spontaneously experienced quality of tones; it seems adequate, in order to simplify the reasoning, to use the term pitches for tones of a certain perceived pitch.

<sup>33</sup> Note that the subject in question could be a repertoire, musical style or music culture etc. Here is however the analysis of a single melody in focus.

common in e.g. Western art and popular music from the 18th and 19th centuries to find chromaticism only in successive leaps. One can perhaps regard chromaticism as a structural variation of the glissando phenomenon as a temporary inhibition of categorical pitch perception. On an instrument with fixed pitches with limited possibilities of continuous pitch variation, it is possible to give the impression of a glissando through a chromatic leap. Further, the existence of polyphony in Western music makes also the need to maintain the stability of MPC in melodies by melodic means less important, which leaves space for melodic challenge of the integrity of the MPC level<sup>35</sup>.

- (3) **Chromaticism within heptatonic, pentatonic or other MPC settings with less than 12 notes per octave are herein regarded as an exception of the general rule that different pitches in immediate succession belong to different MPCs. It is assumed that conception of chromaticism require categorical differentiation in relation to the MPC level, determined by the restriction of melodic motion within MPCs to immediate succession.**
- (4) **Chromaticism are assumed to be conceived as a temporary suspension of the MPC structure, comparable to glissando.**

### 4.4.3 Ideal sets of melodic pitch categories

From ethnomusicological studies it is evident that certain numbers of MPC sets per octave are extremely prevalent (see e.g. Scabolci 1956; Nettl 1956, Kolinski 1947,1961; Sachs & Kunst 1962, Hood 1971; Dowling & Harwood 1986; Krumhansl 1990). The most frequent and typical seem to be sets of three, five or seven MPCs per octave, known as tri- penta- and heptatonic scales.

In the 20th century, the introduction of the dodecaphonic MPC set has had great impact on Western art music. Note, however, that the number of pitches available did not increase; it was instead a method of composing in a manner which would give every note of the Western chromatic set the same melodic independency or stability – making them categories – and was thought to be a way of developing as well as avoiding tonality in the traditional sense. This can be interpreted as implied increase of the number of MPCs per octave. According to the general assumption of the validity of categorical present, MPC sets with more than 9 categories would have to be cognized as composite MPC sets of local MPC sets (c.f. Dowling 1978, Krumhansl 1990:253ff).

It is common in Western musicology to assume that the number of MPCs is constant in different octaves, even if this notion can be challenged from an ethnomusicologist point of view (See e.g. Kolinski 1961) However, it seems reasonable to assume from the evidence I have come across that the **maximum** number of categories per octave is constant, even if the number can actually change between octaves. This would not necessarily imply octave equivalence on a more specific level regarding pitch, but it makes it possible to use the octave range as a means of analysis.

---

<sup>34</sup> C.f. the suggested quarter tone melody in figure 1 in this chapter

<sup>35</sup> This is obviously a reason why this method has limited validity for melodies, which are dependent on a polyphonic structure.

- (5) I assume that most MPC sets in melodies in different cultures can be delimited by the octave range.
- (6) I assume that there are certain typical maximum numbers of MPCs per octave, most frequently three, five, seven and twelve. I also assume that the maximum number is constant between octaves.

Another property of MPC sets is that there typically seems to be limits to the extent of variation of MPCs (Hornbostel 1948, Wallin 1982). The limit of pitch variation within a MPC is evidently not a result of formulated agreement within a musical culture. Further, the acceptance of variation in intonation does evidently vary greatly even within a musical style, as can be deduced from e.g. what is regarded as acceptable variation of intonation among instrumentalist and singers within Western classical opera or popular music. (Burns 1999:248) Still, the degree of variation within categories is mostly limited to about a minor third in relation to a reference pitch in the material I have come across. To be able to account for melodies where MPCs never come in immediate succession it seems plausible to assume this as a maximum range of intonation variation of MPC.

From properties of intonation practice in different cultures, it can also be deduced that division of a pitch range in near-to-equal intervals is preferred, even if perfectly equidistant intonation and tuning systems are not prevalent.

- (7) I assume that there are typical limits to the variation of intonation in relation to a reference pitch within MPCs in different musical styles. For analytical reasons, I have chosen a tentative limit of 1,5 diatonic steps (minor third).
- (8) I assume that a division of a pitch range into near-to-equal intervals is preferred, within the intonation limit given in (7)

#### 4.4.4 Division of the octave – perfect fifth/fourth and equidistant tuning

The division of an octave in a perfect fifth and a fourth in a heptatonic MPC set is a property with obvious cultural limitations. There are for instance Far Eastern intonation practices and theoretical concepts (Thai, Javanese etc.) that seem to be based on the division of an octave in equidistant parts. East and Central African tuning practices have also frequently been as proceeding from a concept of equidistant pentatonic division of an octave.

However, in musical styles where the perfect fifth and fourth division of an octave is used or where such border intervals occur, these frame intervals seem to have a typical number of intermediate MPCs. Since these properties do not interfere with the principles of equidistance, it is possible, for analytical reasons, to initially assume that a perfect fifth typically ranges over five categories (three intermediate) and a perfect fourth over four categories (two intermediate). If the initial analysis shows that perfect fourth – fifth division of the octave is not present, the MPC set can be recalculated.

- (9) When the pitch set in the melody conform to a division of an octave into a perfect fifth and a perfect fourth in a heptatonic MPC set, the number of MPCs within perfect fifth is assumed to be limited to 5 (three intermediate MPC) and correspondingly four per perfect fourth.

- (10) **When this initial assumption is in conflict with the above rules the MPC set is recalculated according to the assumption of an equidistant MPC set (including the possibility of dodecaphonic MPC set)**

#### **4.4.5 Additional analytical considerations**

From the initial definition of MPC as determined by melodic motion it follows that the categorization is dependent on melodic context. Therefore, I take neighboring melodic motion into consideration in the method of analysis.

I also assume that the implications given by melodic motion and interval size do not always have to be in concordance, which makes the categorization a result of more than one melodic transition.

- (11) **Categorization of MPC is a result of melodic motion with the possibility of different implications. The method of analysis has therefore to take more than one pitch shift (transition) into account**

Further, I assume that the implications of MPCs are symmetrical in relation to melodic direction, i.e. that the pitch difference between two tones gives the same implication regardless whether the transition goes from a higher to a lower pitch or vice versa. This is a consequence of the assumption of the stability of the MPC set.

- (12) **Implications of pitch set structure are symmetrical in relation to interval direction**

In accordance with the general model of melodic conception presented initially, I assume that the conception of the MPC set emerges gradually and can be reconsidered from basis of new, divergent information. With recurring similar melodic information, an MPC set becomes an established framework through which melodic information is filtered. Thus, the preference for forming a new MPC set based on divergent melodic information – instead of interpreting this experience as deviation from or exception to the existing set – is inversely related to how strongly the MPC set is established, i.e. the number of pitch transitions involved.

The structural implication of pitch change makes it reasonable to assume that repetitions of tones and continuous alternations are not primarily structurally significant for MPC conception.

- (13) **The conception of MPC set emerges gradually and can be established through recurrent stimuli**
- (14) **The tendency to form a new MPC set based on divergent melodic information rather than to experience this as a deviation from an existing MPC set is inversely dependent on how well the existing MPC set is established, i.e. the number of pitch transitions involved.**
- (15) **Since the pitch changes are the fundamental means of creating a MPC set, pitch repetitions, returns and chromatic passages are omitted from the initial analysis.**

## 4.4.6 Individual differences in categorization – limitations of the method

The concept of MPC is not uncomplicated. It encompasses the stability of the pitch categorization independent of melodic motion as well as the dependency of melodic motion for the categorization of pitches, which is the medium for the categories to emerge.

In this conflict between stability of pitch categories and melodic dependency, it is possible to take different stands, to conceive the categorization differently. It is really a matter of sensitivity to local melodic turns versus sensitivity to pitch consistency and can be regarded as a parallel to individual different sensitivity to change and stability in tonality cognition. A person with her attention directed towards tonality may conceive variations of intonation in relation to her perception of changes with regards to tonal center. In contrast a person with her attention directed to melodic Gestalts may perceive a different categorization, where alterations of MPCs may be conceived as alterations/deviations from a general pattern, without the need for a stable and consistent temporal or tonality relationship.

Hence, I assume that it would be impossible to give one interpretation of all individual and cultural different MPC categorizations in all musical contexts, even if the cognitive interpretation will be common in most musical contexts.

- (16) **Melodic pitch categorization can be interpreted differently among individuals and in different musical styles. Hence, a method of analysis of MPCs can only be a prediction of MPC conception (with the exception of formally defined MPC concepts).**
- (17) **This method of analysis of MPC aims to give a weighted analysis based on both context dependency and pitch identity**

Another apparent limitation of the method is a result of the possibility within a given musical style or culture to presuppose a certain melodic pitch categorization, not giving enough information within the melody to make the analysis without the cultural/stylistic context.<sup>36</sup>

- (18) **The absence of cultural and stylistic input sets limits for the validity of the predictions given by the method of analysis. The validity is dependent on the amount of information given by the melody and presupposed within a musical style/culture respectively.**

However, this limitation makes it possible to study to what extent style and cultural knowledge influence melodic pitch categorization.

## 4.5 Method of Analysis in outline

### 4.5.1 Overview of the procedure

The method of Analysis can be summarized as follows:

---

<sup>36</sup> The reason why cultural/stylistic input are left out of the method of analysis is generally discussed above (see Chapter 1, section 1.2)

- a) MIDI notes are transposed and truncated into one octave (see Assumption no. 5)
- b) Immediate pitch repetitions, returns and chromatic leaps are omitted (see Assumption no. 15)
- c) MPC categorization is made sequentially according to a set of conditions (for a detailed overview of the initial pitch categorization conditions, see below, Table. 1) (see Assumptions no. 1, 2, 6, 7, 8, 9, 11, 12, 13)
- d) The MPC set is checked for double categorization of pitches into different MPCs (categorical integrity, Assumption no. 3)
- e) If double categorization occurs, further reduction of the melody is carried out by the removal of more ambiguous intervals from the melody. This is semitone leap (which could possibly be chromatic) and possible diminished or augmented interval combinations. Then steps c) and d) are repeated (categorical integrity, Assumption no. 3)
- f) If e) The MPC set is checked for double categorization of pitches into MPC (categorical integrity)
- g) If e) and f) and double categorization occurs, a sequential analysis is carried out, which is based on the assumption of a change of melodic pitch set throughout the melody. Hence, the analysis stops when double categorization occurs, and a second analysis starts at this point in the melody line and so forth throughout the melody. (change of MPC set, Assumptions no. 1-14)
- h) The MPC sets are tested for heptatonic symmetry, i.e. that the number of MPC per fifth doesn't exceed the maximum limit (according to assumption no. 9). If this is the case, the melody is assumed to be dodecaphonic and the midi notes are categorized sequentially according to the principle of the least number of chromatic alterations. (See Assumptions no. 9 and 10)
- i) If chromatic and other notes have been left out of the previous analysis, they are placed into categories in another categorization according to a set of conditions (Table 1. Chromatic conditions). (See Assumptions no. 3 and 4)
- j) The MPCs are named according to common Western notation based on the enharmonic spelling principle with the smallest number of accidentals possible.
- k) Following the general model, I assume that in reality the categorization of pitches evolve in real time, not as in the method with repeated calculations. It would evidently be possible to create such a computer model, which would be able to recalculate previous assumptions based on new information. However, this design, is for practical reasons, performed by repeated calculations.

## 4.5.2 Interval implications

The crucial point of the method is evidently the conditions according to which midi notes are placed into categories. The preliminary category implications are also the most problematic stage of the analysis, where the assumptions may need to be recalculated based on what happens later in the melody.

This is done according to the general principle that, within a heptatonic system based on perfect fifths and/or fourths (such as e.g. diatonic MPC sets), diminished intervals are implied through the addition of 'inner' semitone steps creating 'inner' second, third, fourth and fifth frames, which extends the implied number of intermediate MPC.

Conversely, augmented intervals are implied through the addition of neighboring 'outer' semitone steps, creating 'outer' third, fourth and fifth frames, which reduces the number of possible intermediate MPC. (Assumption no. 9)

In addition to this principle, preference for near equidistant division of intervals is assumed, such as division in major seconds of a tritone implies an augmented fourth. (Assumption no. 8)

The table below displays the conditions for preliminary interval categorization in terms of common notation in order to benefit musical comprehension.

### *Table 1. Explanations*

The table should be read as a set of exclusive conditions, ranging from smaller intervals to larger. For, if the conditions of a diminished third are not fulfilled, a whole tone step will be interpreted as a major second. Below the title of interval an abbreviated description of the principle follows. "Outer frame" refers to if neighboring intervals creates a frame interval for the interval to be interpreted.

The implication of each interval is displayed through an example input interval (arbitrary notation), which is the two first notes to the left. If the interpretation requires an interval context, the interval is displayed in the context. To the right of the arrow, the resulting interpretation is displayed. When a context interval is octave equivalent, both octave variants are displayed as a 'chord'. The actual pitches in the examples are arbitrary and are relative within each paragraph.

*Additional conditions* are displayed after the result in the form of AND or OR clauses.

The term *nine-list* designates occurrences within the nine neighboring MIDI notes (within the categorical present,  $7 \pm 2$  rule). *Member nine-list* means that the note(s) have to occur in nine-list, *avoid nine-list* means that avoidance of the particular notes is demanded.

*Avoid single MPC* designates that a categorization of a pitch, which has already been categorized into another category, should be avoided as single member in a MPC.

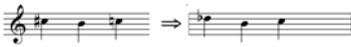
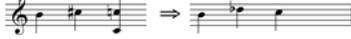
The *numbers  $\pm 2$*  etc. designates allowed variations of interval distance expressed in quarter-tone-steps, since the MIDI notes are encoded in quarter-tone steps. When the term *step is* used it refers to whole tone steps.

## Interval categorization conditions

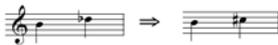
MIN2nd MINOR SECOND (not chromatic mode)  
 1/2 steps are initially interpreted as minor seconds



DIM3rd DIMINISHED THIRD  
 "Inner second"



MAJ2nd MAJOR SECOND  
 General interpretation



AUG2nd AUGMENTED SECOND

1. 'inhibited fork': outer maj3rd + added fourth/fifth frame

basic cond      result:      *Additional conditions*

*A.* nine-list      or      avoid in MPC list

*B.* both      avoid 2a) nine-list      OR      both      avoid 2b)      either missing

2. outer maj3rd + added fourth/fifth frame

basic cond      + member nine-list      or      result:      *Additional conditions*

*A.* avoid single MPC

*B.* avoid 1 nine-list      OR      avoid 2 nine-list      both      either missing

+lower fifth frame

basic cond      + member nine-list

+ lower fifth frame

basic cond      + member nine-list      result:

maj3rd frame + upper fifth / + upper fourth

basic cond      + member nine-list

maj3rd frame + upper fourth / + upper fifth

basic cond      + member nine-list      result:

# Chapter 4

a) additional conditions OR

avoid 1 ninelist OR prevprev upper third OR prevprev lower third  
current current

b) additional conditions OR

current + member ninelist OR current + member ninelist OR avoid 1 ninelist

MIN3rd MINOR THIRD  
1,5 steps is gen. interpreted as minor third

DIM 4th DIMINISHED FOURTH

a) "Inner third" 1/2 + 1,5

basic cond min + 2 basic cond max - 2 basic cond max - 2 basic cond min + 2

b) "Inner fourth" DIM5th frame

basic cond max + 4 basic cond min - 4 basic cond min - 4 basic cond max + 4

c) "Inner fifth" MIN 6th frame

basic cond max + 8 basic cond min - 8 basic cond min - 8 basic cond max + 8

d) MIN 7th frame

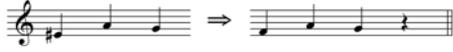
basic cond + 12

Additional conditions A. OR

avoid 1 ninelist OR prevprev OR prevprev  
current current

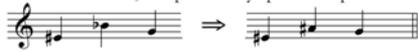
MAJ 3RD MAJOR THIRD

2 steps is normally interpreted as Major Third, 1 intermediate MPC



P5 PERFECT FOURTH

Perfect fourth. 2,5 steps is always prel. interpreted as Perfect fourth, 2 intermediate MPC



AUG 4TH AUGMENTED FOURTH

a) add "outer fifth frame"



b) "Major second -division"



c) "added fifth frame" - rev. dim 4th



Additional conditions for b) and c)



Chapter 4

or and avoid

d) "Disguised outer fifth frame"

DIM 5TH DIMINISHED FIFTH

Diminished fifth gives 2 intermediate MPC

P5TH PERFECT FIFTH

AUG5TH AUGMENTED FIFTH

3 intermediate MPC

a) Maj3rd division

b) Maj2nd division (avoid inner fifth)

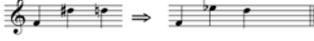
c) outer seventh/outer sixth "reversed diminished fourth"



## Chapter 4

### MIN7th MINOR SEVENTH

5 steps which isn't covered above gives initially 5 intermediate MPC



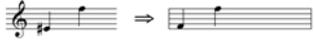
### MAJ7th MAJOR SEVENTH

All major sevenths are initially interpreted as different categories (5 intermediate)



### PERFECT OCTAVE

All perfect octaves are initially interpreted as octaves, six intermediate MPC

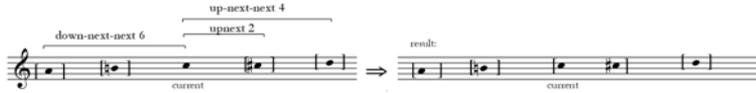


## II. CHROMATIC CONDITIONS

### A. 2 CHROMATIC IN SUCCESSION

#### 1. current with next category

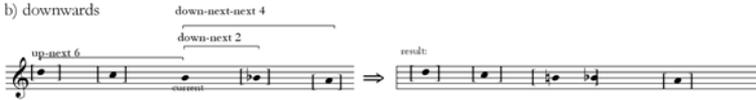
a) "Perf. fourth/MPC size conditions" upwards (brackets show categorization)



additional conditions                      or                      or  
 - current 4                                      - current 6 max prevdown                      up-next-next-next = 6



b) downwards



additional conditions                      or                      or  
 + current 4                                      - current 6 max                      down-next-next-next = 6



c) core-fifth conditions previnterval > 0



d) core-fifth conditions prev-interval < 0



2. Current with previous category

a) upwards size/fourth cond

b) downwards size/fourth cond

c) core-fifths conditions upwards (core-fifths refer to most frequent notes)

d) core-fifths villkor downwards

3. Remaining cases. Chromatic alterations in succession are alterations of previous MPC



**B. ONE CHROMATIC ALTERATION**

1. Current with previous category

a) and b) MPC-Size-conditions exactly as above 2a) and b)

c) Core-fifths conditions upwards (based on variable/nonvariable MPC)



d) Core-fifths conditions downwards



## 4.6 Evaluation of the Analysis of Melodic Pitch Categories

### 4.6.1 Means of Evaluation

The reference material for evaluating the success of the MPC analysis has predominantly consisted of notations. It is evident that notations do not necessarily reflect cognition of melodic surface structure. It is, however, reasonable to assume that the discrepancies between cognized MPC structure and notated MPC structure, i.e. pitch spelling, will not significantly affect the results of the evaluation. The reason for making this assumption is that notation is basically a means of communicating the pertinent level of melodic structure<sup>37</sup>. Thus notation can be regarded a representation of cognized melodic surface structure at a fundamental level. In other words I assume, that possible notations of a melody, which would significantly differ from the conceptions of the composer/transcriber or would be insignificant to the reader, generally will be rejected.

Besides comparison with musical notation, I have also evaluated the success of the MPC analysis, by applying the MPC analysis output in the subsequent parts of the method of analysis. This has not been done in any systematic manner, but is still of importance, since deficiencies in the MPC analysis will seriously affect the result of the subsequent analysis of melodic structure. Thus, the results presented in the subsequent parts of the thesis support the validity of the MPC analysis.

To make this point clear I will in the following give some examples of the behavior and significance of the MPC analysis.

---

<sup>37</sup> given the exclusion of pure *tablature* notation.

## 4.6.2 The significance of the MPC analysis at the level of melodic surface structure

The assumption behind the concept of melodic pitch categories is that there is a categorical difference between structurally significant and structurally non-significant pitch shifts or interval differences (see section 4.4.1). This assumption is, as has been mentioned above, strongly supported by empirical findings of melody recognition. The goal of the analysis of melodic pitch categories is thus to differentiate between structurally significant and non-significant pitch differences. According to this concept a pitch shift on the MPC level results is significant for the conception of melodic contour, while a change of interval below this level does not affect the conception of melodic contour. Thus, an incorrect MPC interpretation distorts the melodic contour.

The significance of the MPC interpretation for the representation of melodic contour is illustrated in the following example.

The image shows a musical score for the melody part of Boléro by M. Ravel. It consists of seven staves of music. The first staff begins with a tempo marking of  $\text{♩} = 96$ . The music is written in a single melodic line. The fifth staff is circled with a large, thin oval, highlighting a section where the pitch spellings (accidentals) are used in a way that distorts significant changes at the level of melodic contour. This section shows a sequence of notes with various accidentals (sharps, naturals, flats) that do not clearly represent the intended melodic contour.

*Figure 4-4. The melody part of Boléro by M. Ravel. Preliminary input notation before MPC analysis. The enclosed part displays pitch spellings that distorts significant changes at the level of melodic contour.*

Figure 4 displays the input representation of the melody part of Boléro by M. Ravel.<sup>38</sup> The initial spelling of MIDI notes is tentative and not the result of an analysis. The initial pitch spelling is, however, not coherent with regards to melodic contour. In the enclosed area (and remaining part) the initial pitch spelling assigns notes, which probably will be conceived as belonging to different categories, into the same category. (See e.g. the notes a#-a in the beginning of the enclosed area). Correspondingly, pitch shifts between what will probably be conceived as neighboring categories (e.g. a#-c) are spelled as if there was an intermediate pitch category. This spelling distorts the melodic contour of this part of the melody, resulting in primarily pitch repetitions at the categorical level.

When the MPC analysis is performed the output is in accordance with the pitch spelling in the original notation of the theme, which represents the melodic line as pitch shifts between categories.



Figure 4-5. Excerpt of Boléro. Output pitch spellings after melodic pitch category analysis. Refers to the enclosed area in the input representation above.

<sup>38</sup> The graphical interface of the MODUS implementation the current method of analysis is limited. It does e.g. not display any beaming between MIDI notes. Nor does it assign accidentals according to common practice: All notes that are altered are preceded by an accidental and consequently all notes not preceded by any accidental are naturals. NB Metronome value and time signature is given by the input file and is not significant in this part of the analysis.





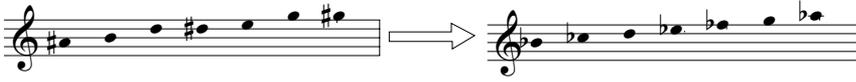


Figure 4-11. Scale representation of the input MPC representation (to the left) and the resulting MPC interpretation after MPC analysis.

### 4.6.3 Chromaticism and extra-MPC alterations

Chromaticism within a non-dodecaphonic MPC set, as have been suggested in section 4.4.2, should be regarded as a sub-categorical level to the central MPC level. The obvious problem regarding melodic pitch categorization in melodies with chromaticism is thus to separate these levels. In other words, to determine which notes belong to the same category.

A popular melody that exhibits chromaticism within Western music is the famous Habanera from the opera *Carmen* by G. Bizet.



Figure 4-12. Input pitch spelling of Habanera from *Carmen* by G. Bizet. The melody is reduced and transposed from the original.

To obtain a hierarchy of structural stability on different levels the MPC analysis first disregards pitches within chromatic leaps, based on the assumption of chromaticism being a temporary suspension of the MPC hierarchy. In the next step the possible interpretations of the notes within chromatic leaps according to the conditions of the maximum number of pitch categories within interval frames and the preferred size limits for MPC categories. Besides that it makes use of the basic assumption that chromatic steps within categories basically can be interpreted as an alteration.

In the case of the Habanera the pitch categorization concurs with the original pitch spelling.

Chapter 4



Figure 4-13. Output pitch spelling of the Habanera version after MPC analysis

The importance of a consistent interpretation of chromaticism, with regards to MPC level for the determination of melodic contour, is demonstrated in the example below.



Figure 4-14. Input pitch spelling of constructed melody. Parallel melodic segment is indicated by brackets.

In this constructed melody the melodic parallelism is somewhat obscured in the initial pitch representation since the parallel melodic lines are interpreted differently in the original presentation and the transposed repetition. The change of pitch spelling in the output of the MPC analysis interprets the second measure as an ascending scalar motion parallel to the transposed repetition.



Figure 4-15. Output pitch spelling of constructed melody after MPC analysis. Parallel segments are indicated by brackets. Extra-MPC alteration is enclosed.

This example also contains yet another feature of the MPC analysis. The interpretation of chromaticism as a temporary suspension of the MPC integrity allows for the analysis to handle conflicts between the different MPC implications in terms of local extra-MPC alterations. Typically this concerns conflict between the fundamental MPC rule, which states that pitch shifts occur between categories, and the integrity of the MPCs. In this case the higher local stability of the f category allows the analysis to make the segments parallel, with regards to MPC interpretation.

The interpretation of chromaticism is difficult in melodies with a fundamentally pentatonic structure and it is relatively uncommon in such styles. However, this is rather a typical trait of American blues idiom and related styles, such as jazz and some American popular song of the 20<sup>th</sup> century. In those melodic styles it is not uncommon that the structural stability of the chromatic variants becomes so strong that it is possible to conceive variants as different MPCs.

The output pitch spelling of the initial theme of “Stormy Weather” (Koehler/Arlen, “Vocal Real Book” version), identical to the reference notation, demonstrates this problem.



Figure 4-16. Beginning of “Stormy Weather” (Koehler/Arlen). “Vocal Real Book” version. Output pitch spelling after MPC analysis.

Although this melody is not pentatonic, it exhibits pentatonic traits, by frequent interval combinations of thirds and seconds. Thus, the interpretation of semi-tone steps becomes ambiguous. According to general pitch stability and the chromatic alteration rule the two first

notes of the melody are to be interpreted as variants within the same MPC – in terms of tonality a minor-major variation of the third scale degree. But according to the fundamental MPC rule, which states that pitch shifts occur between categories, they are to be regarded as belonging to different MPCs. In this case the fundamental rule applies, (which is in accordance with the reference notation), but since the MPC analysis is highly contextual, a structurally non-significant alteration of the melody may result in a different interpretation.

Both interpretations do occur in contemporary notations, which may reflect different cognitive interpretations or just differences in notation practice. Viewed in this perspective one could perhaps describe this stylistic trait as an application of 19<sup>th</sup> century Western chromaticism on a fundamentally pentatonic structure with variable MPC intonation.

#### 4.6.4 Evaluation of the success of the Analysis of Melodic Pitch Categories

The evaluation material has consisted of more than two hundred melodies of different stylistic origin. The main part (ca 120 melodies) has been melodies belonging to the Scandinavian folk music idiom including both vocal and instrumental melodies. In this repertoire the input notations have included quarter-tone alterations. Apart from this bulk of melodies the reference material includes Arabic classical songs (10 melodies), Carnatic Ragas (10 melodies), Balkan tunes (20 melodies), 20<sup>th</sup> century popular songs and jazz standards (15 melodies), the Sonatas and Partitas for solo violin by J.S. Bach with a predominantly melodic texture<sup>39</sup> and other melodies from the repertoire of Western classical music (25 melodies). Single examples of African and Far Eastern melodies have also been included.

The total number of melodies in this reference material in which the result of the MPC analysis is consistent with the reference notations is 195, while the number of analyses clearly inconsistent is 5. Of those, three belong to the repertoire of Western Classical Music and two belong to Jazz and popular song. These inconsistencies were due to changes of tonality that would require a change of MPC set.

Those examples show that the model fails when the melody contains sudden shifts between very distant tonalities. It suggests that change of MPC set should be included in further development of the implementation.

It should be noted that the reference material does not contain any dodecaphonic melodies. The reason is that the MPC interpretation does not require any categorization of variants of pitch categories in dodecaphonic melodies. Pitch spelling is thus merely a technical question of graphical significance.

The overall result indicates the validity of the proposed model of analysis of melodic pitch categorization. Furthermore, it indicates that this level of melodic organization may to some degree be regarded as cross-cultural.<sup>40</sup> It should be noted, however, that Western common notation is very limited as a means of representing the MPC level, such as pentatonic or dodecaphonic MPC sets. The representation of the MPC analysis in Western common notation is a matter of convenience.

---

<sup>39</sup> Where the omitting of chords do not radically alter the musical structure.

<sup>40</sup> A discussion of cultural implications of interval interpretation can be found in Valkare 1987

# 5 Metrical Analysis

---

## 5.1 Meter – general concepts

*“It would be a hopeless task to search for a definition of rhythm which would prove acceptable even to even a small minority of musicians and writers on music” (Apel 1945 quoted after Cooper & Meyer 1960)*

### 5.1.1 Meter and Gestural Rhythm

Practically all music theoretical terms are used in a number of different significations in different contexts. This is especially true for terms used both in music practice and in musicology. Both the different use of the same terms in different languages and their historical background contribute significantly to the terminological confusion. Standardisation of terminology exists within branches of musicology, but it is evident that since musical terminology relates to ideas about the nature of music it is hard to achieve consensus.

One area where the confusion of terminology is extremely prevalent is the temporal dimension of music, for which the term *rhythm* is most frequently used. To my knowledge, there is no contemporary consensus regarding the use and definition of the term *rhythm* within the musicological community, which makes term problematic. One of the central issues is whether rhythm should be connected with periodicity and/or grouping. For example, some writers (e.g. Gabriellson 1984:249ff, Blom 1989:33ff) presume periodicity to be a prerequisite for rhythm conception, while other writers (see e.g. Cooper & Meyer 1960:6, Lerdahl & Jackendoff 1983:18) assumes conception of grouping to be independent of periodicity. Another problematic issue is the division between rhythm and form (see e.g. Clarke 1999:476).

As have been explained in Chapter 3, section 3.3 *Dimensions of Melodic Surface Structure*, I have found it useful to distinguish between two different aspects of rhythm, ‘*Gestalt conception*’ (grouping) and ‘*conception of periodicity*’ (meter) (Ahlbäck 1983-1996, 1986, 1989; cf. Cooper & Meyer 1960:6, Lerdahl & Jackendoff 1983:18, Clarke 1999:476, Jones 1987, London 2002:531). I have also postulated that these conceptions may be independent but more typically interact in the cognition of temporal structure in music.

To clarify the definition I will repeat the figure (see fig. 1) presented in section 3.3 with the addition of some common musical terms categorized according to the division in meter and gestural rhythm/grouping. The term meter is here used to designate the conception of regulative periodicity at all structural levels of rhythmic significance. Regulative periodicity refers to the conception of a periodical temporal pattern, to which temporal events will be related. The definition of meter given by London (2002:531) following Jones et al, as a “*stable, recurring pattern of temporal expectations, with peaks in the listener’s expectations coordinated with significant events in the temporally unfolding musical surface*” is analogous to the definition given above.

- (1) **Meter is used here in the sense of perceived and/or structural periodicity in music, to which musical events are conceived as related. Meter is thus, whenever conceptually present, a primary level of the temporal organization of music.**

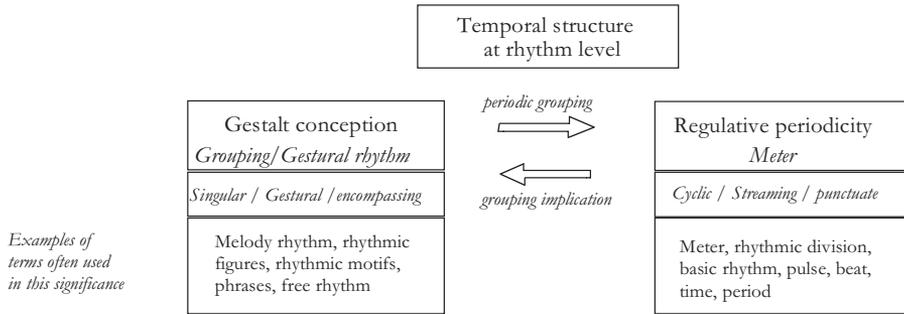


Figure 5-1 The two dimensions of temporal structure in music at rhythm level. Examples of related common terms.

As have been mentioned above, many writers of musical rhythm seem to exclude music without metrical structure from the rhythm concept (see e.g. Gabrielsson 1984:251-252, *rhythm and non-rhythm*). This seems to be a narrow perspective since there are many musical styles to which the concept of meter is not applicable. Examples of such genres are different kinds of calls, laments, recitative etc., which are generally not conceived as metrical. In such music the temporal structure is often described in terms of *free rhythm*. To exclude *free rhythm* from the rhythm concept is to imply that temporal relationships are insignificant, which seems to be a dangerous limitation. (See further discussion in Chapter 7, *Non-metrical structure*, section 7.1)

Conversely, there are examples of music where the structure does not support cognition of gestural rhythm. Examples of such musical structures can be found e.g. in late 20<sup>th</sup> century Western art music (see e.g. “*Pulse music*” and other works by S. Reich, “*Pleiades*” by I. Xenakis) and in contemporary dance-music styles such as some Techno music, where the structural emphasis lies entirely on periodicity. The existence of such cyclic music suggests that it is possible to conceive pulse streams without grouping, even though our tendency to experience grouping seems to be innate and spontaneous.

The level of metricity can also be regarded as relative. It is extremely common in e.g. European vocal folk music styles that a metrical structure can be conceived, but is treated very freely. Such songs may include parts with loosely regular structure, but also fermata points, where the pulse stream can be conceived as temporally terminated. This is often with a term proposed by Bela Bartok (Bartok 1920:11) often referred to as *parlando rubato*. Here, we are in a border area between metricity and non-metricity, with no clear perceptual or conceptual distinction.

However, here it is assumed that when a metrical structure is conceived, it serves as the primary temporal grid to which musical events are related. Thus, a melodic structure is either regarded metrical or non-metrical, although the method takes e.g. local metricity in non-metrical contexts into account and include the possibility of extra-metrical events in a metrical context.

- (2) **Conception of meter and gestural rhythm in music can be simultaneous and interrelated, but meter and gestural rhythm can also exist independently of one**

**another, such as in music without conceived meter or music without conceived gestural rhythm.**

- (3) **Conception of metrical structure, the degree to which a perceived periodicity is conceived as a temporal grid, can be regarded as relative . There can be extra-metrical events within a generally metrical context.**

A foundation of this model is also that these conceptions may be regarded both as phenomenal and cognitive entities. The phenomenal and cognitive structure may correspond, but they do not have to. The closest sounds will not always be grouped together. Nor is there always a strict correlation between phenomenal periodicity and perceived periodicity.

According to this view **to walk** is a periodical action, while **taking a step** is a gestural action. In this sense meter can be regarded as the ability to perceive and relate to periodicity as a means of coordinating actions. Assumed here to be a fundamental human ability, this may be interpreted as the kinaesthetic correlate to the psychological concept of streaming.

- (4) **The ability to conceive meter in music is assumed to be innate and spontaneous, closely related to the ability to relate to periodicity as a means of coordinating actions and can be regarded as a streaming phenomenon.**

But our conception of periodicity is not limited to phenomenal periodicity. It is a well-known fact that we possess the ability of perceiving periodicity even when it is not phenomenally present or obvious (see e.g. Gabrielsson et al 1983). Furthermore, there are many examples of musical styles, such as e.g. Balkan dance music, where dance patterns and musical structure indicate an asymmetrical beat pattern, with recurrent, significant changes in beat periodicity (see e.g. Singer 1974:379-404). The herein proposed definition of meter incorporates quasi-periodicity, hence both walking and limping are considered fundamentally metrical actions.

Referring to the discussion above regarding the distinction between metrical and non-metrical structure, it seems reasonable to assume that it is easier to conceive periodicity when phenomenal periodicity exists, and hence that regularity and consistency of regularity of musical structure enhances metrical experience. Thus I assume that the level to which music is conceived as metrical is related to level of regularity/isochronicity and consistency of isochronicity.

- (5) **Meter is here used in the sense of conceived periodicity, including quasi-periodicity, which implies that events that are not phenomenally isochronical can be conceived metrically. It also implies that a quasi-periodical temporal grid may be conceived as regulative, involving expectations of temporal attention according to a significant scheme of period fluctuations.**
- (6) **Metrical conception is in concordance with the general model of melody conception assumed to be evolving gradually through the listening to a piece of music**
- (7) **People are assumed to be sensitive to physical periodicity (regularity) in sound patterns (timing) and to the consistency of periodicity.**

I assume that the tendency to relate metrically to music differs among people on an individual, cultural and social basis.

- (8) **There are individual differences in the tendency to relate metrically to music. Such differences can possibly even be culturally or stylistically discrete.**

The fundamental assumption that metrical conception is spontaneous (*Assumption no 2* above) is strongly supported by empirical research (see e.g. Clarke 1999:483). This could be interpreted as if we are recognising periodicity actively, searching for regularities. Viewed herein as an exponent of the fundamental assumption of *cognitive economy*, the search for regularities is unconscious and continuous (Cf. section 3.3). Thus it is logical to assume that any perceptually recognizable dimension of musical sound can be used a source for conception of periodicity.

- (9) **A periodical change in any perceptually recognizable dimension of musical sound can be conceived metrically.**

### 5.1.2 Pulse

Pulse as a musical concept is generally much used in most analyzes of music, yet it is hard to find definitions of pulse in traditional works on music analysis. I have not managed to find such a definition within the music theory literature, which has proved to be useful for my purposes.

Pulse is often described in terms of a conceived flow (see e.g. Blom 1989). A *'flow'* of what? Usually a flow of *'beats'*. But then, what are beats? *Beats* commonly regarded a periodic quality that we extract from the sound that is not always present in the physical sound. In the words of London (2002:531) "*metrical articulation or time points may be read as the temporal locations of the peaks of attentional pulses; [...] moments of attentional energy*". But *beats* can also be a property of the sounding music, marked by physical beats or dynamic articulations. *Beats* could possibly then be regarded as conceived periodical markings of time in music. But then we turn the next question: – Are all periodical markings of time in music regarded as beats? With regard to the use of the concepts of *pulse* and *beats*, this does not seem to be the case. (Cf. London 2002) A metronome, which in the Western world often is used for measuring the tempo of pulse, usually has a scale ranging between 40 to 240 MM.

It is common to find shorter periodical sounds than 240 MM. in music with a designated tempo of the tactus of e.g. 120 MM. in Western Music, which are not generally conceived as pulse. It seems to be common in e.g. Western and Asian music to find both shorter and longer periodic sounds in music than what is conceived as pulse (or meter in general). This is in compliance with the concept of meter as a time grid for events to happen. Shorter sounds than perceived pulse is experienced as happening between the beats and longer sounds as encompassing more than one beat (cf. London 2002).

Generally, it seems as if the concept of pulse is related to pulse frequency - pace / tempo. There is empirical evidence (e.g. Parncutt 1994), which indicates a tempo range for pulse conception between 1800 ms to about 200 (33 MM- 300 MM)<sup>41</sup>. Regarding the close phenomenal relationship between musical pulse and motion, it would be no surprise if there is a connection between the possible pace of human actions and pulse conception. (see e.g. Clarke 1999, London 2002). Pulse seems to be most frequently used to designate the shortest periodicity of referential function, events either occur on beats or between beats. This implies an atomic quality of pulse as a fundamental temporal grid of the music. The main referential pulse is in the following designated central pulse level or tactus.

---

<sup>41</sup> Higher tempo values have been reported (see e.g. Björnberg)

The concept of pulse is further suggested to be related to the perceptual present. It is herein assumed that, for a pulse to serve as the basic temporal mapping, the periodicity must be perceivable within the basic scope of attention limited to the perceptual present. This assumption implies that the upper limit of duration of pulse period is depicted by the perception of a continuity within the perceptual present; hence the perceived pulse accentuation must occur more than twice during a perceptual present. Given a typical value of the perceptual present between 5 and 6 seconds this limits the maximum duration of a pulse to around 1800 ms, which is upper value indicated in the aforementioned empirical studies of preferred pulse rate.

Besides the tempo dependency of the pulse concept – following the concept of pulse as the fundamental temporal mapping of the musical flow – I assume that there are ideally periodical levels with both shorter period duration and longer period duration, to form a stable pulse conception (cf. London 2002). Then the pulse functions as a temporal coordinator of shorter events and denominator of longer events, thus a basic temporal map which makes it possible to handle events within the limits of the categorical present (see section 3.1). Translated into the pulse concept this implies that a beat should ideally handle a maximum of between two and four higher and lower periodicities. A consequence of this is that a perceived periodicity at one end of the category spectrum, for example, representing the shortest durations of the melody but only 1/8 of the longest durations, would not be an ideal pulse and we would rather look for a periodicity at a higher level.

To summarize, pulse can be defined as a periodical flow of beats, where beats are the shortest periodic perceived or phenomenal markings of time in music, which can be conceived as the basic units to which temporal events relate.

- (10) **Pulse is the periodic flow of beats in music, the fundamental structure of conceived periodicity in music, whenever present it constitutes the basic temporal mapping of music**
- (11) **Beats are perceived or structural periodic markings of time (i.e. pulse periods) or points of temporal expectation in music, and can be conceived as the basic units to which temporal events relate (including complex metrical events)**
- (12) **If a perceived periodicity is conceived as pulse relates to tempo (pulse frequency). Periodicity is perceived as pulse within a certain tempo-range.**
- (13) **It is herein assumed that, for a pulse to serve as the basic temporal mapping, the periodicity must be perceivable within the basic scope of attention limited to the perceptual present.**
- (14) **If a perceived periodicity is conceived as pulse is related to interonset structure, which duration categories that are present. A pulse is ideally a central duration category and conforms to the limitations of the categorical present.<sup>42</sup>**

### 5.1.3 Metrical complexity

A central assumption within this model is that meter can be conceived as multi-layered: That we possess the ability to perceive regularity/periodicity on different levels simultaneously.

---

<sup>42</sup> Note that the above definitions are operational to be evaluated regarding their usefulness and generality as used in this method as a model of segmentation of monophonic melodies.

This notion is, for instance, supported by the presumed simultaneous existence and experience of pulse and time/meter in common Western notation.

A central concept in Western music theory is the assumption of the singularity of pulse perception, pulse as *tactus*. Some theorists have extended this concept to assume the simultaneous existence of more than one pulse level, thus designating *tactus* central pulse level. Theoretical concepts like polyrhythmicity and experimental studies (Meyer and Palmer 2001, quoted in London 2002:541) indicate that pulse can simultaneously be perceived on different levels and people can choose which pulse level to be the *tactus*. On the basis of this evidence one can question if the conception of central pulse level should be considered as axiomatic in all contexts among listeners. Studies of metrical experience such as tapping studies seem to presuppose the existence of a *tactus* through the experimental task of finding one pulse level. It is not uncommon for performers in different dance music traditions (such as e.g. French-Canadian and Norwegian traditional dance music) to use both feet tapping at different paces, both of which would be conceivable as pulses. This phenomenon could be regarded as representing a complex pulse conception.

On the other hand, the need for a referential pulse in polymetrical conception and performance is well known, at least Western musicians in general tend to experience pulses at the rates of 4:3 as something else than 3:4 (see e.g. Jersild 1975) and often having difficulties in deliberately changing between the conceptions in performance. But musicians who denote much attention to such structures (such as e.g. many percussionists) seem to acquire a skill in experiencing simultaneous pulses, which results in a conception of simultaneous pulses at different pace. This is to my experience also often reported by listeners to music with focus on the metrical structure, as a conception of a 'package' of pulses at different paces creating a metrical conflict or 'swing'.

Hence, even if it is possible to always experience one pulse level as referential, as *tactus* or central pulse level, the singularity of pulse experience does not seem to be unambiguously supported by either experimental studies or musical behaviour. Therefore, I have rather found it useful to regard pulse experience as potentially complex; to assume that it is possible to conceive parallel pulses, streams of basic units on different levels simultaneously and that it is possible to have a fluctuating conception of central pulse level in a stable complex of pulses.

**(15) Meter is assumed potentially be conceived as complex in a hierarchical manner, as more than one level of periodicity in simultaneous coordination. This can be expressed as conception of superimposed pulses (polymetricity) or the simultaneous experience of time and pulse. One of these levels are often conceived as the central pulse level or *tactus*.**

Thus, since meter, whenever present, is assumed to be perceived as the fundamental temporal organization of the musical structure, the metrical levels have to be perceived as related to each other. Accordingly this relationship does not need to be simple. There are many examples of peoples ability to conceive sound or action patterns as guided by different tempos where the events on different levels do not always coincide, often labelled polyrhythmic experience (see e.g. Jones 1959).

**(16) The ability to conceive metrical conflict between pulses of different tempo in a complex metrical conception is assumed.**

There are at least three different types of complex pulse conception, each with different structural and possibly conceivable levels of metrical conflict;

- a) complex/superimposed meters conceived as **even** tempo relationships, such as 2:1, 3:1, 4:1 etc., where beats of slower tempi always coincide with beats of faster tempi;
  - b) complex/superimposed meters conceived as **uneven** tempo relationships, such as 3:2, 4:3, 5:2 etc., where beats of slower tempi periodically do not coincide with beats of faster tempi;
  - c) complex/superimposed meters conceived as **phase displacement** of meters of identical tempo, 1:1., where beats of the different tempi never coincide.
- (17) **Metrical complexity at pulse level can include conception of superimposition of pulses as well as composite beat groups / beat asymmetry. Pulse should however be possible to conceive as the basic periodical measurement of the music, while it has to conform to the conditions given by the perceptual present.**

The concept of complex pulses poses a central question, namely the distinction between time (meter at measure level) and pulse.

Within Western music theory and common notation practice, there is not a clear division between pulse and higher levels of metricity such as time. But one could perhaps speak of time and other higher levels of metricity as conceived periodical “grouping of beats”, as composite metrical levels as opposed to atomic/basic metrical levels.

Conception of metrical complexity at beat level can regard both superimposition of pulses and composite beats, i.e. additive rhythm (Sachs 1953). It would be an impossible task to make a clear conceptual division between these composite pulse levels and the time concept. Examples of this problem are frequent in notation practice in e.g. transcriptions of European folk music. You can, for instance, find waltzes notated in 3/4, 3/8 and sometimes 6/8, depending on the orthographic tradition and the conception of the transcriber. This indicates that transcriber have experienced as a pulse what the other has experienced as time, grouping of beats. The division between the time and pulse concept in this work is defined by the pulse concept, which implies that a pulse must be possible to conceive as the basic periodical temporal grid in music. Periodicity at measure level – time – thus can be defined as regulative periodicity determined by periodical grouping of beats.

This implies a relationship between grouping and meter, which is not generally accepted (see e.g. Lerdahl & Jackendoff 1983). Meter is generally interpreted in terms of accent structure. In the current work it is suggested that metrical conception at all levels above interonset level is inferred from conception of grouping. Grouping is thus assumed to implicate the accent structure at higher metrical levels. This suggestion is based on the fact that meter can be conceived by listeners when exposed to a sounding stimuli consisting of tones with equal interonset intervals shorter than the conceived meter and with no differentiation in phenomenal articulation (see e.g. Chapter 6, section 6.2.6.3, *experiment 3a*). This implies that the metrical experience must be derived from the sounding structure by other means than interonset interval or phenomenal accent structure. The results of the referred experiments demonstrates that this can be interpreted as based on the conception of grouping structure, i.e. that there is no exclusively metrical pitch accentuation that determines the metrical conception of the listener. The relationship between meter and grouping is even more articulated when higher metrical levels are considered. As will be demonstrated in the evaluation section in this chapter and in chapter 6, it is impossible to derive periodicity on measure level from phenomenal accent structure alone.

On the other hand, analytical models of metrical structure which exclusively make use of interonset information for determination of the metrical structure, seem to perform well on lower metrical levels in musical examples with durational contrast. (See summary in e.g. Temperley 2001:27-54, Clarke 1999:482-489). The relative success of these models suggests that interonset structure plays a crucial role for metrical cognition.

The relative influence of grouping and durational patterns leads to the assumption that the influence of duration patterns (interonset patterns) for the determination of metrical structure is inversely relative to the length of the structure in terms of number of onsets: The more onsets the greater the influence of grouping for the cognition of metrical structure and vice versa.

There is, however, a problem concerning the herein suggested interrelationship between meter and grouping. The crucial point of Lerdahl & Jackendoff's theoretical concept was that metrical structure and grouping structure must be treated separately since they do not always concur. Since this is true, how can the suggested relationship given above apply? A possible solution to this problem is indicated by Cooper & Meyer (1960:8), who use the terminology end-accented and beginning-accented grouping.

In this work I am suggesting that grouping on the level of musical surface structure is generally not conceived only in terms of a time-span, encompassing a number of events. Musical terminology, such as upbeats, anacrusis etc. reflects the conception of something that occurs before virtual start of the phrase. I thus suggest that the conception of a group may include determination of the group start in two dimensions, as determined by the end of the previous group and as determined by the structurally first gravely accented event in the group, designated start- end end-oriented grouping (see further discussion in chapter 6, section 6.1). This concept implies that both structural accentuation and structural distance can imply grouping. A group boundary may thus be ambiguous, providing the possibility for the listener to attend to either the end of the previous or the start of the impulse point of the current. I further assume that listeners may be more or less directed to one of the conceptions. This is strongly supported by the results of the grouping experiments in chapter 6.

The important implication of this view on grouping for the analysis of metrical structure is that start-oriented groupings, i.e. the virtual start of a conceived group, is what is input from the conception of grouping structure to metrical structure. Conversely metrical structure implicates virtual starting points which affects conception of grouping through the accent pattern – implying moments of “attentional energy”.

- (18) **Pulse at atomic level is structurally defined mainly by periodical interonset intervals, while higher metrical levels such as composite pulse and time are dependent to larger extent on periodicity in grouping of events, which involve other musical dimensions such as pitch. The strength of durational accent is assumed to be inversely relative to the number of events within the period.**
- (19) **Time (meter at measure level) is defined as regulative periodicity determined by periodical grouping of beats**
- (20) **Meter is inferred from grouping by the perception of group starts as impulse points, as structural accents. Grouping is inferred from meter by the perception of temporal positions determined by the metrical grid as impulse points, implicating group start.**

Note that this definition excludes composite meters which cannot be perceived as periodical in any sense. This is important to point out since time signatures often are used solely for notational purposes in e.g. modern Western Art music.

### 5.1.4 Accentuation and the impulse quality of meter

The concept of measure start is so common throughout at least Asian and Western music traditions that I am assuming it to be cross-cultural. It is obviously possible for us to conceive periodicity without implying a starting point of the movement, such as pendulum movements, wave motion etc. (cf. Levy 1983) But, since the pulse level can be perceived as periodical points of attention, they can be regarded impulse points, with a punctuate dimension in analogy with the pulse concept.

Hence, measures are here assumed to be counted from the beat, which is conceived as having the most prominent accentuation within each measure.

- (21) **Meter at measure level is in analogy with the general concept of meter assumed to have punctuate dimension, implying impulse points. Thus measures, in terms of pulse groups, will be counted from the beat conceived to have the most prominent accentuation.**

The last statement may seem rather dubious, since it is a common experience among music listeners that phenomenally non-accentuated beats like rests can be experienced as measure starts. One should recall this refers to conceived accentuation, and according to the general model of melody cognition (Chapter 3), the pattern conception does not need to be constantly supported by phenomenal input.

The basic statement regarding metrical marking is that any periodical change in any perceptible dimension of the musical sound can function as a metrical marking (*Assumption no 9*). The generic term for metrical markings is accent. Accents can according to the terminology proposed by Lerdahl & Jackendoff be classified as phenomenal, structural or metrical. A phenomenal accent is according to this terminology points of local intensification cause by physical properties in the stimulus such as changes in intensity (dynamic accents), note density, register, timbre or duration. Structural accents refers to accents derived from the cognition of the musical structure, such as changes in tonality, melodic parallelism etc. Metrical accent refers to accentuation implied by the metrical scheme.

I have found it useful to distinguish between two types of accent quality, here designated **grave** and **acute** accentuation, by which the quality of any of the three types of accents can be

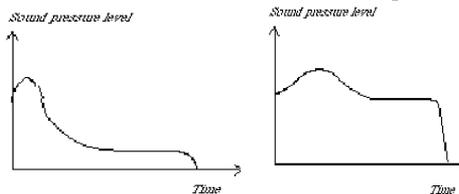


Figure 5-2. Typical 'acute' dynamic accent Figure 5-3. Typical 'grave' dynamic accent

described. A *grave* accent is conceived as *grave*, i.e. sustained (long), heavy, round, deep, while an *acute* accent is conceived as *acute* (short), light, sharp, high. The typical sound envelopes for the two accent qualities as phenomenal dynamic accents are shown in the figures above:

These categories resemble the behaviour of light and heavy objects. A heavy object, like a car, does not move immediately. We apply pressure to it, and once it starts rolling it takes time or force to stop it. A light object like a straw can be moved by a snap of a finger, but it quickly stops moving since it has little mass.

If we compare these categories with human motion, the *grave* accent can be likened to the typical change of pressure on the foot when we take a step: The maximum pressure/weight is not achieved immediately, but gradually, until the foot starts to leave for the next step to come and the pressure is lowered. The *acute* accent can be compared to a jump, it reaches its maximum almost immediately, requiring a sudden impulse for us to lift, but, due to gravitation, we will quickly return to ground. Hence, the impulse phase is most evident. Walking and jumping, as well as typical behaviour of light and heavy objects are common experiences, and I assume that the ability to experience structural similarities between accent types (acoustic/structural) with human physical behaviour and the behaviour of objects of different weight on a subconscious level is general. This is above all supported by the strong correlation found in many types of dance music between dance patterns and musical structure.

Note that this is an ideal dichotomy and the concepts of '*gravity*' and '*acuteness*' are relative. Therefore, it is rather a question of an accent being relatively '*grave*' or '*acute*', a quality of accent, than of absolute qualities.

What is interesting with this typology of accents in the current context is its importance for the theory of beat- and measure starts. To determine the impulse point is crucial for determining the metrical levels; beats, measures, periods. This is in turn crucial for determining grouping structure if the basic assumption that meter, whenever present, is the fundamental temporal mapping of the music.

It is well-known that in certain musical contexts e.g. the loudest dynamic accents and often also certain types of structural accents are on beats which are not conceived as measure starts. This is evident in e.g. much Western popular music where 'backbeats' are acoustically most marked in the musical sound, however still conceived as 'backbeats' and not as 'downbeats'. This can be viewed in the light of the *grave/acute* classification of accents. If *grave* accents are conceived as more prominent than *acute* accents, it wouldn't matter "how hard the drum will be hit" - it would still be a backbeat. The results of a previous study of relationship between phenomenal accentuation and conceived measure start (Ahlbäck 2003, in press) indicates that people prefer to put the measure-start at the *grave* accents rather than at the *acute*, in a context with periodical differences in dynamic accentuation of beats which can be described according to the *grave-acute* dichotomy. (See also section 5.2.8)

- (22) **It is assumed that people possess the ability to experience structural similarities between accent types (phenomenal/structural/metrical) with human motion schemata and the behaviour of objects of different weight on a subconscious level. The categorization in *grave* and *acute* accents is thus assumed to be conceptually relevant.**
- (23) **It is assumed that *grave* accents are more metrically prominent than *acute* accents. Thus, metrical groups are more likely to be counted from beats with *grave* accent, than from beats with *acute* accents.**

## 5.2 Metrical analysis – Method design

### 5.2.1 Outline

The two main tasks of the metrical analysis are (1) to separate non-metrical melodies from metrical and (2) to identify the metrical structure in metrical melodies. The former is of fundamental importance, since existence of metrical structure provides a primary structure to which gestural rhythm is assumed to relate, which is not the case in non-metrical melodies. Hence, the method of analysis of phrase structure is designed differently for metrical and non-metrical contexts respectively.

According to the central concept of multi-layered metricity there is a distinction between atomic, fundamental pulse and composite pulse. This distinction has bearing on the design of the method of analysis, for atomic pulse level is mainly dependent on interonset patterns, while composite pulse levels depend on periodical grouping, which in turn can depend on other musical dimensions (*Assumptions no 17-19*).

There is, however, yet another complication. Musical performance involves typically variations in event duration (see e.g. Gabriellsson 1983) which due to the assumed spontaneous search for periodicity can be disregarded during the listening process, allowing the listener to conceive meter even when strict periodicity do not occur. There is rather strong psychological evidence that quantization is performed by people while perceiving music (Fraisse 1956, Povel 1981, Lee 1991, Parncutt 1994 et al). Therefore, the input durations have to be quantized when the input originates from performance, in order to evaluate if the structure may be conceived metrically or not.

Consequently, the metrical analysis is performed in three steps: (1) *the metrical quantization analysis*; (2) *the beat atom analysis* (3) and then, if metricity is found and there is possibility of composite pulse levels, analysis of higher metrical levels is performed by the *complex metrical analysis*.

However, if the *beat atom analysis* results in a metrical interpretation at central pulse/tactus level, the analysis of higher metrical levels is performed through metrical interpretation of the result from the subsequent metrical phrase and section analysis. This implies that the complex metrical analysis is employed only when the beat atom analysis do not provide a result at central pulse level. The reasons for this technical procedure is that the *metrical phrase and section analysis* requires a primary analysis of metrical structure at central pulse level, since this level is assumed to be the fundamental temporal grid of the melody.

### 5.2.2 Metrical Analysis I - Metrical Quantization

#### 5.2.2.1 General concept

As mentioned above, it is important to consider how the model relates to unquantized input in order to evaluate it as a model of metrical cognition.

What is the position of the initial quantization in the process of meter perception and cognition? According to the general model of melody perception (see 3.2), the process starts by the analysis of local information within the time span of the perceptual present interpreted

by previous conception of melodies. This in turn gives rise to a global conception of the melody on a global and local level. This process is repeatedly going on through and after listening to the melody, a process in which the initial concept can be emphasized or reconceived.

The heart of the meter concept is periodicity. To experience periodicity, we need to perceive relationships between durations of events in terms of a common divisor, which needs to be perceived as continuous, i.e. continue beyond the scope of a perceptual present. In that way, the perception of a common divisor of durations is both prior to and a necessary requirement for the conception of meter. The problem regarding quantization is, however, that pertinent information for the conception of meter and phrase structure, inherent in the micro-durational variations is omitted.

The model is concerned with the lowest level of common time units, which, in practice, generally can be regarded as the level of beat division. However, if beat division is inconsistent, it may give a result at beat level or above.

It is based on the assumptions of the general concepts of meter described in section 5.1, with the additions of specification of absolute limits for significant duration difference, ornamentation, typical beat rate etc., that are described in Chapter 3, *The general model of melody cognition*.

The basic assumption behind the analysis of metrical quantization is that rhythmic relationships are evaluated sequentially within the perceptual present at the lowest level according to the categorical proportional rule. (see section 3.1 etc.)

### 5.2.2.2 Method design

The model of the process of quantization can be described briefly as follows:

#### ***A. Initial interpretation of rhythmic relationships between the first three events onsets and/or offsets***

The first relationships between three contiguous event onsets and offsets, i.e. the durations of two events, are timed to relationships that can be interpreted as a ratio of a common denominator. This interpretation is guided by a set of preference rules that assumes: (1) strong preference for simple ratios and duration equality; (2) preference for ratios under the upper category limit of the perceptual present (max 1:8 relationship based on the categorical present); (3) preference for divisions above the level assumed rhythmical significance; and (4) preference for a division which is consistent with preferred central pulse level. The latter is adapted from Parncutt (1994), and implies, as interpreted here, that the highest division must be categorically consistent with the limits of preferred pulse rate. This means that divisions of the shortest possible beat duration in more than nine parts are not considered, according to the categorical present.

Two events are considered equal in duration when their duration is more similar than dissimilar. For durations at division level (below the limit of preferred central pulse level) this level is depicted by the categorical proportion rule, which states that two durations are equal when they differ with less than one third of longest duration. Durations over preferred pulse rate must differ with less than 1/4 when a shorter denominator is identified. Duration differences below 150 ms are generally not considered as significant.

Durations shorter than 120 ms or pairs of events with a shorter total duration than 300 ms are generally treated as ornamentation and clustered into longer durations or if extremely short even omitted from the analysis. Ornamentation is also determined by pitch analysis. Repeated alternation between neighboring MPCs are treated as ornamentation when the total durations for two succeeding events are below 300 ms.

Rests (OOIs) are treated as articulation and included in the duration of the preceding note event if the duration of the rest is less than  $1/3$  of the preceding note or if the time difference is below 150 ms.

Significant differences in duration between successive events are timed in relation to the obtained common denominator into the nearest integer ratio.

### ***B. Continued interpretation of the next pair of events***

The common denominator of these two events is perceived to continue in the examination of the next pair of events, and this next time-span is considered in relation to the first if it falls within the time limit of the perceptual present. An event duration which extends more than the two preceding events is timed in relation to the neighboring events. When there is no acceptable common denominator between the closest successive events, events can be clamped allowing the total duration of successive events with a common denominator to be compared with the next group of successive events. The number of events considered at a time is confined to the time and category limits of the perceptual present.

When the division is less than half the typical lower preferred pulse, compound divisions (super-divisions) are continuously checked for, and when such a division is found to be unambiguous, the events within each group or super-division are timed in relation to the total duration of the super-division. In this way metrical hierarchy at division-level is implemented in the model, which is of crucial importance so as not to get out of phase in the analysis and can be said to reflect the time-frame of the perceptual present. Whenever such a super-division is found, it is regarded as a time-frame in which the tempo cannot be radically altered.

### ***C. Timing at super-division level***

If a common denominator persists new sets of events are timed in pairs, and in groups if super-divisions exist in relation to the existing common denominator and their internal difference in duration. The common denominator can be reconsidered when a durationally significant difference appears (in relation to preferred pulse rate and time discrimination limit). This is a feature which is primarily interpreted as a change of the general time denominator, secondarily as a start of a new time denominator or thirdly an inhibition of a common time denominator.

To summarize, relationships in duration of events are examined in pairs as to obtain a common denominator, which can express the durational relationship between the two events. Successive pair-wise relationships between at least three events are considered at a time, but more events up to the limits of perceptual and categorical presents can be included in this comparison. When the main division is very short, groups of divisions (super-divisions) can be considered in addition to the above. Simple relationships are strongly favored. Preference for equalization of durational difference is applied; the rule states that if two durations are more similar than dissimilar, i.e. the difference being less than one third of the total duration, they will be regarded as equal, as well as the rule of minimum time limit of duration difference of

150 ms as the core rules. Shorter notes are treated as ornamentation of notes, while short rests are treated as articulative rests.

This model share many common features with earlier models of quantization, e.g the meter program developed by Sleator and Temperley (2001), which obviously have been developed in parallel to mine. Their model, however, includes medium level metrical analysis, which is performed separately in my model. Above all it seems more adjusted to the analysis of metricity in multipart music, in the use of event simultaneity in the evaluation of metricity. To employ my method on multipart music, it would have to be adjusted for e.g. the problem of simultaneity between events in different parts.

### 5.2.2.3 Examples of the performance of the quantization method

I will present three examples of the performance of the method of quantization in order to demonstrate its abilities and shortcomings.

#### Vals efter Pål Karl



Figure 5-4. waltz after Pål Karl

**Example 1 Waltz with tempo changes**

One circumstance, which is crucial for the subsequent part of the metrical analysis, is that perception of tempo change are equalized by the quantization. This action assumes that we prefer to experience different durational density according to a variable common denominator rather than referring to an absolute value of the common denominator, which would demand new denominators when the original value is obsolete.

In this case, a Swedish folk music waltz performed in traditional style on MIDI keyboard was recorded as a MIDI file by a notation program. In the current example the tempo was slow in the beginning and increases towards the end, changing from ca 272 and 545 ms, which represents a tempo change ranging between 110 and 220 MM. Minor fluctuations in tempo was also common.<sup>43</sup>

The output of the initial quantization analysis can be regarded below:

Figure 5-5. *Quantization analysis of waltz after Pål Karl*

<sup>43</sup> The somewhat unusual default tempo setting (440 MM) is due to practical restriction regarding small note values in my application.

Chapter 5

This represents a typical melody with relatively low density in division, which, on the other hand, requires a great deal of generality of the method to take account for tempo changes. The method in its current version gets into trouble, however, when the tempo changes are too sudden, which can be seen in the following version of the same melody. The input version is here both quite sudden and extreme with regard to tempo changes. The beat ranges between a whole note (two half notes) in the beginning to only about a half note eleven notated measures later, it then turns back to an even slower tempo, only to speed up at the end in a way which is extreme even regarding performance. The resulting quantization makes one mistake, which is of course of crucial importance and would result in an erroneous metrical analysis.

This points to the limitations of this method in its current form. To be able to account for such extreme variability, assuming it would be perceived as tempo changes by listeners, the subsequent parts of the analysis would have to be run in parallel to the quantization analysis; successively evaluating the melody as suggested in the general model of melody perception (see above). Another possibility would be to allow alternative interpretations of critical situations which would be maintained through the stepwise analysis.

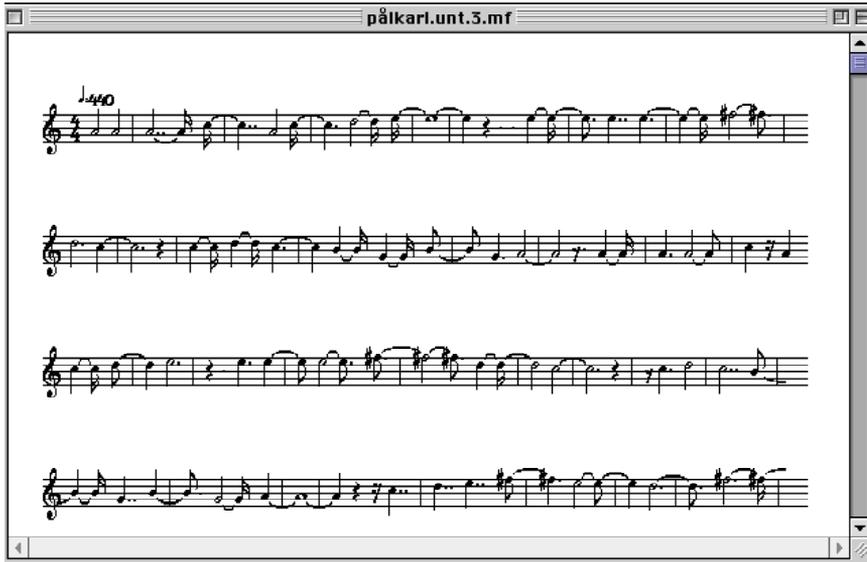


Figure 5-6. Untimed input, waltz after Pål Karl

In this case the final metrical interpretation on measure level cannot be obtained without involvement of the phrase analysis, i.e. implying analysis of pitch similarity at phrase level.

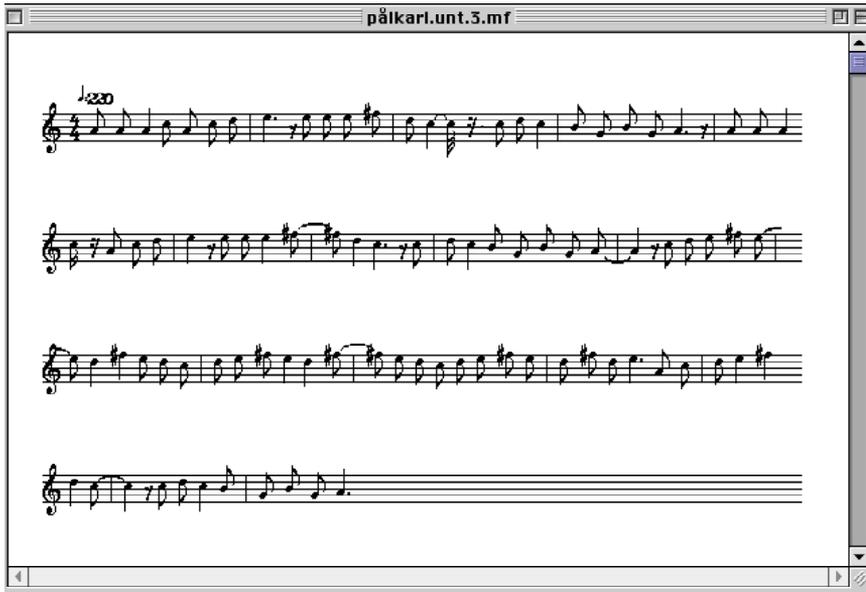


Figure 5-7. : Timed output, error in measure 2 (unrecognized change of duration). Waltz after Pål Karl

### Example 2 Significant changes in beat length and ornamentation in polska tunes

In *polska* tunes, especially from the western parts of Sweden and southern parts of Norway, it is common with significant but variable changes of length of beats in combination with rich ornamentation and diversity in terms of division. (Ramsten 1982, Ahlbäck 1986) These styles provide challenging examples for the current method of quantization.

One main problem here is to separate between ornamentation and division. To which division unit does a note, which can be classified as ornamented, belong? Does it belong to the preceding or succeeding unit? The way of handling this problem in the current method is to cluster ornamented notes but not to attach them initially to any other division unit, except for single notes, which are generally attached to the preceding unit. Another general problem arises as to when durational differences should be regarded as significant or not. In this style, the beat length at tactus level varies considerably and more or less periodically. This can be more easily demonstrated if we apply the subsequent levels of metrical analysis, the metrical beat atom analysis, the metrical grid analysis and the beat group analysis to this example. (see fig. 8 and 9).

As can be viewed in fig. 10, the timing of a parallel sequence of pitches in measures 3 and 7 respectively is different from each other. In the first case, they are interpreted as extending over two beat-atoms, while, in the second case, they are timed as occupying three beat-atoms. What is important for higher level analysis, the phrase and section analysis, is that they both equals one beat group, whatever the size. The variation of beat group length is, as can be seen,

Chapter 5

considerable, and, at the next to last measure, the quantization makes an quantization error, because of a sudden accelerando at the end, which causes the method to lose one beat-atom.

### Polska efter Spaken

from live performance

Figure 5-8. Untimed input file, polska after Spaken

Figure 5-9. Quantized output, polska after Spaken

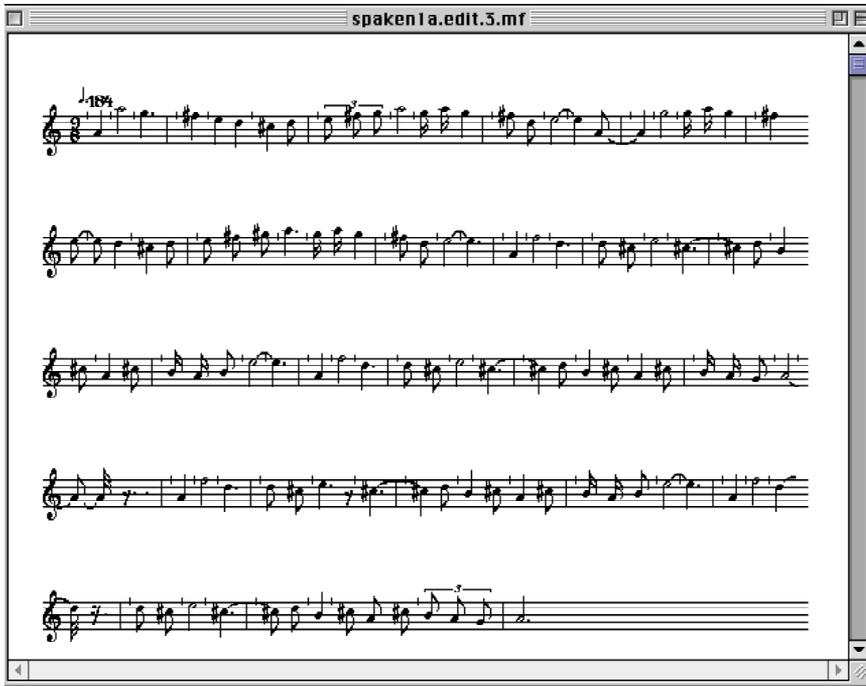


Figure 5-10. Output of metrical analysis. Polska after Spaken.

It is questionable whether the level of quantization here, which allows relationships between successive events as complex 3:4 / 4:3 when the denominating division is shorter, is too precise in its timing of events. However, the variation of beat length in this kind of music today is, within the culture, regarded as significant, even though it is not generally expressed in ratios, but rather in terms of short and long beats. (Ramsten 1982, Ahlbäck 1986). It would be perfectly possible to obtain an even more generalized description, which does not recognize differences at this level by changing the limit values.

The most hazardous points in the quantization process are the events of relatively long duration, which encompass more than one division unit, since tempo changes at these points cannot be easily tracked and the exaggeration and levelling of events of long duration is an important expressive tool (Bengtsson & Gabrielsson 1973, Friberg 1995). The problem of tracking these differences is related to the lack of onset input. The way of handling this problem in subsequent metrical analysis has been to allow a greater value of variation of periods when the beat-atom duration are at division level, which of course moves the problem to evaluation between different possible solutions. Since the consistency of a metrical pattern is one of the main factors in this evaluation, it has been a quite successful method.

Chapter 5

Yet another example is a polska tune originally played by the great fiddler Gössa Anders Andersson (1878-1963). Here, one finds the same kind of local tempo variation as in the preceding example, but here with even more elaborated ornamentation.

### Polska after Gössa Anders Andersson

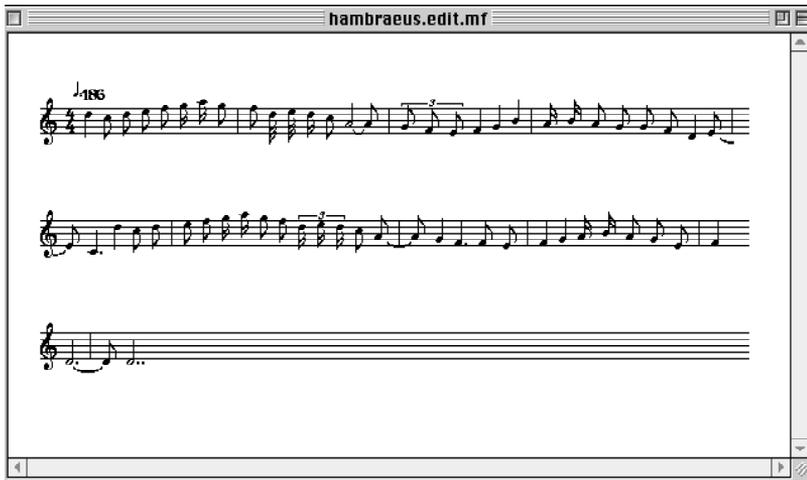
from live performance on fiddle converted to midi

♩ = 220



The image shows a musical score for a polska tune. It consists of four staves of music. The first staff begins with a tempo marking of a quarter note equal to 220. The music is written in a single melodic line on a treble clef staff. The notation includes various rhythmic values and ornaments, such as grace notes and slurs, which are characteristic of a fiddle performance. The score is presented as a MIDI file.

Figure 5-11. Untimed input file. *Polska after Gössa Anders Andersson*



The image is a screenshot of a MIDI editor window titled "hambraeus.edit.mf". It displays the same musical score as Figure 5-11, but the notes are now quantized to a regular grid. The tempo marking at the top is now ♩ = 186. The window has a standard interface with a scroll bar on the right and a status bar at the bottom.

Figure 5-12. Quantized output. *Polska after Gössa Andersson.*

As can be seen when comparing the first measure of the input file and the resulting quantization, notes of about the same length become strongly equalized division units in the beginning of the measure (notes 2-3), while they are classified as ornamentation in the end of the measure (notes 6-7). This classification is based on the concept of ornamentation as pitch variation (melodic return), which is true for the latter notes, but not for the former. It also relies on the quantification of groups of events; the notes 2-3 have a total duration which is greater than the limit value of an ornament.

## 5.2.3 Metrical Analysis II – Beat Atom Analysis

### 5.2.3.1 General concept

The beat atom analysis performs metrical analysis based exclusively on information of event duration, with regards to IOI and OOI intervals. It is thus directed towards metrical analysis on lower metrical levels (see assumption no. 17), specifically the tactus level. However, metrical hierarchy on all levels within the scope of the perceptual present is regarded.

The analysis is accomplished by searching the melody for periodicity in MIDI note *interonset intervals*, (IOI) i.e. the time-span from one onset to the next onset, and offset-to-onset intervals (OOI).

Since pulse is assumed to be tempo dependent, the tempo range is initially set between ca 10 (5625 ms) and 300 MM (200 ms). Even if tempi over 350 MM have been reported (see Björnberg 1996:75) and children's spontaneous tempi might range well over 300 MM (Hugardt 2001), such tempi seem to be exceptions to the general rule and are normally connected with a metrical hierarchy of higher metrical levels.

The value 5625 ms is the tentative value of the perceptual present used throughout the model. This value is considerably larger than the assumed limit for preferred pulse (see section 5.1.2 *Pulse*). However, since metrical hierarchy is considered to be influential for pulse conception within the scope of the perceptual present, higher metrical levels are limited to this value.

The chosen duration of the perceptual present is directed related to the chosen duration of central pulse period and vice versa. (see Chapter 3, *The General Model of Melody Cognition*, section 3.1 and *Assumption* no. 13 ). According to the findings of Parncutt (1994) there is a peak in pulse salience around 80-100 MM. A value around 100 is often mentioned as an average standard tempo in studies of spontaneous tempo (Hugardt 2001).

Given a tentative preferred pulse period duration of 625 ms (96 MM), the typical value of the perceptual present can be expressed as the maximum number of categories depicted by the categorical present (the  $7 \pm 2$  rule), which gives nine times the typical pulse period which makes 5625 ms. This value conforms to the most commonly noted estimations of the perceptual present.

For practical reasons, some limitations related to conventions within common Western notation are set, such as a maximum period equivalent to 9 quarters and a minimum beat length of  $1/32$ . These limitations are tentative, not essential to the method of analysis, and can be changed to meet the needs of another notation practice.

An important consequence of the assumption of pulse being conceived as the fundamental temporal grid of music is that pulse events are determined by the capacity and conditions given by short-term memory and can thus be viewed in the sense of a perceptual present (see assumption no. 13). We have to be able to relate to the temporal grid at every position in the music to conceive it as a temporal grid. According to the concept of the categorical present there is an upper limit of nine items ( $7 + 2$ ) that can be processed simultaneously, which may be interpreted as a maximum composite beat length and/or limit for obscuring beats. This is hence used as a constant for maximum group level and events obscuring beats.

### 5.2.3.2 Method design

The design of the method of basic (atomic) pulse tracking is as follows: Beat-atom refers to basic beat level. N.B. This description presents the method of analysis in terms of the procedure within the computer implementation of the method of analysis.

1. The default beat-atom length is set to the beat-atom length given by the maximum tempo (200 ms)
2. The maximum period value (max-beat-atom-val) is set to a constant maximum period (5625 ms), which equals the total duration of nine quarter notes at a tempo of 96 MM.
3. The increment value is adjusted to the minimum duration during the first 5.625" of the music
4. The first 5.625" of the music is preliminarily checked for minimum MIDI-note durations and standard values are hypothesized based on this check.
5. The central outer loop begins and subsequently checks occurrences of every beat-value up to maximum beat-value at every position determined by the increment value. Both the time from each starting-point to next event and the number of events to next position in the tested metrical grid are measured. Variables measuring rhythmical discreteness (r-cue-true and female-list) are set.

Then the grid position is checked according to the following set of conditions:

- Next event-start is on next grid-position.
- Next event-start is on  $n \times$  beat-atom-length, where  $n < 10$ , that is there are no more than nine beats to next event-start. This is based on the assumption that period has to be established within the perceptual present (see *Assumption no. 13*)
- Next grid-position is on a event-start but it is  $n$  event-starts before next grid-position (beat-atom-length), where  $n < 10$ , which means that there are not more than 9 events between two beats. This is based on the assumption that period has to be established within the perceptual present (see *Assumption no. 13*)
- When neither next event-start nor next grid-position is on event-start, the metrical grid is evaluated depending on what happens until the next coincidence between event-start and metrical-grid-position. If the intermediate time equals two beats or two or three event-starts its regarded as syncopation, i.e. temporary deviation from the metrical grid
- If the time to next event-start is less than the time to next grid-position, but the duration of the next event is equal to beat-atom-length, it denotes syncopation or a superimposition of pulses in **phase displacement**. This can include up to nine beats (see *Assumption no. 15*)
- If the time-span between current position and next coinciding event-start and grid-position is divided into a maximum of five 'super-beats' above three beat-atoms or three super-beats above four beat-atoms it can be regarded as **uneven superimposition of pulses**, hence within the metrical grid. (see *Assumption no. 15*)

- Other temporary deviations over three or four beats can, under certain circumstances, be allowed

If none of the above conditions applies, the loop is stopped and the grid is regarded as not valid.

When the grid is valid, the grid is tested for maximum- and minimum-division of beats respectively. The maximum limit is consistently nine events and the minimum limit –9 events according to the categorical present<sup>44</sup>. If the grid fulfils these conditions the values of the grid are pushed into an index-value-list.

6. Every valid grid is thus collected in a list and evaluated by a calculation of an total-index-value for each grid based on the following conditions:
  - The absolute onset support of the grid, i.e. the number of occurrences of events at grid-positions expressed as an increment of an index-value by the number of occurrences ((+ 1 (/ (third grid) 100))
  - The relative onset support of the grid, expressed as the grid-cnt-/cnt-ratio, i.e. the ratio between events at *grid-positions* and total number of possible *grid-positions* ((\* (/ (third grid) (fourth grid)) 100))
  - The temporal prominence of the grid, represented by the position of its start, giving an index value which favors earlier positions in the input file. The input file includes rests as well as notes.
  - The absence of serious onset deviations from the metrical grid, expressed as occurrences of extremely obscure syncopations giving depreciation of the index value (inverted value).
  - Durational support/prominence for inversion of the grid, expressed as occurrences of repeated short-long starts (“feminine” starts) giving depreciation of the index-value (inverted value). This is a consequence of the assumed prominence of grave accentuation as a metrical determinant (see Assumption no. .
  - The group-size consistency of the beats of the grid according to Millers’ Law expressed as max- and min-division of the beats, giving a depreciation of index-value for extreme figures (division-value)
    - a) minimum of five midinotes per beat (min-div) or a maximum number of notes per beat below 0 (max-div) gives a division-index of 0.5
    - b) min-div > 1 and max-div > 4 or  
max-div < 1 and min-div < -3 gives a division-index of 0.75
    - c) min-div > 1 , max-div > 4 and (max-div – min-div) > 2 or  
max-div <= 1, min-div < 0 and (max-div – min-div) > 2 gives a division-index of 0.8
    - d) max-div < 3 or min-div > 1 or max-div > 6 and min-div >= 0 gives a division-index of 0.8333... (5/6)
  - The tempo of the grid in relation to ideal pulse tempo, expressed as depreciation of the index-values by extreme tempo-values:

---

<sup>44</sup> This is, however, probably an unnecessary high number, since the period has to be established within the perceptual present, which means that at least two beat impulses must fall either within the range of the nine elements or within the possible time limit of the perceptual present 5-6 “. Ideally the periodicity would be established within this range, and would give an ideal number of two

- a) tempo > 240 or < 30 gives a tempo-index of 0.5 (beat-atom-len-val)
- b) tempo > 160 or < 50 gives a tempo-index of 0.75
- c) tempo < 75 gives a tempo-index of 0.8
- The ratio of onset dissupport of the metrical grid, expressed as occurrences of obscuring of the metrical grid, such as uneven or displaced pulse superimposition gives depreciation of the index-value based on the symmetrical dominance assumption (see *Assumption* no. 7).
  - a) If the obscured beats are more than 2/3 of the number of events at grid-positions the obscure-cnt index is set to 0.5
  - b) If the obscured beats are more than 1/3 of the number of events at grid-positions the obscure-cnt index is set to 0.75
- The rhythmical support and prominence of the metrical grid is expressed through occurrences of rhythmically discrete grid-positions, i.e. supported by change in rhythmic duration at grid-positions, which increases the index-value by its inversion

These different index-values are multiplied and result in a total index-value.

7. The evaluation of the grids is then performed using both the index-value and hierarchy of beat-atom-grids, under the assumption that a consistent metrical hierarchy supports to a metrical grid.
  - Metrical ambivalence is checked through the search for metrical grids with incompatible beat-length but identical start-position as the grid with the highest index (first-choice). When there is a competing metrical grid with higher index-value than 2/3 of the index-value of the highest ranked grid (first-choice) and a grid with shorter beat which is compatible with both the competing grid and first-choice, and either the max-division or the tempo of first-choice is extreme (> 4) or it doesn't exist any grid with a beat-length which equals half the beat-length of first-choice, then the competing grid is chosen instead of first-choice.
  - Then the grids with compatible beat-lengths and starting-positions are collected in a list, which is used to check potential metrical hierarchy, and evaluated on basis of the index-value and hierarchical conformity and added to the first-choice list.
  - If the metrical grid with the highest index has a tempo below 37 MM or encompasses more than two standard beats ( (\* (truncate max-beat-atom-val 9) 2) ) with a max-division > 2, this is regarded as an extremely low tempo/extremely long beats and its index-value is reduced to 1/4 of the original value and the first-choice list is resorted according to index-value, which may result in a new first-choice.
  - If the grid second in rank has half the beat length as the highest ranked grid, > 4/5th of its index-value, with high event-start/grid-ratio and acceptable division and tempo-values, the rank is changed between first and second ranked grids. This precaution is taken so as not to obscure possible asymmetrical composite meters, which are not included at this level of analysis.
  - As another precaution, a second metrical choice is set to the first metrical grid with start-position different from first-choice and with compatible

beat-length with the minimum beat-length of the first-choice hierarchy. If such a grid does not exist second-choice is set to the first grid with equal beat-length as the highest ranked first-choice but different start-position. Finally, if there is no such grid, second-choice is set to the first grid with the same length as the second ranked grid in first-choice.

8. The result of this evaluation, first-choice and second-choice, is returned to the main function and evaluated regarding the need of further analysis.<sup>45</sup>

---

<sup>45</sup> Change of basic/atomic metrical grid, such as a sudden change of tempo within the melody or another change of metrical grid which results in two incompatible grids, with no common beats of the same tempo in the same phase at any level, is not yet implemented in the program.

The reason for this is that it is not needed at present to manage further structural analysis in the main body of melodies which is relevant to this work. It would, however, be easy to add this feature to the current method simply by continuing the search for metrical grids throughout the melody when a given grid is discriminated by lack of feedback from the melodic structure.

## 5.2.4 Metrical Analysis III – Complex Metrical Analysis

### 5.2.4.1 Overview

If the *Beat-atom analysis* results in an atomic pulse level, the method continues to evaluate complex metrical levels. Here the analysis is based on the assumption that more complex metrical levels are created through difference in accentuation of beats/events, groups being defined as spanning from one accentuated beat/event to another. To perform this analysis, an initial successive evaluation of the accentuation of the events of the melody is carried out.

However, in many cases such a complex analysis is not needed at this point, since the atomic/basic pulse analysis often result in a pulse which is at a typical central pulse or tactus level (see ass no 13), where there are shorter and longer *interonset* intervals in accordance with the categorical present. Then more complex metrical grouping, such as periodicity at measure level / time, is dependent on phrase analysis and thus will be discussed in that context. (see Chapter 6, *Metrical Phrase and Section Analysis*)

There are, however, certain metrical structures for which the beat-atom analysis is not sufficient, such as: (1) Melodies in which the melodic structure consists of many events of equal length, as in a *Moto Perpetuo*; (2) melodies with *interonset* aperiodicity on a complex level, as in melodies with additive rhythm such as in Balkan dance-music; (3) or superimposed pulses of incompatible beat length as in polymetric music such as much African music. For such metrical structures the result will be an atomic pulse on division level which does not have the category requisites of a complex pulse, i.e. without sufficient division of beats or too short period to be a denominator of long *interonset* intervals within the melody structure according to the categorical present.

In such cases *complex metrical analysis* is required to detect a pulse level, which fulfils the demands of a central pulse level, which is in turn needed to perform the phrase analysis. The following chapter applies only to melodic contexts in which the analysis of the atomic pulse level results in a pulse period which is sub-division of the central pulse level, and thus should be grouped.

Here is a central concept of the current model introduced, namely the method of analyzing melodic segmentation by melodic similarity (parallelism) and melodic discontinuity. (see Chapter 3, *General model*, and below). Through the entire method of analysis of melodic structure, similarities are assumed to be superordinate (prominent) to discontinuities. In practice, this means that the melody is first searched for *short* sequences of pitch- and duration patterns (according to *the categorical and perceptual present*). The starting-points of the groups of these “*strong*” *sequence*’ are given numerical values to reflect the assumed group accentuation. Local discontinuities i.e. changes of pitch, duration etc., are then analyzed, which also results in numerical values reflecting assumed accentuation.

On the basis of these values, possible composite metrical levels are calculated, all of which are evaluated on the basis of consistency, symmetry, salience/*Prägnanz* and total values of accentuation based on segmentation by melodic sequencing and discontinuity. As with the atomic beat analysis assumptions of ideal pulse rate are also used in this evaluation.

Through this analysis a definite analysis of pulse, levels of pulse and a preliminary analysis of time (measure level periodicity) is carried out. (Observe that the latter through its inter-

contextuality, i.e. dependency on higher grouping levels, is dependent on section- and phrase-analysis and thus must be reanalyzed at a later stage in the structural analysis.) This preliminary analysis can result in duple or triple time, but also in compound time with asymmetrical grouping. This is demonstrated by the transformation of the input score (fig. 13) from its notated pulse and time-notation, to the calculated metrical structure displayed in the output score (fig. 14). In the output are beats (grouping of events into beats) at intermediate tactus level separated by inserted hooks at the top of the system. The preliminary central time is displayed by the time-signature and inserted bar lines as is shown in figure 14 below.

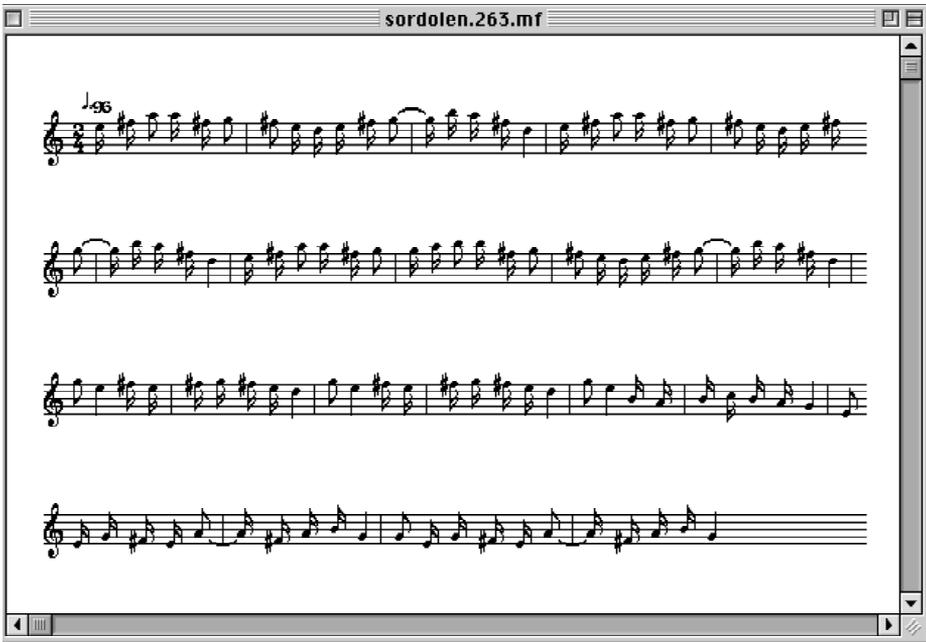


Figure 5-13. Input score adapted before metrical beat atom analysis of the tune "Sordölen", played by Eivind O. Hamre, Bygland, Norway adapted from a transcription by V. Lande (Lande 1983). Chordal notes, ornaments and micro-tonal alterations omitted. The metrical interpretation in the original notation has been altered.

Figure 5-14. Output of Beat atom analysis of “Sordölen” after Eivind O. Hamre. Beats are designated by hooks at the top of the systems between each beat. Observe the change of metrical interpretation at measure level in relation to the input file.

#### 5.2.4.2 Analysis of low-level grouping – general principles

The first step of the *complex metrical analysis* is the search for ‘short’ or low-level pitch-contour and rhythmic sequences.

Low-level grouping, refers to *primary grouping* of individual events. The significance of low-level groupings in the context of metrical analysis is that group starts can be perceived as impulse points or structural accents. If the group starts occur periodically within the span of the perceptual present, they are assumed to imply a metrical conception. (*Assumption* no.20). The point is that such sequencing can function as metrical accentuation, as accentuation of beats, hence the rule of recurrence within the perceptual present applies for metrically perceived primary sequencing.

Since grouping and meter are generally regarded as independent qualities, this assumed linkage between grouping and meter requires further discussion. In the fundamental outline of this theoretical model of melody cognition (Chapter 3, *General Model*) I am distinguishing between *gestural rhythm* and *meter*, where gestural rhythm is intimately linked to melodic

segmentation, such as motifs, figures, phrases and sections. As was discussed in section 5.1.2 some researchers in music theory and cognitive psychology (e.g. Lerdahl & Jackendoff 1983, Clarke 1999) from a similar distinction, makes the assumption that the cognition of meter and grouping are not interrelated. However, this is problematic in relation to higher levels of metricity, such as time, which can be regarded as periodicity implying grouping of beats; Regularity in grouping can according to the current proposal be conceived as implying metrical impulse, thus the group can be regarded as implicating metrical unit, a period. Thus, grouping and meter can be interlinked, periodicity regarded as a quality of grouping, while if we focus on the gestalt quality of individual groups grouping and meter can be regarded as two independent structural qualities. The results of experiment 2, trial 4 (See section 5.3.4.3) demonstrates this linkage between low-level grouping of events and metrical conception. In this experiment subjects extracted time and beat-level metricity from pitch structure.



Figure 5-15. Metrical levels

It is here possible to experience at least three metrical levels, all of which seem to be apparent to most of the subjects participating in the test. One question to be raised regarding this result is whether the intermediate level, indicated by beams, can be regarded as a metrical level, as a pulse.

The metrical preference and wellformedness rules, proposed by Lerdahl & Jackendoff (1983:68ff) and implemented in a computer model by Temperley (2001), excludes every deviation from isochronocity from the concept of meter. However, as is acknowledged by Lerdahl & Jackendoff (1983:97), this is problematic in at least two important respects:

- (1) People seem to be able to perceive a regular pulse even when it is not phenomenally isochronous, which is apparent e.g. in many studies of deviation from isochronicity in performance of notated music (e.g. Bengtsson & Gabrielsson 1977, Gabrielsson 1983). Thus, perceived periodicity allows perceptible deviations from isochronicity.
- (2) There are frequent examples of musical styles where the time of certain dance music types are conceived as combinations of short and long beats. (see e.g. Singer 1974) The most well known example of this to Western listeners is probably dance music from the Balkans and Turkey, where systematic alternations between short and long beats corresponds to the dance steps. Thus *beats* are connected with *steps* in a regular way, which is analogous to the correspondence between beats and steps in other dance types. This phenomenon, perceived regularity in deviation of beat-length, can be regarded as “*quasi-periodicity*” or *metrical asymmetry*, which relates to isochronous or

symmetrical meter in a way analogous to the relationship between limping and walking.<sup>46</sup>

Hence, this is an area where melodic segmentation (phrasing, gestural rhythm) and metrical structure concur. We can conclude that conception of metrical structure can be drawn from low-level melodic segments, which can be regarded as the primary grouping of individual melodic events. From this example we can also conclude that in a metrically conceived melodic context the fundamental grouping can be in concordance with the perceived metrical structure (beat-groups). My interpretation of this relationship is that primary grouping within the typical timeframe of pulse conception, in metrically conceived melodic contexts has to be congruent with perceived metrical structure.

In their influential work, Lerdahl & Jackendoff (1983) take a quite radical view in separating grouping and metrical structure on every level. It is demonstrated in a passage from the beginning the 40<sup>th</sup> symphony by W.A. Mozart, with metrical structure according to Lerdahl & Jackendoff displayed as dots below the system and the lowest grouping level showed by brackets above the system.<sup>47</sup>



Figure 5-16. Metrical structure and grouping structure in the opening of Mozarts G-minor symphony K.550, first movement

This demonstrates a low-level grouping that starts on an unaccentuated beat to encompass an accentuated beat. The segmentation proposed by Lerdahl & Jackendoff can be drawn from durational distance and proximity; notes close in duration are grouped together, while the metrical accent can be derived from durational accent.

If the metrical structure was derived from conceived grouping, as suggested above, this would result in a metrical structure which would not concur with what can be assumed to be the general conception of metrical structure in this piece, given the grouping structure proposed by Lerdahl & Jackendoff. If so, how can then grouping and conceived impulse points interact?

In another influential work, the aforementioned “The Rhythmic Structure of Music”, Cooper & Meyer provides an example of how they regard phenomenal accent to be metrically determinative.

<sup>46</sup> The conceived regularity of limping and walking is semantically inherent in these words, since they are plural by definition. To take one step is not to walk, rather it refers to recurring action. This analogy is apparent in e.g. the Turkish ‘aksak’ meter, which literally means ‘limping’.

<sup>47</sup> It is placed under the system in the original example. Higher grouping-levels are omitted.



Figure 5-17. Variation of “Twinkle twinkle, Little Star”. Example demonstrating how the accent structure, determined by durational accent, changes the meter according to Cooper & Meyer (1960:19-20).

“Notice that though the original barring of the tune has been maintained in Example 16, the change in rhythm and the resulting change in melodic emphasis have created a change in metric organization. The opening tones should now be written as upbeats to the D, as in Example 17.” (Cooper & Meyer 1960:20)

This example is problematic for me, since the change in rhythm does not result in any change in my metrical conception of the tune. The relationship between rhythmic and metric structure implied in the *A* version is on the contrary typical to many tunes. Another well-known traditional folksong, Jingle Bells, has exactly the same rhythmic structure in the opening two measures. Still, notation and lyrics implies that a metrical interpretation of the rhythmic structure as was suggested by Cooper & Meyer would be inconsistent with the general conception of the metrical structure:



This example shows that even though the song “Twinkle, Twinkle, Little Star” is well-known to both Cooper & Meyer and me, we can interpret the metrical structure differently from the same structural cues. This is evidently true also for grouping structure, as is demonstrated by the results of experiment 3 (See Chapter 6.3, *Evaluation of Metrical Phrase and Section Analysis*) and by similar tests performed by Höthker et al (2002a-c).

The only way I can change my metrical conception of “Twinkle, Twinkle, Little Star” in Cooper & Meyer’s variation is to speed up the tempo considerably. How can this be explained?

My interpretation of this phenomenon is that the long duration changes its function in my conception. At a slow tempo it will function as a closure of the melodic motion, determined by greater durational distance to the next event than to previous events, which makes the subsequent note an impulse-point. At a fast tempo it changes its function to become an impulse point, determined by durational salience, hence changing my conception of metrical structure.

My conception of the grouping at a slow tempo would consequently concur with my metrical conception. Hence, grouping can imply metrical accent when the group boundary is conceived as an impulse – as an accent. According to this concept groupings can be classified as either start- or end-oriented, where a *start-oriented* grouping interpretation has concurrent

group boundary and accent, while an *end-oriented* grouping interpretation may include an upbeat to the virtual starting point.

Returning to the Mozart example given by Lerdahl & Jackendoff, it is perfectly possible to conceive the grouping structure at primary level in a different way than have been suggested by them. Given the high tempo which is typical of most interpretations of the Mozart G-minor symphony I can easily conceive it the other way around., reflected in the following notation:



Figure 5-18. *Alternative grouping interpretation of opening of the Mozart G-minor symphony.*

Which grouping is right? Grouping by durational salience or grouping by durational proximity? This is obviously a meaningless question since the two parameters coexist in the conception of the groups; one binding the short notes to the long, the other creating a grouping between durationally salient, hence accentuated notes.<sup>48</sup>



My own experience, referred above, when trying to conceive the metrical structure in “Twinkle, Twinkle, Little Star” as predicted by Cooper & Meyer, as well as the evidence referred in section 5.1 suggests that durational salience is a stronger structural factor when the period between the accents is relatively short. This can be regarded an extension of the concept of interonset interval, which implies the perception of a tone and a rest as a singular event, the offset-onset interval being insignificant. In this sense, an interonset can be regarded a reduction of two events into a single gestalt. At shorter periods, it may be easier to regard short or non-accentuated events as insignificant.

The difference between the two grouping interpretations of the Mozart excerpt can also be described in terms of the previously suggested accent qualities, *grave* and *acute* accents. (see section 5.1.4). The start-oriented interpretation, beginning on the note of longer duration, can be regarded having the quality of a grave accentuation.

This distinction is very important for this analysis, since the goal of the metrical analysis is to extract the metricity from the analysis of grouping indications. Hence, all primary groups of metrical significance are here considered to start on the most ‘gravely’ accented event in the group within the typical level of beat conception.

I am assuming that this level of primary grouping of events is governed by general cognitive principles of pattern recognition, that can be expressed in terms of the Gestalt laws (see below). The model assumes that we are tracking periodicities on the basis of input from primary melodic segmentation including interonset patterns, which is in turn affected by

<sup>48</sup> Interestingly enough, Lerdahl & Jackendoff suggests the term metrical grouping: “*But once metrical structure interacts with grouping structure, beats do group one way or the other*” (1983:28)

perception of periodicity. This relationship can be expressed in terms of statistical support for and undermine of a pattern of periodicity: When a metrical pattern is more supported than undermined by primary melodic segmentation/grouping, metrical grouping controls primary melodic grouping by other principles.

### 5.2.4.3 Principles of low-level grouping

If we accept that low-level sequencing can function as metrical accentuation, it follows that metrically conceived groups should ideally be as 'short' (in terms of duration and number of elements) as possible. The shorter, the more recurrences of a certain period within the perceptual present, the more salient the metrical grid becomes. This notion could be considered as a reason behind preference for the conception of primary groups consisting of two or three elements. For example, a group length of three elements allows three groups within maximum nine elements of the categorical present.

Thus I assume preference for primary grouping in groups consisting of two or three ungrouped elements. There is psychological evidence pointing in this direction, that subdivisions of larger groups in subgroups of two and three are favored (see e.g. Clarke 1999:474).

- (24) **Metrical grouping concerns primarily low-level, primary grouping of elements within the limits of the perceptual and categorical present**
- (25) **Shorter groups are preferred to longer. This applies for the total duration as well as the number of consisting elements.**
- (26) **Primary grouping in two or three elements is preferred. Conception of larger groups as composed of subgroups of 2 or 3 elements is preferred.**

A fundamental hypothesis in this method is that the principles which govern the experience of grouping as well as any aspect of structure can be reduced to just one principle, one question – same / not same. This contradictory pair can be formulated as other more specific contradictories along different parameters, with different and important applications for melodic structure, similarity – dissimilarity (same/not same shape/qualities), continuity – discontinuity (same/not same stream) and proximity – distance (same/not same location). The basic idea in this method is that primary grouping is dependent on perception of either the positive or the negative side of the basic principle, that elements perceived as 'same' are grouped together and while a perception of 'not same' indicates group boundary. The first case refers to grouping by similarity, continuity and proximity and the other to grouping by dissimilarity, discontinuity and proximity. In other words the positive and negative sides of same/not same. In the case of melody, similarity is often used as a contextual grouping principle, meaning that grouping is conceived when the whole passage is heard or when a dissimilarity is reached. Discontinuity, on the other hand, refers to the change of a 'same' conceived as a continuum. It concerns mainly the local context, the change creates a border between groups, it is the structural break, the 'not same' which is in focus.

On the basis of the above distinction as well as experiments (see section 5.3.4) I have elaborated a set of gestalt theory-based principles for melodic grouping. They are ordered in three groups primary, secondary and tertiary grouping principles, which is meant to reflect their role in the process of grouping: The primary grouping principles are thought to create

grouping, the secondary to infer grouping (dependent on earlier creation) and the tertiary grouping principles refers to grouping prominence.

Primary grouping principles

**(27) Grouping by similarity, continuity and proximity – 'sameness':**

**Similar and close events tend to be grouped together.**

- repetitions of singular pitches, intervals or repetitions of durations are grouped
- similar, continuous or close elements with regard to pitch, melodic direction, interval size or duration are grouped
- sequences, i.e. arrays of pitch- and/or change of duration, which are repeated create grouping

low-level grouping priority:

Sequences are dominant and overrule local repetitions; repetitions overrule local similarity. Sequences which includes both durational changes and pitch change, overrule duration sequences, which in turn overrule pitch sequences.

**(28) Grouping by discontinuity, dissimilarity and distance – 'difference'. Change / difference indicates group boundaries**

a) dissimilar and distant events indicate grouping boundary (non-sequential)

- distant, dissimilar or discontinuous events with regard to pitch, melodic direction, interval size or duration are not preferred to be grouped together

b) change (sequential difference) in structure indicates grouping boundaries

- change of pitch set
- change of melodic direction
- change of interval size
- change of onset-offset - tone/rest
- change of duration/*interonset* interval

low-level grouping priority:

Change (sequential difference) is dominant, and overrules non-sequential difference. Change of pitch set is overruled by change of melodic direction, which is overruled by change of pitch interval size, which is overruled by change of tone-rest, which is overruled by durational changes.

Secondary grouping principles

**(29) 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**

- A sequence of repeated grouping creates an expectation of the same grouping to continue. A strongly expected grouping can be inferred. (metricity)

**(30) Grouping by symmetry. Symmetrical grouping is more prominent than asymmetrical grouping and can be implicative in the case of grouping by hierarchical symmetry.**

- Groups which are symmetrical, i.e. equal or similar regarding duration or number of elements contained, are preferred to asymmetrical
- Groups which contain two (or multiples of two) elements are prominent before groups containing three elements. Duple division is prominent to triple division.
- If the established grouping exhibits symmetry, symmetrical grouping can be inferred.

Tertiary grouping principles

**(31) Grouping by perceptual prominence / foreground, i.e. articulation, impulse and gravity: Grouping between most articulated events is prominent.**

Articulation refers here to articulation by pitch or duration, since this the two sole parameters in the input. This rule has special significance for the interpretation of group start (phase of sequences etc).

- A periodic/ metrically perceived articulation within the limits of the perceptual present indicates start in relation to its gravity and group-length (see above section 5.1.4).
- A non-metrically perceived articulation indicates primarily ending (see also proximity-distance rule above), in relation to its gravity and group-length.
- Onset is more articulated than offset, start on tone is preferred before start on rest

**(32) Grouping by integrity/'prägnanz' / contrast: Discrete grouping is prominent.**

- Inner '*prägnanz*': groups with a salient shape, which contain a few contrasting elements are more prominent than groups with low inner contrast or many divergent elements.
- contextual '*prägnanz*': groups which are unique, contrasting, discrete in relation to the surrounding context are prominent.
- contextual '*prägnanz*' is dominant.

The complementary nature of the two primary grouping principles leads to a possibility of using either the positive or the negative side of a structural dimension as an indicator of segment border. But how useful a certain measure is varies for different dimensions. A sequence represents order similarity, adjacent sets of similar pitch and/or *interonset* changes, which can provide precise indications of segment border. The opposite to sequence is lack of recognizable order and non-adjacency, which is not useful as a means of tracing segment boundaries. The converse holds for discontinuity in relation to continuity, where local structural discontinuity is a more useful measure to determine segment limitations than the measure of local similarity.

From this we can draw the principle that more unambiguous, precise, delimiting indications are superior to ambiguous, vague indications. But then, what is superordinate, 'same'-based indications as sequence-start or 'not same'-based indications as discontinuity?

While sequences can override discontinuities as structural delimitations, through the possibility of incorporating discontinuities within sequences, sequences are superordinate to

discontinuities. On the other hand, the interpretation of sequence-start is dependent on discontinuity, while discontinuities are superordinate to sequences as delimitation indications.

This relationship is evidently a matter of structural context; sequences are the dominant factor in large-scale context, globally, while discontinuity dominates locally: If the integrity and identity of the sequence is preserved, the sequence can incorporate a discontinuity which would locally override the sequence. Thus, the influence of grouping by sequence is dependent on the length of the sequence in relation to the discontinuity. Hence, regarding primary grouping, which is local, discontinuity is superordinate to sequence as delimitation factor.

However, since delimitation by sequence supported by discontinuity is superordinate to delimitation as such locally, there is a substantial local impact of indication of segment border through sequence.

- (33) **More unambiguous and precise structural indications of group boundaries are superordinate to structural features which give more ambiguous or vague indications of group boundaries**
- (34) **Indications of group boundaries which involve more elements are superordinate to structural features which involve less elements.**
- (35) **Combinations of different indications are superordinate to single indications**

#### 5.2.4.4 Implementation principles

The implementation of this hierarchy of principles in the method of analysis is performed in three steps:

1. Segmentation by sequences (similarity), evaluated by discontinuity/change in structure, secondary and tertiary grouping principles.
2. Segmentation by structural discontinuity/change, evaluated by secondary and tertiary grouping principles.
3. Segmentation by metricity (good continuation) and symmetry.

These order of the above three steps of the analysis is due to the nature of the principles. Most important, by sequences can include changes in structure, which in isolation would be perceived as group boundaries. Segments created by sequences, successive repeated patterns of duration- and/or pitch-changes, are contextually dependent, and hence must be analyzed before analyzing phrase boundaries dependent on change. Segmentation by metricity and symmetry requires input from primary grouping and hence must rely on prior analysis of primary grouping.

In real-time melody experience, these three levels of group perception are assumed to be performed simultaneously, so that initial grouping by structural change can be inhibited by grouping by sequencing, which in turn can be inhibited by segmentation by good continuation (metricity). For practical reasons, this is implemented in the computer program in three separate steps, where sequences which can overrule grouping by change are tracked, before tracking group boundaries by structural change. In a third step, metricity and symmetry are tracked and evaluated.<sup>49</sup>

---

<sup>49</sup> It is important to note that this methodology does not influence the possibility to evaluate the model of melodic perception, since the interrelationship between the different steps is defined.

On the primary level, duration (*interonset*) changes and patterns are assumed to overrule pitch change patterns (see description of the principles above). This is supported by various psychological findings (see e.g. Deliège 1987). It reflects the fundamental principle of priority of *interonset* information in primary melodic grouping. ('Note on' is thought to be the primary source of melodic information).

In summary, sequences of pitch and duration change overrule sequences of duration change, which overrules pitch change sequences. This analysis of sequences is for practical reasons performed in three separate steps.

#### 5.2.4.5 Segmentation indication by sequences – method design

##### Overview

Analysis of duration and contour sequences is performed by three different functions, which are run in succession: *rhythm-contour-sequence-search*, *rhythm-sequence-search* and *contour-sequence-search*.

The purpose of this stage of analysis is to reveal primary sequences, which can overrule local boundaries determined by change in melodic structure, here designated strong primary sequences. Thus, they must incorporate a minimum of two changes of pitch and/or duration, i.e. three events.

It is important to note that the analysis at this stage is about **indications** of sequences which *can* possibly overrule other local boundaries. The indications of sequences have to be weighed together with other indications of primary grouping at a later step in the metrical analysis.

The analysis is performed in two steps. First, note-starts are successively tested for sequences consisting of up to nine events (excluding articulatory rests) and minimum three events. When similar changes of structure are detected at two points the possible sequence is tested for similarity according to certain conditions, which are different for the three different types of sequences.

The conditions for strong primary sequences are summarized below:

##### Conditions for strong primary duration sequences

- a) The groups must contain at least three elements
- b) The groups must be either identical regarding duration (*interonset*-values) or have duration changes that are categorically unambiguous (difference within categories smaller than between categories).
- c) The groups must contain at least two different duration values.
- d) The groups must have a discrete set of duration values.
- e) The groups must contain only tones (articulatory rests excluded) but the last event of the group.
- f) The groups must be adjacent.
- g) The start- and end-positions of the groups must be discrete regarding either duration or melodic contour.

##### Conditions for strong primary contour sequences

- a) The groups must contain at least three elements.
- b) The melodic direction must be identical within the groups.

- c) The interval relationship (at large) must be preserved between the groups, i.e. if the first note in the second group is higher than the first note in the first group all correspondent notes must have the same relationship.
- d) Within the sequence the groups must be discrete in the sense that the interval, pitches or the change of direction *between* the groups and either *before* or *after* the sequence, are unique in relation to the intervals, pitches or changes of direction *within* the groups.
- e) The sequence must be discrete in the sense that at least two of the changes that constitute the groups of the sequence have to be different from each other.
- f) The sequence must not be overlapped or “disguised” by other strong grouping indications, as overlapping sequences, melodic return, repeated tones etc.

The overall principle behind these conditions is the integrity of the pattern, which is determined by how discrete the sequence is expressed in terms of the difference between changes within groups and between groups and the surrounding structure.

The design will be exemplified in more detail below.

### **Duration plus Contour – sequences**

Duration- and contour-sequences are examined in three main steps:

The sequence is first checked for similarity regarding duration-change, melodic direction and corresponding interval relationship between groups. As mentioned above identical duration-change, identical pitch direction (up, down, same) and identical pitch direction between groups is required. The number of duration categories within the group is measured, determining the rhythmic complexity of the group. (see below)

In the second step the ending of the sequence is checked for duration integrity. If the duration change at the end is not discrete in relation to the sequence, the starting point of the sequence must be examined further.

The third step is an evaluation of sequence-start, which is applicable only for sequences with a rhythmically non-discrete ending. This is designed differently for simple and complex sequences respectively. Complex sequences are assumed to be more discrete, since they are determined by a greater number of different unique elements according to the ‘*prägnanz*’ law. Simple duration sequences, consisting of only two duration categories, are considered as more unstable and require additional information from pitch contour.

Sequence-start is for simple sequences (two duration categories) examined through a set of conditions ordered according to the number of short durations preceding the longer duration. This is due to a central assumption of start- and end-orientation of primary groups respecting, namely that the tendency to perceive a long note as start and not end of a group is inversely related to the number and total duration of short notes in the sequence. The greater the number of short notes in a sequence with one long note, the greater the tendency to regard the long note as marking the end of the sequence. A long note in connection with only a few short notes is more likely to be perceived as an accentuated note among non-accentuated, hence indicating start of sequence. Likewise, a sequence starting on a long note succeeded by short notes is regarded structurally stable if the total duration of the shorter notes is less than or (if certain other conditions apply) equal to the longer note and the total number of elements is less than six.

The basic features of this analysis can be summarized as follows:

- Preference for *start on long note* ('strongest marked element') in sequences with few elements, such that if the short notes are less than four with a maximum total duration which equals the total duration of the long note *start on long* is preferred.
- Conversely the preference for *start on short before long* is related to the number and total duration of short notes in relation to long, such that when the number of notes either is larger than three, either has a total duration larger than the duration of the long note *start on short notes* is preferred
- Preference for grouping according to duration and pitch proximity/continuity and hence preference for *start after distance* regarding duration and pitch proximity/continuity. The strongest factor regarding grouping by pitch proximity/continuity is immediate repetition, after which comes melodic return, pitch change continuity etc.
- Preference for discrete groups regarding pitch and duration. Durational integrity can compensate for pitch integrity and vice versa, when both dimensions are involved.
- Avoidance of (negative preference for) ungrouped, single elements.
- Preference for symmetrical grouping, both regarding duration and frequency of occurrence.
- Preference for grouping/subdivision in 2 elements before larger groups.
- Conversely avoidance of start on locally unmarked element (by pitch-or duration-change).
- Preference for start on longer groups (3 elements) rather than shorter (2 elements), 3+2 is more stable than 2+3 depending on number of elements involved see above preference for *start on long note*
- Preference for grouping in good continuation of previous grouping and symmetrical grouping indication.

This is implemented through a set of rules which are different for simple and complex sequences and different for each number of short notes involved.

Some aspects of grouping by pitch proximity/continuity is given certain significance, such as grouping immediate pitch repetition and melodic return, which can invoke durationally unmarked elements to be sequence-start. Pitch and duration discontinuity/distance within two elements from sequence-start is also especially significant, since it determines whether initial grouping will be recognized. In the case of duration-sequences this is more important than the pitch integrity of the sequence.

Example: Simple sequences

- a) Sequence-start at melody-start, on first note after rest or immediately succeeding established sequence starts on first element
- b) Sequence which cannot be expanded starts on first element, except for *start on long* when  $\text{long} < 1/3 \text{ sequence-length (total-duration)}$  or  $> 5$  elements
- c) Sequence with one short before long 

Preference for *start on long*, *start on short* before long only when rep. to long, discrete pitch register, alt. defined group ( $> 2$ ) after long.



Chapter 5

d) SSL-start, repetition to mid, NOT OK  $\Rightarrow$  start on long



d) SSL-start, repetition to mid, NOT OK  $\Rightarrow$  start on long



d) SSL-start, repetition to mid, NOT OK  $\Rightarrow$  start on long



d) SSL-start, repetition to mid, NOT OK  $\Rightarrow$  start on one before long



d) SSL-start, marked start OK  $\Rightarrow$  start on sequence



d) SSL-start, unmarked start NOT OK  $\Rightarrow$  no grouping



Figure 5-20. Example of grouping indications for duration and contour sequences. Two short notes before long, 2:1-relationships

e) Sequence with three *short before long* 

Slight marker-dependent preference for *start on sequence*, depending on short-long relationship (3:1 gives equal). Short-long start demands repetition to long.

f) Sequence with four *short before long* 

Preference for *start on sequence*. *Start on long* demands pitch integrity or repetition from long. Start on first, second and third element marker-dependent.

g) Sequence with five *short before long* 

Preference for *start on sequence*, marker-dependent in relation to second and third pos.. Avoidance of *start on long*, or on *short before long*.

h) Sequence with six *short before long* 

Preference for *start on sequence*, marker-dependent in relation to second and third pos..  
 Avoidance of *start on long* etc.

i) Sequence with n *short before long* 

Strong preference for *start on sequence*, marker-dependent. Strong avoidance of *start at end of sequence*.

The above durational description provides examples of the consequences of the rules for simple duration- and pitch-contour sequences with 2:1 relationships between long and short.

When the relationship between short and long tones is above 3:1 the distance–rule makes the sequences end-oriented or too ambiguous to be regarded as overruling boundaries given by local structure and will be disregarded.

non-complex sequences (4:1 relationships)

c) SL-start. hidden start by repeated tone ambiguous grouping



Figure 5-21. Pitch and duration sequences: Ambiguous sequencing with 4:1 relationship

Complex sequences are in general more salient than simple sequences. However, when the number of categorical duration-changes is above three, the sequences become to complex and obscure, according to the ‘*prägnanz-rule*’, to overrule local boundaries. Hence they will be disregarded.

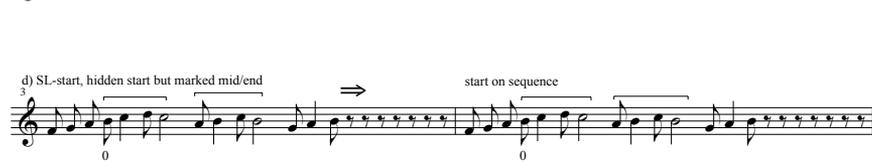
The same rules of *discrete sequence-start* apply for complex sequences as simple sequences. But durational integrity can also compensate for pitch integrity of the sequence, which generally supports *start on sequence*.

complex sequences (2:1-4:1 relationships)

c) SL-start. hidden start by repeated tone start on long



d) SL-start, hidden start but marked mid/end start on sequence



e) SL-start. unmarked start by chained return start on long

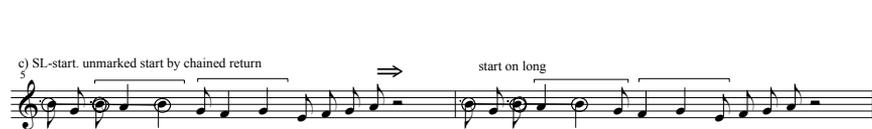


Figure 5-22. Pitch- and duration sequences. Short complex sequences



Figure 5-23. Long complex sequence, start on longest tone

## Duration-sequences

Indications of strong duration-sequences are evaluated through an analogous method as the duration plus contour sequences. The conditions and determinants are identical to those as described above.

There are, however, more limited conditions for pure duration-sequences to be accepted as *strong sequences*, since the melodic contour does not help in determining the integrity of the sequence. This is especially true for longer duration sequences, which need pitch integrity to be recognized, i.e. have to be discrete regarding pitch register.

## Contour-sequences

Since pitch change is a weaker structural parameter than duration on a primary level (see section 5.2.4.3) contour-sequences demand a high level of pitch integrity to give evident indication of sequence-start.

The basic conditions are the same as the aforementioned:

- No change of duration is allowed within the groups except for the last element
- The direction of the pitch changes must be identical in the two groups, except for the last pitch change.
- The direction between every corresponding element of the groups must be identical, except for the last element

This condition implies that the actual size of pitch change can vary between the groups, as long as the change of direction within and between groups is identical.

When the basic conditions are fulfilled the sequence is evaluated according to a set of conditions, which are described and exemplified below. Note however, that the metrical implications given by this part of the analysis can be overruled by subsequent stages of the analysis.

The below rules share many properties with e.g. the rules proposed by Narmour (1990). A fundamental difference is, however, that this primarily regards grouping by sequencing. These rules have been evaluated by and derived from results of listener tests. (see section 5.3.4.2, *Experiment 1*).

## Obscuring melodic return

*Melodic return* groups the notes involved together. When melodic return (e.g. notes a-b-a) occurs from two notes before the sequence to the first note in a 'chained' sequence, the sequence start is obscured and the sequence starting on the second note will be recognized.



Figure 5-24. Examples of the implication of obscuring melodic return

**Hidden sequence: Two-element sequence overlaps three-element sequence.**

This condition occurs when a sequence of two tones overlap a three-tone sequence such that the first two tones in the 3-group actually form a 2-group. If the 3-tone sequence is chained a 3-tone sequence can count from the third tone.

Another variation of the same condition applies when a 2-tone sequence overlaps the 3-tone sequence such that the last two tones in the 3-group form 2-group. Then the sequence is not recognized.

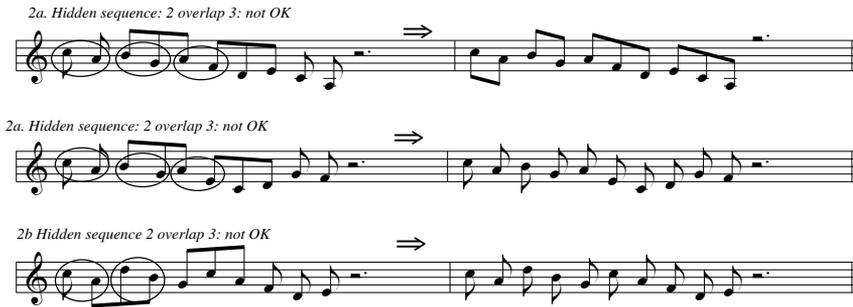


Figure 5-25. Examples of implication of the “Hidden sequence”-condition

**Chained sequence – non-discrete start**

The first variation of the third condition applies to *chained sequences* without discrete start with change of direction within the sequence. For the sequence to count it requires discrete mid- and end-borders are required.

Chapter 5

3a1. chained seq.: not OK

3a1. chained seq.: not OK

3a1. chained seq.: OK

3a2. chained seq. not OK

Figure 5-26. Examples of implication of the first “Chained sequence”-condition

The second variation of the third condition applies to unidirectional *chained sequences* without discrete start.

3b1. chained seq. not OK

3b2. chained seq. not OK

3b3. chained seq. not OK

Figure 5-27. Examples of implications of the second “chained sequence”-condition

**Repetition Condition**

This condition applies only to sequences with three notes, which contain one pitch repetition. This causes a built-in ambiguity, since one repetition of a pitch indicates grouping in two. For such a sequence to count, a discrete ending after the second group is required.

4a. Repetition condition not OK

4a. repetition condition OK

Figure 5-28. Examples of implications of the “repetition” condition

**End Repetition**

- Repetition of pitch over end-border invalidates sequences which do not contain repetition.

5. end repetition not OK

Figure 5-29. Example of implication of the “end repetition” condition

**Unidirectional sequences without discrete start**

Unidirectional sequences without *discrete start* must not be preceded by *change of direction* on the preceding note, or *gap-fill movement* to the preceding note, if they are not discrete regarding step-size at mid- and end-position.

6. not discreet start. not OK

6. not discreet start. OK

Figure 5-30. Examples of implication of the “Unidirectional, non-discrete start” condition.

**Unidirectional sequences without discrete ending.**

Unidirectional sequences without discrete ending must have at least change of step-size or of direction at end-position to count.

Chapter 5

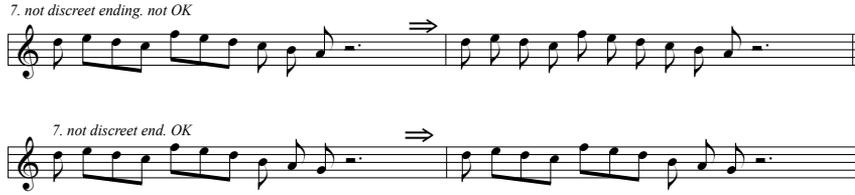


Figure 5-31. Examples of implication of the “Unidirectional, non-discrete ending” condition



Figure 5-32. Example of implications of the fundamental conditions for a “strong” contour sequence

**Fundamental conditions**

The fundamental conditions (see fig 32), which apply to all sequences not covered by the special conditions listed above, states that a sequence counts if all of the following conditions apply:

- The step-size between the groups (mid-boundary) is not both smaller than the maximum step within the sequence and larger than the minimum step within the sequence.
- Either it must involve change of duration at any group-boundary or not the same step-size between groups as within groups.

- Change of duration or direction at either start- or mid-boundary. and either
- Change of step-size, duration or direction at mid or end-boundary. and either
- Change of direction at two subsequent boundaries
- Repetition and another direction within sequence
- Repetition at start and change of direction at all boundaries
- Discrete step-sizes at two subsequent boundaries, such as step-size > maximum step within sequence or < minimum step within sequence.

If these conditions apply, the sequence is said to count.

#### 5.2.4.6 Segmentation indication by discontinuity/change – method design

This part of the analysis of composite metrical grouping is performed through a *local* discontinuity analysis typically involving four beats and a *global* discontinuity analysis. The latter is concerned with major structural changes typically encompassing more than four beats.

Both local and global discontinuity analysis is based on the classification of relative pitch and duration as significant structural determinants of melodic structure. Pitch and duration are classified relative to the atomic pulse, as beat events, which reflect the fundamental assumption of the pulse as the primary temporal mapping of the musical flow, i.e. every event is perceived in relation to the pulse. (See *Assumptions* no 1 and 10)

A central issue regarding the analysis of discontinuities is what structural changes will be recognized as discontinuities? The different levels of classification of pitch and duration reflect different levels of structural generalization which are used to detect structural changes.

The following apply to duration:

- **Dur-class** classifies the event-durations in relation to beats: Dur-class value 0 (labelled **ti**) represents no onset or offset at beat-start; Dur-class value 1 (labelled **short**) represents shorter duration than beat-length on beat-start; Dur-class value 2 (labelled **long**) represents duration onset or offset at beat-start of beat length; Dur-class value 3 (labelled **xlong**) represents duration onset or offset at beat-start exceeding the length of the beat.

This classification reflects the prominence of the metrical grid (*Assumption no 11*), thus labelling the durations only in relation to its relation to the atomic pulse.

- **Division** classifies the event-durations in relation to beats somewhat more specifically through a numerical representation of the number of events (onsets and offsets) at a beat.

This equals e.g.  and  as a division of a beat.

This classification reflects a principle of variation of duration dependent on the prominence of the metrical grid, which is frequent in e.g. Western Classical Music (See e.g. Variation examples by Quantz, quoted in Narmour 1999:447).

- **Dur-change**, in addition to the *division* classification specifies the relative successive relationships between the durations within a beat, into the categories longer, shorter, equal to next tone.

This equals e.g. the following figures,  which reflects the relative nature of the perception of duration.

- **Duration** of events (*interonsets* and *interoffsets*) per beat expressed as midi clicks.

Pitch-classification is also performed beat-wise:

- **Step-size-dir** encodes the relative pitch-changes connected to a beat as a list consisting of
  - a) The melodic pitch category (MPC) distance from the first MPC of the beat to the first MPC of the next beat, expressed as the difference between the initial MPC of the next beat and the initial MPC of the current beat.
  - b) The MPC distances within the beat expressed as the successive differences between the next MPC and the current.

This implies that change to higher relative pitch will result in positive values and conversely negative numbers for change to lower relative pitch. A rest will result in a NIL value. This classification reflects the punctuate nature of beat conception, the start of a beat as a marked or accentuated event. (*Assumption no 20*). Thus the initial event of a beat will be structurally more significant relative pitch changes within the beat.

- **All-succ** (*all-successive-mpc:s*) encodes the relative pitch changes as MPC:s related to the first MPC of the melody as differences between current MPC and the referential first MPC. It presupposes that the pitch relationship persists over rests and grand pauses, i.e. the existence of a local absolute pitch reference within the realms of a melody.

This reflects (as does step-size-dir classification) the prominence of MPC changes as the basic structural pitch changes of the melody (see Chapter 4, *Analysis of Melodic Pitch Categories*). The categories *leaps* and *steps* hence are interpreted as dependent upon MPC structure; steps designating changes between adjacent MPC:s and leaps changes between distant MPC:s.

- **All-notes** encodes the changes between the initial relative pitches as MPC distances as described above regarding step-size-dir above, but the pitches inside the beat are represented as MIDI-note numbers, hence reflecting local absolute pitch memory within melodies (see Chapter 3, *General Model*)

These structural classifications in relation to beats are used throughout the entire analysis of melodic structure, with the addition of some further structural classifications relating to beats.

Based on the above classification of duration and pitch structure, every event of the melody is evaluated with respect to local and global discontinuity. This evaluation results in higher values at possible segment starts after possible segment borders; the higher the value, the greater the possibility of a segment start.

As mentioned above this includes local as well as global discontinuities, which will be described more thoroughly below.

The values derived from the previously described sequence analysis and the discontinuity analysis are the weighed to form a total value of the segment start indications of each beat.

### **Analysis of local discontinuities**

This analysis is formed in the computer implementation performed by the function marker-at-p-5, which is one of three similar discontinuity-rating functions within the computer

model. This particular function is also used for determining phrase borders through local discontinuities.

This function is basically a set of conditions ordered so as to reflect relative structural importance from more prominent discontinuities based on onset/offset and duration changes to less prominent discontinuities such as change of melodic direction or relative pitch change.

In the previous section analysis of *segment start indication* through strong contour sequences were described. An analysis of shorter sequences of lesser strength is necessary here to evaluate the local discontinuities. A sequence can override certain structural changes and be disguised by others, but it can also enhance or disguise other discontinuities.

Hence, the melody is once again searched for short duration- and pitch sequences. These sequence analysis differ from the earlier performed sequence analysis in a number of respects: First, it is based on the different classifications of beats described above, while sequences are identified on a number of different levels of specificity ranging from dur-class sequences to actual duration sequences, near contour identical sequences to pitch sequences. Second, every sequence is counted and no extinction of overlapping sequences is made. Instead they are merged into a total sequence-list where identical sequences of lower level of specificity are deleted.

Thus the list of sequences can contain sequences at different levels of specificity.

This sequence identification is performed in generally the same way as the sequence identification in the phrase analysis and will be described in detail in that context (see Chapter 6, *Metrical Phrase and Section Analysis*, section 6.2). It should be noted here that contour-sequences of two or three atomic beats are counted, while specific duration sequences may contain up to 24 atomic beats, which is regarded as a general upper limit of number of beats per measure<sup>50</sup>.

The resulting sequence list serves together with the classification of pitch and duration changes as input to the analysis of local discontinuities in marker-at-p-5.

### ***The principles behind the valuation***

1. Since this analysis concerns metrical grouping it presupposes that segments start only on beats (atomic beat level).
2. A local structural change involves typically three transitions, – before change, change and after change – and thus typically involves events of four beats. There are exceptions, like the end of a melody, which does not require the involvement of four beats to be recognized. Likewise, there are situations, which involve more than four beats, e.g. regarding the influence of sequences.
3. From the condition of *metrical grouping* concerning composite beats or time, it follows that we are looking for impulse points, structurally accentuated events. (*Assumption no 11, 20*) Beats as well as measures are counted from the most marked event (*Assumption no 20*). This makes duration not only a grouping factor by distance, but also by salience: A longer duration can be perceived as a structurally more salient event, indicating start, as well as indicating end, through the *interonset* distance. This is quantified in the method through stronger or equal prominence for longer duration as a start indication in relation to end

---

<sup>50</sup> If certain conditions apply, such as extremely short beats, measures up to the limit of 40 beats can be tracked by the model.

indication up to the length of three times the atomic beat value. Greater values indicate that it will encompass more than one metrical subgroup (due to the 2- and 3-group subgroup preference rule, (principle of low-level grouping, ass no 3)), which implies the suspension of composite beat onset, hence suspension of start accentuation.

4. Regarding pitch changes distance is still the primary factor determining segment borders, indicating start after. However, since this analysis mainly concerns the primary grouping of events immediate pitch repetition and return is of greater importance than what is normally the case in phrase border analysis. On this primary grouping level, a repetition of a pitch can be perceived as a prolongation of the pitch, hence comparable to a longer duration of a tone and conceivable as a divided longer tone. Analogically, a return can be perceived as an interrupted prolongation of a tone.
5. The nature of primary low-level grouping also leads to the relative prominence of onset and offset information as segment border determinants in relation to pitch information.
6. Non-changing structure, i.e. structural similarity between the current beat event and the previous, gives, of course, low group-start values, while persistent structural change, i.e. structural similarity between the current beat event and the next, gives higher group-start values.
7. Combinations of different types of discontinuities, such as indication by a combination of change of melodic direction and change of duration, gives stronger start-indications than single discontinuities.
8. The strength and conditions for determining structural discontinuities depend globally on the rhythm-metrical structure, i.e. the level of division of beats and maximum duration of sounds in relation to beats. The underlying assertion is the assumption of the category dependency of pulse conception (*Assumption no 13*). Therefore, an analysis of the rhythm-metrical structure is performed by the function `sparse-rhythm-p`. This function sets a variable `super-imposed-meter-alert` which is set to true if the rhythmic structure is sparse in relation to beats, i.e. if beats are divided into a maximum of two or three equal time intervals and the duration structure is dominated by events of the same duration as beats or longer. Then there is supposed strong possibility for super-imposed pulse conception. Typical examples is the eight note pulse in pieces notated in 6/8, which often implies a beat level of three eight notes or the quarter note beat in most waltzes notated in 3/4, where it is often possible to feel the entire measure as a beat. If `super-imposed-meter-alert` is true, it follows that more beats have to be taken into account in e.g. the metrical analysis, since structural discontinuities can regard the changes between subgroups and not only single beats and the number of events regarded is less than in a complex beat structure.

The above principles applied to melodic context are summarized below:

Primary grouping indications

- a) Grouping by change of **duration / interonset** interval
- b) Grouping by change of onset - not onset - offset
- c) Grouping by pitch pattern similarity: Sequence
- d) Grouping by pitch continuity 1: Repetition (note prolongation).
- e) Grouping by pitch continuity 2: Return
- f) Grouping by pitch discontinuity: Melodic direction
- g) Grouping by pitch distance: Step-size / register
- h) Composite pitch change 1: Melodic direction and step-size

## i) Composite pitch change 2: 'Roundabout' / intervallic return

These structural indications can be regarded as relatively 'blunt' instruments in determining primary groups. Nearly every parameter has at least two different interpretations, either due to different applicable principles – *does a long note represent an articulated event or does it represent greater interonset distance and hence a segment border* - or due to the ambiguity of the identification of change – *does the change take place from the last event of the previous structure or the first event of the new, before or after*. Hence, the interpretations of these indications are contextually dependent, relying primarily on secondary and tertiary grouping principles (see section 5.2.4.3 and below).

This leads to the need for flexible implications of grouping factors as well as complex evaluation of these implications. Below the implication for each factor will be exemplified and discussed. Numerical values (below note examples) represent a valuation of the segment border indication; higher values represent stronger indication of segment start at position.<sup>51</sup>

a) **Grouping by change of duration/interonset interval:**

The segment start implication of change of duration exceeding the atomic beat level is twofold and converse: *Interonsets* of longer duration are structurally prominent, hence indicating start on event. Conversely, *interonsets* of longer duration imply longer distance between onsets, which according to the grouping by proximity principle implies start after the event of longer duration. Which of these two interpretations is prominent depend briefly on the length of the duration in relation to the atomic beat level. Change of duration within atomic beat level implies change of *interonset* intensity, giving prominence to divided atomic beats (ornamented beats) relative to undivided atomic beats. The above principles are exemplified in the following example. Under ordinary conditions, a duration which involves more than two atomic beats is interpreted according to the distance principle, giving the strongest segment start implication after the long event<sup>52</sup>. As opposed to that, a duration of about twice the length of an atomic beat is interpreted as having a stronger 'grave' accent quality, giving stronger segment start implication value to the longer note. A divided beat is interpreted as an accentuated beat – ornamentation through *interonset* frequency - in principle giving segment start indication to the ornamented beat.

However, as previously mentioned above, the interpretation of these general indications is strongly contextual. For example; Sequencing can revert the prominence implications shown in example a) , b) and c), so can interplay with pitch indications, not to mention implication by secondary and tertiary grouping principles. The interplay with other primary and secondary grouping principles in the initial valuation can be read from the table of segment border valuation (Table 1, below)

---

<sup>51</sup> The following valuation of prominent group boundaries, share many properties with the group boundary preference rules suggested by Lerdahl & Jackendoff, the Implication-Realization interpretation of melodic expectation suggested by Narmour. The fundamental approach is, however, different, since the current regards metrical accentuations.

<sup>52</sup> However, the duration of three times the atomic can under certain conditions, e.g. for high tempi or sparse onset structure, be interpreted as dominant segment start indication

DURATION

a) xlong > 2 X beat

b) xlong = 2 X beat

c) divided beat/ornament



b) **Grouping by change of onset/not onset/offset (related to a)):**

Note onset is prominent segment start indication, 'music on' being more prominent than 'music off'. Not at onset or offset is strongly avoided, giving negative segment start indications, as can be read from the example.

a) onset-not onset

b) before and after rest

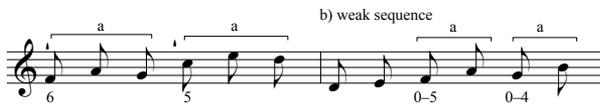
c) before and after xlong rest



There is, however, a difference between the indication value of the last atomic beat-position before the next onset and the non-onsets on the next beat-position, the former having a slightly higher value (-1). In fact, as will be demonstrated in section 5.3.2.7, it can, by contextual support, get as high reach value as 5. The reason for this ambiguity is that it is the last beat event before a change, which can be interpreted as the event from which the change takes place. This contextual duality of the last beat event before onset is of course extremely common in much Western music, not to the least in Swedish fiddle music and other European dance music traditions. A segment start without onset on the next event is much more rare, albeit contextually dependent.

c) **Grouping by pitch pattern similarity: Sequence**

Sequence means that the similar contiguous sets of pitch and/or duration changes are perceived as groups, which implies that the perceived beginning of these groups are conceived as more prominent starting points than events within the groups. Here are also the 'weaker' sequences not taken into account by the strong sequence analysis considered. (min and max values depending on context are displayed in the figure)



d) **Grouping by pitch continuity 1: Repetition (note prolongation).**

Repeated notes are grouped by pitch continuity, which implies that *start on first note* of repetition is more prominent than *start within repetition*. According to the discontinuity rule, the last note of a repetition – the turn of melodic direction / point of change of step-size – is regarded as a more prominent indication of segment start than notes within the repetition but less prominent than the first note of the repetition.



e) **Grouping by pitch continuity 2: Return**

Return (pitch interval and its reversal) can be regarded as an extended repetition with interruption, thus grouping by pitch continuity. Start of return is most prominent, where after comes last note of return. Repeated (or ‘chained’) return will according to the pitch continuity principle be perceived as an extension of pitch continuity of the notes involved, which generally gives more ambiguous indication and less segment start indication to the notes within the chained return.

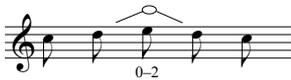
RETURN



Note that the return indication normally is in conflict with other principles like change of melodic direction or gives ambiguous indication through chained or overlapping returns.

f) **Grouping by pitch discontinuity: Melodic direction**

Change of continuous pitch change implies the grouping of elements involved in the continuous pitch change, which in turn implies prominent group start at the turn of melodic direction. The strength of this indication is evidently dependent on the perceived generality of this change, i.e. how many events are involved in the two opposed melodic directions.



g) **Grouping by pitch distance: Step-size / register**

Successive tones close in pitch are grouped, while more distant tones indicate segment border, hence indicating most prominent segment start after the largest distance. These changes are perceived categorically in relation to the system of melodic pitch categories: Changes between adjacent categories are regarded as steps, close, while changes between distant categories are regarded as leaps, distant.

Step-size-change gives however in itself a certain segment border indication.

STEP-SIZE CHANGE

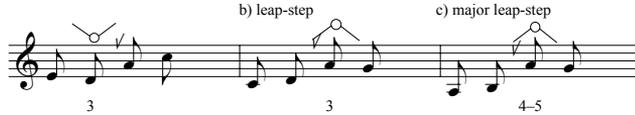
a) small distance

b) large distance



h) **Composite pitch change 1: Melodic direction and step-size**

Certain combinations of indications have special significance. Typical is the step/leap – leap/step and direction change, which gives a relatively strong start indication regarding segment start indications by pitch changes.



i) **Composite pitch change 2: 'Roundabout' / intervallic return**

Another special combination of change of melodic direction, step-size and return with relatively weak indication is the 'roundabout' or intervallic return, implying a larger interval reversed by a minor interval. This functions as articulation of the final tone of the roundabout, defining a center of a pitch space, hence indicating segment start at the final tone.

ROUNDABOUT / INTERVALLIC RETURN



The interpretation of the above segment start indications depend as have been mentioned above on secondary and tertiary grouping principles as well as primary group constraints (see section 5.2.4.3).

The conditions of the local segment border evaluation *including* the impact of the interplay between different factors, is covered in the table below, with examples in common notation. Numerical values (below note examples) represent a valuation of the segment border indication, where higher values represent stronger indication of segment start at position. The values range from -4 to 7, where -4 is a position after the melody and 7 is given at an extremely supported start position (see below). Negative values are given only when there is no onset at a beat.

The order in which the conditions are presented is (as mentioned above) identical to how they are implemented in the method, and why this order represents also the order of relative importance of the different discontinuity factors. The values given also represent this order of segment start indication.

For the different conditions the main implementation is described briefly in the labelling and exemplified in common notation and the most important exceptions are described in brief terms.

*Descriptive signs and terms used in table 1.*

mp	(music-ptr), the event in question repetition or return to an identical tone shift of melodic direction gap/leap between two adjacent notes sequence or subsegment/beatgroup sequence/event change regarded similar	xlong (xl) long (l) short (s) superimposed (superimp.) on-seq-border	dur-class: duration at beat exceeding the beat dur-class: undivided beat dur-class: divided beat conditions which apply only to structures without divided beats. (beat = eighth note) sequence border at mp
----	--	--	--

Table 1. Conditions for Local group boundary valuation

**A1** STARTPOSITION, i.e. start of melody or after xlong rest  $\Rightarrow$  either -2, 4 or 5

a:1a rest at start of melody  $\Rightarrow$  -2      a:1b note at start of melody  $\Rightarrow$  5      a:1c first after xlong rest  $\Rightarrow$  4

**A2** LAST BEFORE END OF MELODY  $\Rightarrow$  0

**A3** XLONG/LONG FOLLOWED BY REST/REST ON NEXT  $\Rightarrow$  gen. -2

a:3d long before rest  $\Rightarrow$  -2      a:3a xl before tail  $\Rightarrow$  3      a:3b xl before rest  $\Rightarrow$  -2      a:3c l before div. rest  $\Rightarrow$  -1

**A4** AFTER PICKUP, i.e. xlong and rest before mp  $\Rightarrow$  7

a:4a first after div. prev with reststart      a:4b xl before prev      a:4d xl rest before mp

superimposed      a:4e omitted startnote after xl

**B1** SHORT REST/TIE  $\Rightarrow$  -1

b:1e short rest/tie  $\Rightarrow$  gen.-1

exceptions:

- short rest on prev after xl  $\Rightarrow$  gen. 2, max 4
- short tie to mp – anticipatio  $\Rightarrow$  gen. 2, max 4
- sequence start variation  $\Rightarrow$  max 2

**B2** LONG REST/TIE  $\Rightarrow$  gen. -2

b:2d

*exceptions:*

- sequence-start-variation  $\Rightarrow$  max 2
- short tie from prev  $\Rightarrow$  gen. 1, max 3
- note-onset on div. next  $\Rightarrow$  -1

**D1** DIVIDED AFTER TIE  $\Rightarrow$  gen. 6 (min 0)

*exceptions:*

- beat-dur-change on next  $\Rightarrow$  gen. 0
- prev divided, short tie to prev, i.e. anticipation to prev.  $\Rightarrow$  gen. 2, min 0
- no marker on mp  $\Rightarrow$  gen. 2
- prev – mp –equality (beat-sequence)  $\Rightarrow$  gen. 0

**D2** XLONG AFTER LONG/TIE  $\Rightarrow$  gen. -1

*exceptions:*

- super-imposed mp after xl (2\*beat)  $\Rightarrow$  max 4

**D3** XLONG AFTER DIVIDED /superimposed: LONG  $\Rightarrow$  gen. 4 (with on-seq-border)

*exceptions:*

- beat-dur-change on next  $\Rightarrow$  gen. 0
- repetition to xlong  $\Rightarrow$  gen. 1
- xlong on next  $\Rightarrow$  gen. 2

**D4** REPETITION TO LONG AFTER DIV./TIE  $\Rightarrow$  0 (max. 5)

*exceptions:*

- chained rep. to mp – dir-change from  $\Rightarrow$  gen. 2
- end through ssl on prev  $\Rightarrow$  gen. 4

- xxl before mp  $\Rightarrow$  gen. 4
- sequence  $\Rightarrow$  gen. 3
- divided next  $\Rightarrow$  gen. 2

**D5** LONG, BEAT-DUR-CHANGE ON NEXT  $\Rightarrow$  0 (max 1)



**D12** LONG AFTER SHORT/TIE  $\Rightarrow$  gen. 5



*exceptions:*

- beat-dur-change at mp  $\Rightarrow$  gen. 7
- next xlong  $\Rightarrow$  gen. 0 / 2
- superimposed: long after tie  $\Rightarrow$  gen. 4
- xlong after divided  $\Rightarrow$  2 (Note! overlap with D3)
- superimposed: beat-dur-change on next  $\Rightarrow$  gen. 0
- long after s-s, s on next  $\Rightarrow$  gen. 1

**E** DIVIDED (S) AFTER LONG/XLONG  $\Rightarrow$  GEN. 5 (L) OR 6 (XL)



*exceptions:*

- beat-dur-change on next  $\Rightarrow$  gen. 1/0
- chained repetition  $\Rightarrow$  gen. 2
- beat-dur-change to mp  $\Rightarrow$  gen. 6
- rep. to mp  $\Rightarrow$  gen. 0/3
- xlong on next  $\Rightarrow$  gen. 1
- rep. to prev  $\Rightarrow$  gen. 4
- rest on prev  $\Rightarrow$  gen. 3

**F** SHORTER INIT. DUR. THAN PREV  $\Rightarrow$  gen. 2



*exceptions:*

- with dir-change 3
- beat-dur-change to mp 6

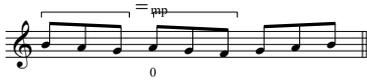
**H** LONGER INIT.DUR. ON MP THAN PREV  $\Rightarrow$  gen. 2



*exceptions:*

- beat repetition to mp  $\Rightarrow$  gen. 1,4, 0
- prev-mp sequence  $\Rightarrow$  gen. 1
- start indication on prev  $\Rightarrow$  gen. 1
- beat-dur-change on next  $\Rightarrow$  gen. 0
- hidden ending on prev/beat-dur-change on mp  $\Rightarrow$  gen. 5

**I** PREV AND MP SEQUENCE IDENTICAL  $\Rightarrow$  gen. 0



*exceptions:*

- mp-tmp equality higher than prev-mp-eq  $\Rightarrow$  gen. 2
- beat- and init-dir-change at mp and not at prev  $\Rightarrow$  gen. 2/1
- reg.change to mp and not at prev  $\Rightarrow$  gen. 2

**J** MP AND NEXT SEQUENCE IDENTICAL  $\Rightarrow$  gen. 3



*exceptions:*

- weak structural change at mp  $\Rightarrow$  gen. 1
- return from next  $\Rightarrow$  gen. 2
- high similarity plus reg-change  $\Rightarrow$  gen. 4/5
- repetition to mp  $\Rightarrow$  gen. 0

**L** FIRST AFTER HIDDEN REP.  $\Rightarrow$  gen. 3



*exceptions:*

- with beat-dur-change on mp  $\Rightarrow$  5
- prev-mp duration equality, long before prev  $\Rightarrow$  gen. 1
- superimposed: sequence from mp  $\Rightarrow$  gen. 4

**M** RETURN/REPETITION TO MP  $\Rightarrow$  gen. 0 (max 4)



*exceptions:*

- repetition from long (ssl or dur-change at mp)  $\Rightarrow$  gen. 2 or 3
- return to mp, leap from mp  $\Rightarrow$  gen. 3
- superimposed: dir-turn over more than 4 beats  $\Rightarrow$  gen. 2
- with on-seq-border  $\Rightarrow$  gen. 3
- beat-dur-change to mp  $\Rightarrow$  gen. 2
- repetition to prev  $\Rightarrow$  gen. 2

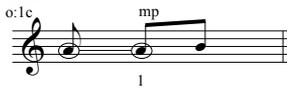
**N** RETURN/REP. FROM MP  $\Rightarrow$  gen. 2 (max 4)



*exceptions:*

- rhythmic structural change at next  $\Rightarrow$  gen. 1
- with rest, leap etc. to mp  $\Rightarrow$  gen. 4

**O** INITIAL REP. TO MP  $\Rightarrow$  gen. 1 (max 4)



*exceptions:*

- hidden long before mp  $\Rightarrow$  gen. 4
- with dir-change/step-size-change at mp  $\Rightarrow$  gen. 3

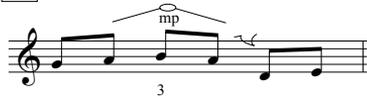
**P** DIR. CHANGE AND/OR STEP-LEAP CHANGE  $\Rightarrow$  gen. 3 (0-7)



*exceptions:*

- only init-step-dir-change  $\Rightarrow$  gen. 1
- beat-dir-change to next > init-step-change to mp  $\Rightarrow$  gen. 0
- xlong on next  $\Rightarrow$  gen. 1
- with major leap to mp  $\Rightarrow$  gen. 4 or 7

**Q** DIR-CHANGE AND STEP-LEAP-CHANGE  $\Rightarrow$  THIRD  $\Rightarrow$  gen. 3



**R** DIR-CHANGE AND LEAP-STEP-CHANGE  $>$  THIRD  $\Rightarrow$  3 or 5



**S** SIMPLE DIR-CHANGE ON BEAT- OR INITLEVEL  $\Rightarrow$  gen. 1 or 0



*exceptions:*

- large leap to mp  $\Rightarrow$  gen. 3
- beat-dur-change at mp  $\Rightarrow$  gen. 3
- hidden end/rep. to prev  $\Rightarrow$  gen. 3
- with beat- or init-dir-change (not covered above)  $\Rightarrow$  gen. 2
- superimposed: with large dir-turn  $\Rightarrow$  gen. 2 (max 4)
- with on-seq-border  $\Rightarrow$  gen. 3 (min 1, max 4)
- weak dir-change on mp  $\Rightarrow$  gen. 0
- superimposed: with on seq-border  $\Rightarrow$  gen. 2 (max 4)
- superimposed: with dir-turn  $\Rightarrow$  gen. 3
- superimposed: with return to mp  $\Rightarrow$  gen. 2

**T** CHANGE OF STEP-SIZE (init-level)  $\Rightarrow$  gen. 1



*exceptions:*

- superimposed: with return from mp  $\Rightarrow$  gen. 4
- with leap to mp  $\Rightarrow$  gen. 3
- with mp-next-equality  $\Rightarrow$  gen. 3
- with on-seq-border  $\Rightarrow$  gen. 4
- superimposed: with dir-turn over min. 5 beats  $\Rightarrow$  gen.3
- with return to mp  $\Rightarrow$  gen. 2 (max 5)

**U** CHANGE OF STEP-SIZE (BEAT-LEVEL)  $\Rightarrow$  6, 2 or 3

u:1 leap to > fifth, step from etc. u:2 gen. mp u:3d superimposed

*exceptions:*

- with on-seq-border  $\Rightarrow$  gen. 4
- repeated step-size-dir from mp  $\Rightarrow$  gen. 4

**V** RETURN FROM MP (SUPERIMPOSED)  $\Rightarrow$  2

v:4a mp v:4b mp v:4c mp

*exceptions:*

- with chained return  $\Rightarrow$  gen. 3
- only stepwise motion  $\Rightarrow$  gen. 1
- with on-seq-border  $\Rightarrow$  gen. 3

**W** ON-SEQ-BORDER  $\Rightarrow$  gen. 3

w:4 mp

*exceptions:*

- indiscrete seq-border  $\Rightarrow$  gen. 0/1
- overlapping rhythm-sequence  $\Rightarrow$  gen. 1
- indiscrete short sequence  $\Rightarrow$  gen. 0

**X** REPETITION FROM MP (init)  $\Rightarrow$  gen. 1

x:1 mp

### Analysis of global discontinuities

This analysis concerns structural changes which typically require the consideration of more than four beats – three transitions. Four types of such structural changes are identified corresponding to the five most important types of structural dimensions regarded in the local analysis:

Table 2. Corresponding structural indications on local and global levels.

local	global
step-leap-change	global register change
local change of melodic direction	global change of melodic direction
duration change	global rhythmic structural breaks
onset – offset change	global rhythmic structural breaks
pitch set change	global pitch set change

#### Global register change

Global register change (global-reg-change) represents the relative pitch register shifts between groups of notes, generally involving a minimum of eight notes or four beats. The fundamental conception is that we conceive pitch differences not only between individual notes but between groups of notes (see Chapter 3, *General Model*, section 3.3): Notes which are close in pitch will be grouped together and separated from groups of notes distant in pitch. This is designated as shifts in pitch register and is a strong structural factor, phenomenally related to perception of change of timbre.

#### *Descriptive signs and terms used in table 3.*

mp	(music-ptr), the event in question	superimposed (superimp.)	conditions which apply only to structures without divided beats. (beat = eighth note)
∨	gap/leap between two adjacent notes		

Actual notes are arbitrary. Registral conditions are considered and when certain differences between initial notes are required they are indicated by lines with rings between notes

Table 3. Examples of conditions for registral shifts as implemented in *global-reg-change-list*

A1 MAJOR REG. SHIFT (diff. > sixth) ⇒ 4/-4

not super-imposed mp super-imposed mp

A2 HIDDEN MAJOR REG. SHIFT ⇒ gen. 4/-4

(not superimp:) min-max-diff > third between and < fifth within

mp

B1 GEN. REG. SHIFT ⇒ gen. 2/-2 (with rest on mp 4)

a) min-max-diff betw. > step and max-min-diff betw. > fifth    b) min-max-diff betw. > third    c) superimp. init-diff betw. > third for four closest beats    d) with long rest on mp 4

mp mp mp mp

B2 DIVIDED REG. SHIFT ⇒ gen. 2

divided leap > third and diff. betw. tmp-min and mp-inits > step

mp

C MINOR REG. SHIFT ⇒ gen. 1

a) tmp-min = mp-max and within-range > third    a2) tmp-min = tmp-max mp-tmp-init-diff > third    a3) tmp-min = tmp-max init-overlap

mp mp mp

b) mp-min = mp-max min-max-diff betw. > step    c) init-diff > step

mp mp

### Global direction change

*Global change of melodic direction* considers melodic direction over a minimum of four beats (three between-beat-transitions). When *super-imposed-meter-alert* is true (see Analysis of Local Discontinuities, *The principles behind the evaluation*) a minimum of 6 beats is required. The concept of global pitch change is based on partly on the grouping of pitches of about the same height (see *global-register-change* and Section 3.3) and partly on the streaming phenomenon. This refers to the experience of a continuous change of pitch in one direction, or a turn of

melodic direction, even when individual pitches or pitches within beat groups do not conform to this general change.

The conceptual basis for this phenomenon can thus be expressed as a unidirectional pitch-change locally during the perceptual present, which here is interpreted as encompassing  $7 \pm 2$  events (the categorical present). This implies that a continuous global pitch-change will be recognized when the sum of pitch changes in the same direction is greater than the sum of contradictory pitch changes, the average pitch changes are in the same direction and are greater than the pitch commonality between the beats and events involved.

For global change of melodic direction to be perceived the above condition of average pitch change and dominance for pitch transitions between beats or events in one direction must be inhibited.

A change of melodic direction gives three possible interpretations of segment border and start-indication, either before, at, or after the beat from which the change occurs. The reason for this is that this beat can be regarded either as the turn of the new direction, the last beat of the previous direction or the first beat of the new direction; although the latter interpretation is given prominence in the method of analysis. The different possibilities of experiencing a turn of melodic direction makes it necessary to give different segment start values to different beats based on the same turn (see table 4 below).

In the table 4 below some examples of the interpretation of global change of melodic direction are given. Examples no. 2 and 4 show the inclusion of events, which do not fulfill the conditions of the unidirectional pitch-change into the identified pitch-change. Inclusion of events applies only when the included segment is shorter than  $1/3$  of the unidirectional pitch-change and contains less than six beats. Inclusion involves the interplay with other discontinuity factors and reflects the principle that global unidirectional pitch change is subordinate to strong global register change, offset-onset and global duration change. However, when a turn of global direction of pitch-change occurs within the next three beats intermediate beats can be included (see table 4 below), since two beats cannot form a structural unit.

Like change in local melodic direction, is change of global melodic direction not the strongest factor for determining segment borders in melodies. This is reflected by the relatively low values given to positions at turn of global melodic direction which are not supported by other discontinuity factors.

Note also that here the implementation is different when *super-imposed-meter-alert* applies, which requires the change of pitch being determined by change between groups of pitches, arbitrary super-beats. (see table 5 below).

Table 4. Examples of rules of global unidirectional pitch change and change of global melodic direction

Markers of global change of melodic direction are indicated by numbers below the systems. Values range between 1 and 3, higher values denoting stronger start indication through change of global melodic direction. A turn of melodic direction is typically given an interpretation involving more than one marked beat, at most 6 marked beats indicating alternative turning points. Equal direction is demonstrated by lines above notes. Markers belonging to the same turn of direction are enclosed.

Events, which included in the direction change in the evaluating analysis because of stronger structural breaks, ambiguity regarding start of direction etc. are marked by dotted brackets.

Ex. 1. Equal direction in both directions. Basic conditions

Musical notation for Ex. 1. A single melodic line in treble clef. The first measure has a '3' in a box below it. The line continues with several measures of eighth notes. A '1 2' is in a box below the line in the 10th measure. Lines above the notes indicate the direction of the melodic movement.

Ex. 2. Equal global pitch change in one direction

Exception no 1: temporary divergence with weak reg. change, divergent initial direction between beats within consistent pitch-change over three beats

Musical notation for Ex. 2. A single melodic line in treble clef. The first measure has a '3' in a box below it. The line continues with several measures of eighth notes. A '1 2 1' is in a box below the line in the 10th measure. Lines above the notes indicate the direction of the melodic movement.

Exception no 2: temporary divergent direction reg. both register and initial direction, within consistent pitch-change over four beats

Musical notation for Exception no 2. A single melodic line in treble clef. The first measure has a '3' in a box below it. The line continues with several measures of eighth notes. Lines above the notes indicate the direction of the melodic movement.

a) structural break through preceding rest moves startmarker backwards

Musical notation for Ex. 2a. A single melodic line in treble clef. The first measure has a '3' in a box below it. The line continues with several measures of eighth notes. A '2 1 1' is in a box below the line in the 10th measure. Lines above the notes indicate the direction of the melodic movement.

b) init. dir. change between beats indicates earlier start

c) init. dir change (+ return) between beats indicates earlier start

Musical notation for Ex. 2c. A single melodic line in treble clef. The first measure has a '2' in a box below it. The line continues with several measures of eighth notes. A '2 1' is in a box below the line in the 10th measure. Lines above the notes indicate the direction of the melodic movement.

Ex. 5 Super-imposed: Super-mp-dir-change, which applies only to regions dominated by undivided beats, is based on test of consistent direction of 2-, 3- or 4-groups of beats

a) equal direction of 4 groups of 3 beats

b) equal direction of 5 groups of 2 beats

c) equal direction of 2 groups of three beats

Musical notation for Ex. 5. A single melodic line in treble clef. The first measure has a '3' in a box below it. The line continues with several measures of eighth notes. Three examples of super-imposed direction markers are shown: (NIL 4), (-5 -3), and (3 2).

d) ex. three groups of three beats (downwards), based on initial values in each super-beat, markers mirror 3-grouping

Musical notation for Ex. 5d. A single melodic line in treble clef. The first measure has a '1 2 1' in a box below it. The line continues with several measures of eighth notes. A '(nil -3)' is written below the line in the 10th measure. Lines above the notes indicate the direction of the melodic movement.

### Global rhythmic structural breaks

Global rhythmic structural breaks are determined mainly by changes of onset and offset (*tone–rest changes*) and duration of events (*interonset* and *offset values*). In addition, global register shifts (*global-reg-change* see above) and pitch repetition are considered in the analysis.

As mentioned above a global rhythmic break typically involves more than the four beats typically considered in the analysis of local discontinuities. The general assumption is that longer *interonset* intervals imply structural breaks, indicating borders after the *interonset* intervals reflecting the gestalt rule of grouping by proximity, division by distance. When *interonset* intervals involve more than two beats, the converse principle of grouping by salience of longer *interonset* intervals which is present in the local discontinuity analysis is suspended, since the missing onsets at beat positions cause metricity (impulse quality) to temporarily be set aside as a governing structural principle.

Therefore, this function is based on the analysis of duration of events rather than beats.

However, quasi-sequencing is regarded in this analysis, such as short duration sequences (or quasi-duration sequences) involving major duration values, which ordinarily would signal major structural break.

Every beat is considered and will be interpreted as a global rhythmic structural break when it is either onset-offset-change (note to rest or vice versa) or duration-change with either xlong (duration exceeding the beat) dur-class or long dur-class (duration of one beat) preceded by a shorter dur-class or a rest at current position.

A maximum of six successive events are considered in the analysis of global rhythmic structural breaks, involving two *interonsets* before actual position and two after. The reason for this is that global rhythmic structural breaks are regarded as singular within the perceptual present, i.e. there cannot be two global breaks within 7 events, then they won't be conceived as a global breaks.

Terms and descriptive signs used in Table 5

**mp** – (*music-pointer*) current event

**tmp, tmp1, tmp2, tmp3** etc. – the events following *mp*

**ppmp, pmp** – (*prev-mp prevprev-mp*) – the events preceding *mp*

♩ ◦ half note or longer – represents xlong duration category, duration exceeding the current beat length

♩ quarter note – represents long dur-class, undivided beat

♪ eighth notes – represents short dur-class, divided beat

♩ half note, quarter note or eighth note rests have analogous dur-class implication

**end** – end of melody

// – alternative dur-classes at position

♩ – global-reg-change to position (see separate explanation)

♩ – pitch-repetition to position

Table 5. Indications of Global Rhythmic Structural Breaks

Principles

d) Start of melody on note indicates start

ppmp	pmp	np	tmp	tmp1	tmp2	tmp3			np
								⇒	5

e) Start on rest and divided beat gives pickup-indication

ppmp	pmp	np	tmp	tmp1	tmp2	tmp3			np	tmp
								⇒	2	3

f) End of melody gives -1 at mp and 7 at end

ppmp	pmp	np	tmp	tmp1	tmp2	tmp3			np	tmp
			end					⇒	-1	7

g) the longest distance gives *end-start-indication* from ppmp to tmp1, i.e. within the range of two events from mp

ppmp	pmp	np	tmp	tmp1	tmp2	tmp3			np	tmp	tmp1
								⇒	0		4

ppmp	pmp	np	tmp	tmp1	tmp2	tmp3			np	tmp
								⇒	0	2

h) no change of duration, gives no *start-indication*

	pmp	np	tmp	tmp1	tmp2	tmp3			np
								⇒	nil

i) pickup-indication (div. beat to whole beat after xl, or tied and div. beat to untied) gives divided start-indication after end

ppmp	pmp	np	tmp	tmp1	tmp2	tmp3			np	tmp	tmp1	tmp2
								⇒	0		2	3

j) rest gives start-indication after rest

Chapter 5

ppnp	pnp	np	tmp	tmp1	tmp2	tmp3		np	tmp	tmp1	tmp2
		d.	o	z	d		⇒	0			4

k) global register-shift indicates segment-start after

ppnp	pnp	np	tmp	tmp1	tmp2	tmp3		np	tmp
		d.					⇒	0	2

l) pitch-repetition (note-prolongation) indicates *start on the beat* which is repeated

ppnp	pnp	np	tmp	tmp1	tmp2	tmp3		np	tmp
		d.					⇒	2	1

m) duration-sequence or quasi-sequence gives start at sequence-border

ppnp	pnp	np	tmp	tmp1	tmp2	tmp3		np	tmp	tmp1
		d.					⇒	2	1	2

n) change of direction gives prominent start-indication after (when on xl)

ppnp	pnp	np	tmp	tmp1	tmp2	tmp3		np	ptmp	tmp
		d.					⇒	1	2	2

o) change of division gives slight start-indication at change

ppnp	pnp	np	tmp	tmp1	tmp2	tmp3		np	tmp	tmp1
		d.					⇒	1	0	

**Global pitch set change**

This analysis is based on the general assumption of local pitch memory within melodies and the general assumption of perception of pitch qualities to groups of pitches (see section 3.3). If a certain set of pitches have been used over a period which encompass the limits of the perceptual present in terms of category perception and this set is replaced by another set which is in use over a comparable period it is likely to be recognized as a global structural change comparable with global register shift.

However, since most such noticeable changes in melodies (usually related to *change of tonality*) do not imply a change of a majority of the pitches in the pitch set, but rather a minority, the local position of the change is usually not unambiguous and leaves normally a

number of beat positions where both pitch sets apply. Therefore, the start-implications of this structural factor will be determined by the interaction with other structural factors.

### Weighted segment start indication

The different structural indications given by strong sequence analysis and the local and global discontinuity analyzes respectively are combined into a weighted value representing the total segment start indication for every beat in the melody.

The numerical impact of the different factors can be read from the following table (ordered from the strongest factor to the weakest)

Table 6. Minimum and maximum impact of different Global structural discontinuities

	<i>min</i>	<i>max</i>
local discontinuity	-2 (after end -4)	7
global rhythm break	-1	4 (end 7)
strong sequence + eval. sequence	0	3+2
strong seq./eval. seq.	0	3
global-reg-change	0	3
global-dir-change	0	3
global-pitch-set-change	0	2

Since the start-indications values interact in the different analyzes, this table does not show the impact of the different structural factors *per se*. The implicit principle is the relative prominence of *interonset* discontinuities over pitch discontinuities, generally reflected in a 2:1 value relation, with 6 being the highest singular value given solely by *interonset* discontinuity and 3 as maximum singular value for pitch discontinuities. (See *Principles of low-level grouping* section 5.2.4.3, rule no 4)

The relative impact of sequence versus discontinuity on this level is also not easily determined by looking at the different ranking of structural factors above: *sequence start* is an influential factor inherent in the analysis of local discontinuities and local and global discontinuities are influential factors in the evaluation of sequences. A strong sequence is determined by discontinuities and the total marker value of a *strong sequence start* will be supported by discontinuity values. Hence, a *strong sequence start indication* can overrule any discontinuity start indication within the sequence if it is supported by strong enough discontinuities at *sequence start*.

*Start indication* entirely by sequence will, however, not be able to overrule the strongest possible discontinuity indication locally, which reflects the notion of the prominence of

discontinuities as singular determining factor locally. (see section 5.2.4.3 *Assumptions no. 31-32*). The estimated maximal impact is about half the value of discontinuity.

Indication solely by sequence		Indication solely by discontinuity	
<i>min</i>	<i>max</i>	<i>min</i>	<i>max</i>
0	8	-2	17

Here is an example of the weighted analysis. The total value is basically acquired through addition of the separate indications noted on the rows below the total value.

Global rhythmic breaks are noted within brackets that show which indications belong to same break.

## Norafjells

played by *Andres Rysstad, Setesdal, Norge*

simplified into a single melodic line

transcr. SA -92

weighted total value: 16 0 3 1 5 - 2 - 1 7 1 6 - 2 - 1 9 4 3 1 5 - 2 - 1 7 1 4 - 1 2  
 global-reg-change: 3  
 global rhythmic breaks: [5] [1 2 1] [2 1 2] [1 2 1]  
 global-dir-change:  
 local marker-value:  
 sequence-values: 5 1 3 1 4 -2 -1 5 0 4 -2 -1 5 2 3 1 4 -2 -1 5 0 4 -1 2  
 3 3

6 3 4 0 5 - 2 - 1 7 1 6 - 2 - 1 9 4 3 1 5 - 2 - 1 7 1 4 - 1 2 6 3 4 0 4 - 1  
 [1 2 1] [2 1 2] [1 2 1]

3 3 2 0 4 -2 -1 5 0 4 -2 -1 5 2 3 1 4 -2 -1 5 0 4 -1 2 3 3 2 0 4 -1  
 3 2 3 2

3 2 3 12 - 2 - 1 11 4 3 7 3 2 1 4 - 1 3 2 3 12 - 2 - 1 11 4 3  
 2 2 [1 2 1] 2 2 [1 2 1]  
 1 2 1 1 2 1 1 1  
 2 0 1 4 -2 -1 3 2 3 3 3 2 0 4 -1 2 0 1 4 -2 -1 3 2 3  
 3 3+2 2+2 1 3 3+2

7 3 2 1 4 - 1 2 2 3 10 - 2 - 1 7 6 0 6 2 0 5 - 1 4 9 3 6 0 8 - 1  
 2 2 2 2 1 2] [1 2] [1 2] [1 2]  
 3 3 2 0 4 -1 2 0 1 4 -2 -1 4 3 0 4 0 0 4 -1 2 3 3 2 0 4 -1  
 2+2 1 2 2+2 2

Figure 5-33. Beginning of Norafjells, played by *Andres K. Rysstad (1893-1984), Hylestad, Setesdal, Norway rec. 1973*. Ornamentation, accompanying notes and articulation removed.

### 5.2.4.7 Tracking periodical grouping – composite pulse and time

#### Overview

The rating of *segment start indication* for each atomic beat described above is the input of the tracking and evaluation of periodical grouping described in the following. The purpose of this part of the analysis is to find a possible composite grouping and to evaluate which of these are most structurally prominent and thus most likely to be recognized by listeners. The output of this exercise can be interpreted as a first (and a second) guess of prominent metrical grouping

The different steps are as follows:

- 1) Collection of all possible strict periodicities of atomic beat grouping, from 2 beats up to 24 (16). The periodicities, here designated '*metrical grids*', are ordered in three different groups representing three different '*beat scopes*', delimiting periodicities within a range of 6 atomic beats ('*basic grid*' added at step 4), 12 atomic beats, 24 atomic beats or 48 atomic beats respectively, with a period length of maximum half the beat scope. The different views that these beat scopes represent are the different levels of local and global periodicities, which concern metrical grouping on a regulative level.
- 2) Collection of periodicities based on regularities in grouping of beats derived from the relative strength of marker-values, *asymmetrical metrical grouping*.
- 3) Successive initial evaluation of all possible strict periodicities based mainly on onset information, resulting in the removal of non-supported periodicities
- 4) Successive evaluation of the strongest marked periodicities based on *marker-values*, consistency of periodicity and consistency in hierarchy of periodicities, resulting in a rank of the most strongly marked periodicities at every atomic beat event at the different beat scopes.
- 5) Final evaluation of the dominant metrical scope: 12-beat, 24-beat or 48-beat, basic-grid or asymmetrical, giving a result with a maximum of three indications <12/24/48-grid> <basic-grid> <asymmetry-grid>. This evaluation is based on the sum of marker-values for each scope, consistency of the grids, the invert number of metrical grid changes at each scope and the degree to which the grids in the scope incorporate the melody.
- 6) On the basis of step 5 a hierarchical metrical interpretation is made, which includes time (signature) and if the result is clear regarding basic grouping, a consistent subgrouping on composite beat level is performed. When an asymmetrical grouping on composite beat level is preferred, this is evaluated by a separate analysis (see below).
- 7) The output is an interpretation of meter at composite beat level, which will be denoted as the new tactus or central beat level and an interpretation of time, represented in the notation by time signatures. Of these two, the interpretation of beat level is final and will be used further in the analysis, under certain circumstances the interpretation of time can be changed due to the phrase and section analysis which is performed after the metrical analysis.

As described above, the result is a rating of periodicities, which result in different possible output of the analysis: For instance, the preference for strict metricity, metrical symmetry, differs quite remarkably between people (see discussion of results section 5.4.4.3) due to perhaps both individual, cultural and social factors. It is perfectly possible to allow these

different preferences be reflected in the result through a global metrical preference variable, which would set the degree to which a symmetrical choice is preferred.

### Method and principles – retrieving metrical grids

The two different means of retrieving metrical grids (steps 1 and 2) referred above, represent not only two different means of extracting periodicity, but also the two different metrical preferences mentioned above, namely symmetrical vs. asymmetrical metricity or periodicity driven vs. gestalt/group driven periodicity. In the first case, hierarchical symmetrical periodicity is presupposed, whereas, in the latter, periodicity is more a result of grouping.

#### **Beat scopes – 12-, 24, 48**

As mentioned above, symmetrical periodicities are initially collected in three different contexts or ‘beat scopes’. These encompass 12, 24 and 48 beats respectively, with a period length of maximum half the beat scope. For the 12-beat-scope the maximum period-length will be 6 atomic beats.

The underlying reason is that both local and global gestalt principles influence the conceived metrical grouping. Which metrical grouping will be discerned depends not only upon the primary grouping principles reflected in the local segment start indications expressed by marker values, but also, to a large extent, on secondary and tertiary grouping principles such as good continuation/constancy, symmetry, articulation and prägnanz. (see *Principles of low-level grouping* section 5.3.2.2). The secondary grouping principles of good continuation and symmetry imply that grouping can even be inferred from regularities in grouping and qualities of groups, while the articulation and prägnanz rules are mainly concerned with the perceptual and conceptual choice among different groupings. In these cases, the influence of the grouping principle will be related to the number of events it concerns.

However, according to the general principle of metrical hierarchy (*Assumption no 15*) a complex metrical level must correspond to a basic metrical level. Time is according to this principle ‘grouping of beats’. Thus, a global periodicity must be in concordance with local periodicity. On the other hand, as noted above, a longer perspective can be needed both in order to discover and evaluate periodicities.

This is related to the perceptual present, which is here defined both in terms of simultaneous category conception (*The categorical present*) and in terms of time-span. (See *Principles of low-level grouping* section 5.3.2.2). For a certain periodicity to function as a metrical grid, a pulse or time in the sense of temporal mapping of the musical events, a prerequisite is that it creates expectation of events to happen in relation to this periodicity. An expectation of continuity we need not only to conceive the period, but also to experience that it is continuous. This requires at least three occurrences, two to constitute the period and one to point at the continuation.

On the smallest *group size levels*, groups of two or three events, this requires at least nine events for group sizes of three event, and six events for groups of two events. To establish a period at least 6 events for groups of three events are needed. This is the primary grouping level for complex pulse at central pulse level in complex metrical analysis.

If this is the central pulse level (*Assumption no 15*), it follows that the next level is the typical minimum level of measure consists of at least two basic groups, which then must

consist of at least 6 beats. Hence, for a group at minimum measure level to be established as a period at least 12 beats must be considered. Three periods of a total of 18 beats are necessary to create expectancy.

This is exactly what is considered within the ‘12-beat scope’, at least three occurrences of a maximum period of 6 beats or within 12 beats. Hence, the 12-beat-scope is regarded as the perceptual present in the sense of atomic pulse level, a melodic ‘now’ or local level. The 24-beat-scope consists of two such ‘nows’. This allows a maximum of a 12-beat-period (medium measure level) within 32 beats; the 48-beat-scope considers periods up to 24 beats within 72 beats, which is used as the upper limit of metrical period at measure level, in the sense of primary temporal map.

The reason for this upper limit is that here is periodicities which is regulative, those periodical levels which have the quality of providing the primary temporal grid on which the phrase structure is formed in focus. Higher level periodicities, such as a 32 bar song period or a 12 bar blues, are formed by the melodic (or sometimes by the harmonic) structure and are not a prerequisite for understanding which events are on beats or not on beats. The argument can be put as follows: The more sub groups a period contains the less it will be possible to experience as a undivided whole, the less its impulse quality. A period of 24 beats decomposed into a unique set of groups of three beats and two beats will cover the largest number of categories of the categorical present, 9 (7+2) groups, involving the largest number of atomic beats per group: Six groups of 3 beats and three groups of 2 beats.

Every level of *periodicities - metrical grid*, which is present in a shorter *beat-scope*, is also allowed in a longer. Thus, a two- or three-beat period grid will typically be present in the *basic-grid-scope*, 12-, 24- and 48-beat-scope. Why then bother with the shorter beat scopes? The reason is that changes in periodicity which occurs within a shorter scope will be disregarded in the evaluation within a larger scope, since consistency factors will influence the result and the search is for periodicities. In the following example taken from the melody in fig. 43, this is actually happening. The 12-beat period consists in the 24-beat-scope and the 3-beat period of the basic grid, while the grouping changes in the 12-beat-scope; the latter in this case in concordance with the melodic structure while 24- and 48-beat-scope ignore changes on that level.

Figure 5-34. Example of metrical grids at the different beat scopes in the analysis of Norafjells (see fig. 43)

**Asymmetrical grid – grouping derived from marker strength**

The search for periodicities is based on the grouping derived from the relative strength of the segment start indication marker values. First, a primary grouping is derived, then the grouping is searched for periodicity. Since no ungrouped beats should occur, the marker value of every beat is considered in relation to the neighboring beats. As in the local discontinuity analysis, the four neighboring beats are considered.

However, the relative strength of the marker values is possible to interpret in different ways and can result in different groupings. If the segment start indication on one beat is much stronger than the previous, but not quite as strong as the following, it can be interpreted not only as group start on the beat in question (because of the change from the previous unmarked beat), but also as group start on the next beat (because of its stronger start indication value). Therefore, ungrouped ‘joker’ beats, can occur in this initial group analysis, as can be seen in the example below.

On the basis of the resulting grouping, the sequence analysis is performed. Certain variability in primary group analysis is allowed if the grouping is more similar than dissimilar. For example, a 2+3 grouping corresponds to a 1+4 grouping in the example below. Thus, the assumption of preference for repetitive pattern recognition is inherent in this analysis. (see *Assumption no 4* et al).

The sequence analysis is performed in the same manner as sequence analysis at phrase level, (see Chapter 6, *Metrical Phrase and Section Analysis*) which in short means that sequences with the strongest marked sequence starts and largest amount of similarity between segments will be preferred.

If a sequence length is periodically repeated (chained) such as it can be divided into large sections of the periodical subsections within the limit of 24 beats (see above), this is considered a metrical grid and will be evaluated in relation to the aforementioned metrical grids.

**Excerpt from "Blue rondo a la Turk"**

reduced into a single melody line

D. Brubeck

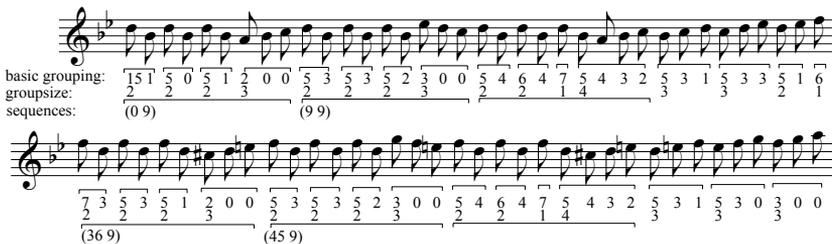


Figure 5-35. Grouping and sequence analysis

The output from the asymmetrical grid analysis is the basic grouping, the asymmetrical grid, and the evaluated sequence-list, the first of which will be competing with *basic-grid*

regarding the primary grouping level – the tactus level. The sequence list will be yet another grid at ‘measure level’.

### ***Initial successive evaluation of metrical grids***

Metrical grids are, as mentioned in the overview, evaluated stepwise. This is due to practical reasons rather than methodological. According to the general model of melody perception the evaluation is considered to be performed in real time when listening to the melody.

The first and second steps of the analysis described above, usually result in a very large number of possible periodicities, and hence the third and fourth steps are obviously crucial in retrieving the periodicities that are actually perceptible.

Initially, grids, which will not be recognized due to lack of support from the onset structure at a certain beat position, are removed from the grid list for that beat position. This is based on the assumption that phenomenal periodicity strongly supports perceived periodicity. Conversely, lack of phenomenal support and *negative support* by interfering periodicities can excise the conceived periodicity.

*Negative support* is connected with lack of onset at periods and is denoted by negative values of indication of segment border at beat positions (see marker values above). These are crucial in the evaluation of which metrical grids should be deleted.

Lack of (positive) support can also be sufficient for the extinction of a metrical grid, when there are competing grids locally. The segment indication value zero means that there is no recognizable change in the melodic structure sufficient to create an expectation of a segment start at that beat.

- A metrical grid is ended (made extinct) when there is repeated lack of onset or offset on periods (repeated negative values) in a melodic context with other onsets locally.
- Grand pause is an exception of the repetition of negative values principle, because there are no competing onsets locally. Hence, there is no competing metrical grid, and the grid can persist during the grand pause.
- Another exception is repeated and periodic ‘offbeat’ onsets, i.e. phase displaced pulse (see ass no 15, metrical complexity). However, this is only possible when the period of the offbeat and beat level are identical and within the range of primary beat grouping, thus not applicable to measure level periods.
- Lack of positive support (negative or zero values) on three successive period starts weakens the sense of periodicity repeatedly, which is regarded sufficient to end the grid. This regards also weak positive support (1) in combination with strong negative support (-2).

### ***Further successive evaluation of metrical grids***

After the initial evaluation, possible metrical grids at each beat position are ranked with regard to measured structural prominence, reflected in a total-grid-value.

This further evaluation is based on

- Segment start indication value
- Continuity and consistency of the periodicity

Chapter 5

- Hierarchical conformity between local and global levels
- ‘Prägnanz’, pattern salience in terms of
  - overall distribution of segment start values at period starts
  - distribution of segment start indications at the start/establishment of the pattern
- Generality, in terms validity of the periodicity in relation to the whole of the melody.

The basis of the total-grid-value is the sum of segment start indication values at grid positions (period starts) within the three different beat scopes. However, these individual values at grid positions can be altered due to continuity and the total sum can be devaluated if the indications generally are low.

The most important deviations from the initial marker values are as follows:

- Negative marker values (non onset) on beat positions when there is onset on the next beat position (short tie, short rest) can be changed into relatively high positive marker values when the surrounding grid positions are strongly marked.

It is a categorical difference between a non-onset position with onset on the following beat position (marker-value  $-1$ ) and non-onset without onset on the next atomic beat position (marker-value  $-2$ ). The first situation implies that it is possible to conceive this position as the position from which a structural change takes place, hence an indication of segment start. However, such an interpretation is context dependent, and is very frequently used in different types of Western music as a metrical marking as in the following example. Here is the grid (0 3), i.e. start-pos 0 and period length 3 beats, marked by a tie to the second grid position and surrounded by strongly marked grid positions, which results in a grid value of 5. (*continuity*)



Figure 5-36. Example of the “Short tie”-rule. Events within a metrical grid with no onset at the grid position but with strong start indications on the surrounding grid position may be conceived as articulated,

- More than one severely ‘feminine’ rhythmic situation within the beat-scope, i.e. short note on grid position followed by an *long* event longer than 2 atomic beats on next atomic beat position, implies a rhythmically unstable grid, which gives half the initial total-marker-value (See section 5.2.4.3 Principles of low-level grouping, applies for all of the following conditions)
- Two grid-positions with strong negative marker  $\leq -2$  within the beat-scope, i.e. non-onset at grid position without onset on the next beat or position after melody end, gives half the initial total-marker-value (*pattern salience*)
- If half or less than the first four markers are weak ( $< 2$ ) the total-value is reduced by half its value (*start salience*)
- If there are two or more strongly negative markers or less than  $2/3$  of the grid-positions within the beat scope the total-marker-value are reduced by half its value (*pattern salience and continuity*)

- If there are two-three weak markers at continuous grid positions the value is reduced to half its value (*continuity*)
  - Combinations of these can lead to a reduction to 1/4 of the initial value
- In addition to these predominantly negative factors involved in the basic calculation of the *total-grid-value*, the local salience, continuity, persistence and hierarchy of the grid pattern can give positive support through increased value as follows:
- The metrical grids for every atomic beat position are compared with the total marker-values for the metrical grids which occur at the surrounding positions, which give prominence to the metrical grid with the highest marker value within five successive beats, which makes it possible to compare all local groupings and thus trace change of metrical grid (*pattern salience locally*)
  - Local, 'basic' grids at composite beat level (2-4 atomic beats) are evaluated in relation to the competing grid based on the grouping based on the highest local values (asymmetrical grid) (*local pattern salience*)
  - The local grouping (basic grid) and lower level grouping generally influences the higher level groupings (*metrical hierarchy – symmetry rule*).
    - Metrical-grids compatible with the current basic-grid, i.e. the primary beat grouping (see above about beat scopes) are assigned an increased value up to 2.5 times the original value).
    - When the metrical grid highest in rank at the 12-beat-scope level is at the length of basic groups (2 or 3 atomic beats), compatible sequences can receive as much as twice of the original value.
  - The continuity and continuous support for a metrical grid influences its value. (*continuity – good continuation rule*)
  - If a metrical grid is successively recurring more than 11 times with period length > 3 atomic beats the total marker-value will be 3 times its basic value. If the pattern is established at more than 3 successive positions two times its value and 1,5 times its value if its been established at two successive grid positions.

### ***Change or persistence of metrical grid?***

The designation of the change or persistence of metrical grid at a beat position depends on the quantification of the total segment-start-indication value described above. This is also dependent on the consistency rules which imply that a change is more likely to be conceived if: (1) it occurs on an established grid position; (2) it is hierarchically concurrent with the basic grid; (3) and/or has a period, that is compatible with the previous grid period. If the new grid with the highest total-value at a position is concurrent with an earlier grid, it is also considered to be more easily established.

When none of these conditions occur, a total marker-value twice as large as the current total grid-value is required to change an existing grid.

### **Final evaluation of metrical grids**

When all metrical grids have been successively evaluated, a final evaluation between grids within different beat scopes is performed. The conceptual background of this procedure is the fundamental assumption of the conception of metrical structure as a learning process which is developed gradually through the experience of the melody. (*Assumption no 6*). It is assumed

that we can reconsider the metrical conception throughout the experience of a melody. There are plenty of examples of the deliberate use of metrical ambiguity in music, where the metrical structure indicated in the beginning of a piece turns out to be something else<sup>53</sup>. The results of experiment 2 (see section 5.3.4.3, *Discussion*) support the notion that metrical conception may change with new information.

The most important means of evaluating the final metrical grids are how dominant the different metrical interpretations are with regards to the whole melody, the salience expressed as total segment start values and salience in the initial phase of the metrical grids and in average at beat positions. A single congruent grid, which can divide the melody into equal periods is more likely to be preferred than a irregular change of metrical grids.

Another principle of evaluation is that a grid is considered more likely to be recognized if local and global periodicity is hierarchically concurrent. Thus, a grid, which is concurrent with a basic grid, is preferred before grids without such hierarchical concurrence.

Yet another principle is the priority for shorter periods over longer periods above basic level, hence a preference for 12-beat-scope grids over 24-beat-scope etc. According to the principle of preference for symmetric grouping (See section 5.2.4.3 *Assumption no 30*) strict periodicity is preferred before asymmetrical grouping.

The priority between beat-scopes and grid types can be read from the following list:

- a) 12-beat-scope consisting of a single or dominant and sufficiently marked grid  
If both conditions are fulfilled.  
or
  - only one grid in the 12-beat-scope list or
  - dominant grid > 3/4 of the melodyor  
and
  - average total-marker-value > 20 (> 4 at each grid position for the initial grid) and
  - either not singular grids at higher beat-scopes or
  - relatively larger average total marker value than at higher beat scopes (24-mean-index < 1,5 X 12-mean index etc.) or
  - singular grids at higher beat scopes possible within 12-beat-scopeor
  - not singular grids at higher beat scopes
  - larger average total marker value than higher beat scopes
- b) 24-beat-scope consisting of a single or dominant and sufficiently marked grid  
equivalent conditions as under a)
- c) 48-beat-scope consisting of a single or dominant and sufficiently marked grid  
equivalent conditions as under a)
- d) basic-grid, singular, dominant and sufficiently marked  
equivalent conditions as under a) plus  
or

---

<sup>53</sup> This phenomenon is frequently used in different musical styles. For example, the introduction of the beat in Reggae songs may be conceived as deliberate use of metrical ambiguity.

- no valid asymmetrical grid
  - higher average total grid value for basic grid than asymmetrical grid
  - higher initial average total grid value (6 first grid positions) than asymmetrical grid
- e) asymmetric grid, singular or dominant through the entire piece
- If all of the following are true
  - the asymmetric grid is homogenous and dominates more than 3/4 of the entire piece
  - the total average grid value > 20
  - the mean marker value per grid position > 5
  - total-marker-value of first grid of 12-beat-scope list is either < 2/3 of the first asymmetric grid or mean index value for asymmetric grid is larger than for grids within 12-beat-scope
- f) 12-beat-scope, with changing sufficiently marked grids
- If all of the following are true  
the average index value  $\geq 20$  (5 beats of at least average marker value of 4)  
larger index-values than 24-beat-scope, 48-beat-scope, basic-grid and asymmetric grids  
or
- one basic grid compatible with all grids within 12-beat-scope
  - less grids in 12-beat-scope than in 24- or 48-beat-scope-lists
  - 12-mean-index-value (average index value for all grids within the scope) > 3,5 or 12-mean-index-value > than 1/2 of 24-mean-index and 48-mean-index (since the mean index value is based on the number of grid positions involved)
- g) 24-beat-scope, with changing sufficiently marked grids  
Conditions equivalent to f)
- h) 48-beat-scope, with changing sufficiently marked grids  
Conditions equivalent to f)
- i) Changing basic-grid  
Requires average total marker value  $\geq 6$ , i.e. mean per grid position > 1.
- j) Asymmetric grid/sequence-list  
Requires dominant grid and average marker value > 4, less than 10 basic groups within the period and higher index value than basic-grid
- k) 12-beat-scope with changing grids  
Counts if the average marker value > 1 per grid position
- l) 24-beat-scope  
Equivalent condition as k)
- m) 48-beat-scope  
Equivalent condition as k)  
The underlying principle of the limit values for average total-marker-values (index values) is a categorization in: (1) very strong segment start indications > 3 at each grid position; (2) strong segment start indications > 2; (3) intermediate segment start indications > 1; (4) and low segment start indication at 1 and below. The 'very strong' category generally

requires either offset-onset indication such as short/divided beat after tie, or combinations of different indications. The ‘strong’ category requires at least combination of step-leap and change of melodic direction, intermediate e.g. average global indications and the ‘weak’ category represents e.g. simple local change of direction or step size.

This means that a grid supported by only e.g. a change of melodic direction is regarded structurally significant only when there are no competing stronger indicated grids.

Regarding larger beat-scopes the output of this analysis typically includes the basic grid and/or an asymmetrical grid, which is sufficiently marked according to the above conditions. It is, therefore, possible to let the interpretation depend on the global preference for symmetrical and asymmetrical grouping respectively. If no such preference is made the method itself chooses between these interpretations based on the evaluation described above.

If an asymmetrical grid is preferred, a final primary grouping analysis is performed, (see section 5.2.2.8 below).

### Examples of metrical interpretations – symmetrical vs. asymmetrical measure structure

The process of analysis and evaluation of metrical grids described above can be difficult to comprehend in terms of musical consequences from the description of the method alone. Here are three examples of how meter tracking is performed in different contexts.

#### *Symmetrical meter – ‘Moto perpetuo’ type*

The ‘obligato’ melodic part in ‘*Jesu Bleibet Meine Freude*’ (from cantata no 147, BWV 147) by J.S. Bach is an example of a melodic structure in which the *Beat-atom Analysis* (see section 5.2.3), which depend on onset information is not capable of presenting a structurally stable metrical solution at central pulse level. The output of that analysis is an atomic beat level of eighth notes given that this is the shortest duration in the MIDI input.

The notation example shows the output of the *Beat-atom Analysis* in common notation and the result of the successive evaluation of metrical grids. Note that the asymmetrical grid is the raw grouping made from the strongest start indications, not the sequence-based analysis obtained from these values.

The basic grid for the entire piece is (0 3) which means that the primary grouping (the central pulse/tactus level) in three atomic beats receives the highest total-marker-value throughout the piece. In comparison to the raw asymmetrical grid, this interpretation is supported in the beginning but with frequent later deviations. Thus, the deviating evidence is disregarded in the light of the initial consistency of the basic grouping and the total support.

Moreover, while the 24-beat and 48-beat scopes give the same singular result at higher metrical levels – measure level – the 12-beat scope gives a changing metrical result. This is due to the fact that a group size of nine beats exceeds the limit of groups within the 12-beat scope. The final choice between these metrical interpretations is as been mentioned performed on basis of an index value comprising the total marker values per grid and beat scope and the dominance and singularity of the metrical grids.

Here there is no dominant-12-grid and hence no dominant-12-region. Conversely, there is both a dominant 24-grid and a dominant-48-grid (which are identical, see above). The

# Jesu Bleibet Meine Freude

excerpt from obligatto melody

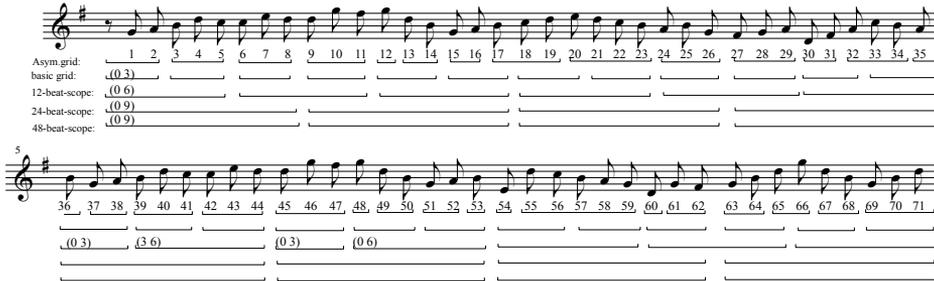


Figure 5-37. Initial metrical grid analysis performed on the obligatto melody from “Jesu, Bleibet Meine Freude” by J.S. Bach. Metrical grids are displayed by brackets below the systems for the corresponding metrical scope.

dominant-24-grid covers 76% and the dominant-48-grid covers 97 % of the melody. The latter is then regarded singular and is chosen at measure level.

The basic grid is also strongly supported with a basic/asymmetry-mean-ratio of  $\approx 8.69$ , which means that the average marking of the grid at each basic position is more than eight times the value of the asymmetrical grid markers. Since the latter are based on the strongest markers this difference is only due to the enhancement of basic grid values due to continuity and conformity with higher levels.

Hence, the result of the analysis will be the (0 9)-grid of the 48-beat-scope list and the basic grid of (0 3), resulting in a metrical notation in 9/8:

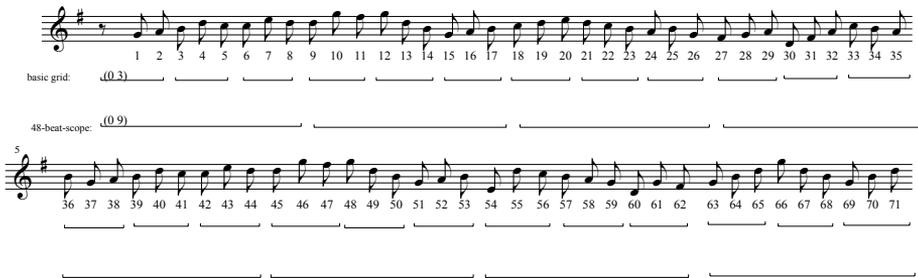


Figure 5-38. Selected metrical grids after the metrical grid evaluation in excerpt from “Jesu Bleibet Meine Freude”.

The figure below displays the output from computer implementation: Note that the hooks at the top of the system indicates beat grouping.

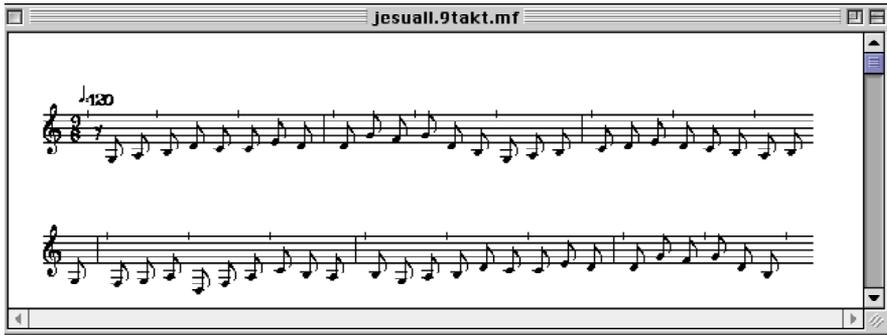


Figure 5-39. Output from computer implementation displaying the result of the Metrical Analysis of “Jesu Bleibet Meine Freude” by J.S. Bach

### Asymmetrical meter at primary grouping level – composite meter

Asymmetrical beat grouping (additive rhythm, composite meter) is often regarded as a typical feature of e.g. dance music in the Balkan and Turkey or Indian art music.

Here is a short melody constructed for test purposes, inspired by the Turkish ‘aksak’ time.

## Test melody for asymmetrical meter

SA/Inspired by Aksak time

Asymm grid: (0.9) (1.6) (0.2) (0.2) (0.2)

Basic grid: (0.3) (0.3) (0.3) (0.3) (0.3) (0.3) (0.3) (0.3) (0.3)

12-beat scope: (0.6) (0.6) (0.6) (0.6) (0.6) (0.6) (0.6) (0.6) (0.6)

24-beat scope: (0.9) (0.9) (0.9) (0.9) (0.9) (0.9) (0.9) (0.9) (0.9)

48-beat scope: (0.9) (0.9) (0.9) (0.9) (0.9) (0.9) (0.9) (0.9) (0.9)

Figure 5-40. Metrical grid indications at different beat scopes obtained by analysis of primary grouping indications

As in the similar example, ‘Blue Rondo a la Turk’, (see fig. 45 above) the asymmetrical grid here results in a chain of sequences of the same total length and with compatible grouping indications:

(0 9): (2+2+2+3) (2+2+1+4); (9 9): (2+2+1+4) (2+2+2+3), (18 9): (2+2+2+3) (2+4+3).

The groups are compatible, but not always alike and have at least two group-sizes/changes in common. The chain of sequences covers the entire melody.

The basic grid is not consistent and, furthermore, has a lower total value than the asymmetrical grid. (The first basic/asymmetry-ratio being 0.624)

At measure level, the single grid of the 24-beat-scope overrides the changing grids of the 12-beat-scope and the result will be the 24-beat-scope grid list and the asymmetrical grid, in this case, giving the same result, dominant meter with a period of 9 atomic beats per period.



Figure 5-41. Final metrical grid evaluation in test melody for asymmetrical meter (see fig. 50)

However, the dominance of the asymmetrical grid and the lack of any basic grid makes it necessary to perform a separate beat group analysis to obtain the central pulse level, which benefits from the asymmetrical grid analysis (This further analysis is described in section 5.2.4.8) The output of the computer analysis, including the beat group analysis, is shown below:



Figure 5-42. Output of the complex metrical analysis performed on the test melody in fig. 50

In certain musical styles, the periods at the measure level can be extensive. A particular example of this is Indian classical music. The raga 'Raag *Suddha Danyasi*', (see fig. 43) from the Carnatic Classical repertoire, provides a good example.

In this example, the larger periods are quite symmetrical. Although the entire piece is considered, the only high-level beat-scope with a singular grid turns out to be the 48-beat-scope, which is dominated by the 16-beat grid. The apparent hierarchical symmetries established in the beginning 16-8-4-2 are later obscured. The 2-beat basic grid, however, persists through the piece. This level is challenged by a more strongly marked asymmetrical grid, which is why the initial output becomes the 16-beat period.

The example in fig. 44 is thus notated in 16/8. Further grouping analysis (see section 5.2.4.8) is necessary to establish the final grouping grouping at tactus level.

# Raag Suddha Danyasi

simplified tr. in western notation/repeats removed

after K Shivakumar

## RESULT OF METRICAL GRID ANALYSIS

asymm-grid: (0 12)  
 asymm-seq: (0 2)  
 basic-grid: (0 4)  
 12-grids: (0 4)  
 24-grids: (0 8)  
 48-grids: (0 16)

3  
 (38 6) (54 8) (0 6)

5  
 (0 4) (89 5) (0 6)

Figure 5-43. Result from initial metrical grid analysis of “Raag Suddha Danyasi”, transcribed from sa-re-ga notation. Repetitions of lines (32 beat level) are omitted.

## RESULT OF EVALUATION OF METRICAL GRIDS

asymm-grid: (0 12)  
 asymm-seq: (0 2)  
 basic-grid: (0 4)  
 12-grids: (0 4)  
 24-grids: (0 8)  
 48-grids: (0 16)

3

5

Figure 5-44. Result from final evaluation of metrical grids in Raag Suddha Danyasi

***Asymmetrical meter at measure level –compound meter***

Many ‘*gangar*’ and ‘*balling*’ tunes of the Norwegian folk music tradition are good examples of pulse-orientated music, where the metrical grouping at the measure level do not have the regulative character that is true in many other musical contexts. The metrical grouping at higher levels can thus be subject to variation. More or less recurrent changes of the size of the periods at measure level, are more result of melodic phrasing than a regulator of melodic phrasing. Consequently the conception of this metrical level by listeners may vary considerably, and the reality of metrical structure at measure level be questioned. (see Blom & Kvifte 1986:515). I will return to these questions in the discussion at evaluation of results (see sections 5.3.3.4 and 5.3.4.4).

Nevertheless, the variation of metrical grouping at higher levels is not arbitrary. The variation of measures length of in the same melody found in the vast majority of the tunes in “*Norwegian Folk Music, series I*” (Gurvin 1959) is very restricted and generally regular.

One of the most important means of determining higher metrical levels in such music is through phrase analysis. Since this is included only to some extent in the segment start indication analysis (see sections 5.2.4.5 and 5.2.4.6), the metrical structure can be altered by the later comprehensive phrase and section analysis. The purpose of the initial metrical analysis is to provide the beat analysis at central pulse level, which is a prerequisite for the phrase analysis. Thus, the focus of the metrical analysis at measure level here is to determine the metrical grouping at measure level to the extent that it affects the metrical grouping at beat level.

The current example is an excerpt from a ‘*gangar*’ tune (see fig. 43). The resulting grids at all parallel scopes are shown below in fig. 45

None of the grids in the figure below, except the basic-grid, is singular within the whole melody. And, the basic-grid (three atomic beats per group) is not dominant enough to rule out the grouping of the asymmetrical grid and asymmetrical sequence entirely since the mean value at the establishing phase is too low. The 12-beat period of the 24-beat-scope is dominant in  $\approx 68\%$  of the melody, which is not sufficient for assumed singularity. Instead, the changing grouping of the 12-beat scope will be the result of this analysis, and together with the basic grid and the asymmetrical grid, it will be subject to further grouping analysis. One of the main features of the grids of the 12-beat-scope is that they all are compatible with the singular basic grid.

The output of this stage of the analysis is in the computer implementation resulting in a barring of the tune, which implies a change of time according to the 12-beat grid. The grouping analysis in this example will be discussed further in section 5.2.4.8 below.

# Norafjells

played by *Andres Rysstad, Setesdal, Norge*  
simplified into a single melodic line

transcr. SA -92

48-beat-scope (0 12)  
24-beat-scope (0 12)  
12-beat-scope (0 6)  
basic-grid (0 3)  
asymm-grid  
asymm-seq-list (0 12)

48:  
24:  
12:  
basic:  
ass:  
asseq:

basic:  
24:  
12:  
basic:  
basic:  
ass:  
asseq:

48:  
24:  
12:  
basic:  
basic:  
ass:  
asseq:

Figure 5-45. All metrical grids found by the initial metrical grid analysis of Norafjells as played by *Andres Rysstad* (see fig. 43)

NORAFJELLS: INITIAL RESULT OF METRICAL GRID EVALUATION

12-beat-scope basic-grid (0 6) (0 3)  
asymm-grid

12:  
basic:  
ass:

12:  
basic:  
ass:

12:  
basic:  
ass:

Figure 5-46. Result of final evaluation of metrical grids in Norafjells after *Andres Rysstad*

### 5.2.4.8 Final analysis of primary beat-grouping

#### Overview

The main purpose of the metrical analysis, which is of crucial importance for the phrase and section analysis, is to obtain a beat analysis at central pulse level. This is because of the assumption that metrical organization, whenever present, function as the fundamental temporal mapping of the music. (*Assumption no 1*).

As pointed out in the examples above (section 5.2.4.7), the successive evaluation of metrical grids does not always come up with a regular primary grouping at central beat level (a *basic grid*). Or, even if it does, it can be strongly contradicted by the grouping indications, which originates from grouping between the strongest segment start indications.

When the metrical analysis have identified an atomic pulse and possibly at measure level but has failed to find a consistent periodicity at composite beat level, the primary grouping at beat level remains to be analyzed.

The previous local grouping analyzes performed within the successive evaluation of periodicity at higher levels does not provide sufficient information for this analysis to be performed, since the existence of a higher metrical level (or the lack of existence of a higher metrical level) affects the local grouping, so that the local grouping has to concur with the higher level grouping. While this analysis is assumed to be parallel with all the other segmentation processes in real listening experience, it is performed in separate steps in the method for the purpose of evaluation of the method.

Thus, yet another analysis of the local grouping must be performed with the new premises given by higher metrical grouping (at measure level) and the non-existence of a salient symmetrical grouping at composite beat level.

This analysis is generally performed in the same manner as the first analysis of primary grouping. In fact, the analysis of “*strong sequences*” is performed by exactly the same algorithms. The successive analysis of local primary grouping is, however, slightly modified, because of the influence of metrical grouping at measure level.

In spite of this modification, the fundamental structural factors for determining segment borders are essentially the same as in the prior local discontinuity analysis. Below, the different factors of primary grouping indications are replicated with the new valuations given in this analysis below. An overview of all the conditions for segment start indication in this part of the analysis can be found in the The most important difference to the prior local segment start indication is the stronger emphasis given to accentuation interpretation relative distance interpretation, regarding segmentation indication by duration. This reflects the assumption that the major segments have already been found – this part of the analysis concern sub-grouping at beat level.

**Primary grouping indications**

- a) Grouping by change of **duration/interonset** interval:

DURATION

a) xlong > 2 X beat

b) xlong = 2 X beat

c) divided beat/ornament



- b) Grouping by change of **onset/not onset/offset** (related to a)):

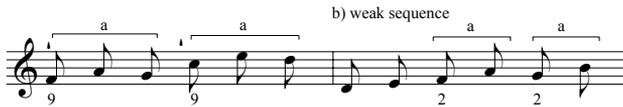
a) onset-not onset

b) before and after rest

c) before and after xlong rest



- c) Grouping by pitch pattern similarity: **Sequence**



- d) Grouping by pitch continuity 1: **Repetition** (note prolongation).

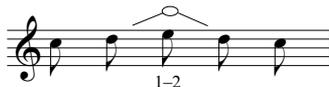


- e) Grouping by pitch continuity 2: **Return**

RETURN



- f) Grouping by pitch discontinuity: **Melodic direction**



- g) Grouping by pitch distance: **Step-size / register**

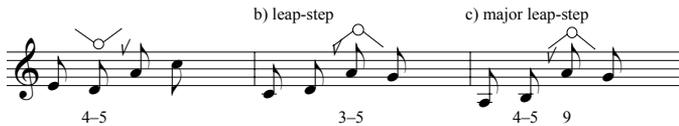
STEP-SIZE CHANGE

a) small distance

b) large distance



h) Composite pitch change 1: **Melodic direction and step-size**



i) Composite pitch change 2: **'Roundabout' / intervallic return**

ROUNDBOUT / INTERVALLIC RETURN



**Evaluation of primary grouping by secondary and third grouping principles**

The interpretation of the above segment start indications depend as have been mentioned above, on secondary and tertiary grouping principles as well as primary group constraints (section 5.2.4.3). These are here replicated for the purpose of clarity:

- a) Primary group constraints (*grouping principles no 24-26*) such as:
  - The primary grouping concerns the grouping of atomic beats. Hence primary groups at central pulse level cannot consist of a single atomic beat.
  - Avoid groups constituted by single events.
  - Primary groups consisting of two or three atomic beats are preferred.
- b) Metrical hierarchy (*Assumption no 14*)
  - Primary grouping at beat level conform to higher level grouping at measure level. Hence primary grouping, which interferes with measure borders, is avoided.
- c) Good continuation (*grouping principles no 29*)
  - Continuous group size is prominent.
  - Continuity in grouping of higher metrical levels is preferred. Thus grouping within measures, which is coherent with earlier grouping is prominent. Grouping is thus assumed to be inherited.
- d) Symmetry (*grouping principles no 30*)
  - Grouping which divide a measure in symmetrical parts is prominent.
- e) Gravity (*grouping principles no 31*)

- Start on articulated, salient event is prominent.

f) 'Prägnanz' (grouping principles 32)

- Groups consisting of a few contrasting elements are prominent.

Primary group constraints and metrical hierarchy are implemented in this part method through sub-grouping by measures, which means that the measures are analyzed one by one. This implies that the position of the event in relation to the borders of the measure is a basis for the evaluation: The first position is given a segment start indication (2), while the positions next to measure borders generally get zero values. Also, symmetry is considered in relation to metrical hierarchy implying that symmetric division of a measure is preferred to asymmetric, which gives higher value to the mid-position of a measure, especially regarding larger measures.

This dependency of measure analysis may seem risky, since the metrical analysis at measure level is preliminary and is supposed to be dependent on phrase analysis. Even if this is true the metrical grouping at measure level is based on structurally prominent *segment start indications* and is interpreted only when these indications are evident enough to give a clear indication. Thus the sub-grouping will be subordinate to this grouping indications.

The group size preference rule is interpreted in a similar fashion as in the successive evaluation of metrical grouping, considering possible group size. This means that the position after a identified group start is labelled an *avoid* position, which requires a stronger segment start indication to get a higher value, reflecting the avoidance of single elements in groups. When the event position is concurrent with the preferred group size (2-4 elements to the previous group start) and is of the same size as the previous group, it is regarded as an *on-group* position, which reinforces its start indication values. Thus is good continuation on sub-group level implemented in the method of analysis.

In fact, the method assumes that every measure can be decomposed into primary groups of 2 and 3 atomic beats, accepting larger groups only when there is no local grouping indication at the 2 and 3 group level.

Continuity in grouping of measures is implemented in the method through the preference for similar grouping of measures of the same size, even when segment start indications at primary level are very weak.

The design of the evaluation of the grouping indications is summarized below:

1. Analysis of the local segment indications for each measure in relation to position in measure (*start-* and *avoid-position*), preferred primary group size (*on-group*) and preference for continuous group size (*on-group*).
2. Evaluation of the grouping indications for the current measure. Possible continuous grouping in 2, 3 and 4 atomic beats (*2- 3- and 4-grouping*) is compared to grouping obtained from the strongest indications in groups of 2 or 3 beats (*asymmetrical / uneven-grouping*) and to non-grouped atomic beats (*beat-grouping*). The asymmetrical grouping in 2 or 3 beats considers the possibility of either groups of 2 or 3 and evaluates which of these interpretations is most strongly indicated by the local segmentation indications from each group to the next. This can result in grouping in 2+3+3+2 etc. If more beats have no segment start indication bigger groups can be formed, such as 2+4+3+5 etc.

When a previous measure exists, also the grouping indication value corresponding to the grouping of the previous measure is compared to the above interpretations (*prev-grouping*).

The evaluation of these different interpretations is based on the rules of good continuation and symmetry through preference for continuous group, and preference for equal grouping of contiguous measures and above all for grouped rather than ungrouped atomic beats. The valuations of different grouping indications can be summarized as follows:

- a) Preference for symmetrical grouping (contiguous groups of equal size) over asymmetrical grouping is implemented through the condition that asymmetrical grouping is considered valid only
    - when the symmetrical grouping positions is not a subset of the asymmetrical grouping positions and
    - the total segment start indication value in relation to number of groups is substantially lower than the total indication value of the asymmetrical grouping. (typically  $< 3/4$  for 2-grouping and  $< 1/2$  for 3-grouping)
  - b) Grouping concurring with previous measure (*prev-grouping*) is preferred when none of the other grouping interpretations have higher average start indication value than *prev-grouping* or when it concurs strongly with the asymmetrical grouping and has higher average start indications values than the symmetrical groupings.
  - c) Symmetrical grouping indications are evaluated based mainly on the average total start indication value per group
  - d) Measures which contains less than five atomic beats falls within the limits of the preferred primary group size (2-4 atomic beats) and are generally not sub-grouped but considered primary groups.
3. The evaluation of the grouping indications can when indications are weak or ambiguous initially result in non-grouped measures, which won't be perceived as such due to the '*good continuation*'/'*constancy*' principle (*sec. grouping principle no 6*) which makes us infer an established grouping when it is expected and not strongly contradicted by the structural indications. The initial evaluation also typically results in numerous different groupings with coinciding or compatible positions, which will be perceived to have the same grouping, also due to the *good continuation* principle.

Which grouping will be perceived is also influenced by the *prägnanz* principle (*third grouping principle no 9*), which states that groups consisting of a limited number of contrasting / unique elements is preferred before groups consisting of many not clearly delimited or non unique elements. This implies that sub-grouping of measures should conform to the principle of a limited number of unique, contrasting primary groups.

Thirdly, the choice between different grouping possibilities is also influenced by the *symmetry* principle (*sec. grouping principle no 7*), which states that consistent equal group size is preferred. This implies that a grouping conforming to this principle will be preferred before a grouping, which contains groups of very different sizes.

Hence, a second step in the evaluation of the grouping indication is performed, where all measures are compared and for each compatible grouping the grouping which contains the least number of unique sized groups, most similar in size, is chosen as the standard grouping. The different compatible groupings are then substituted for the standard grouping.

Ungrouped measures are evaluated regarding the total segment start indications for previous groupings, either the grouping indications corresponding to the previous measure or if there is a sequence of sub-grouping of measures (such as (2+3, 3+3,) (2+3,

3+3)), the grouping indications of the measure at a corresponding position in the sequence are quantified. These grouping valuations are then compared to grouping indications of all standard type groupings and the grouping with the highest relative valuation is inferred, with preference for *prev-grouping* all indications values equal.

Thus, no ungrouped measures will remain except for short measures conforming to the preferred primary group size (2-4).

### Examples of final analysis of primary metrical grouping

The practical implications of the method described above, is perhaps easiest to comprehend from concrete examples.

### Recurrent asymmetric grouping in asymmetric polska tunes

Below is an example of output from the initial successive evaluation. Note that there are three different groupings and one ungrouped measure. The high values of the first beats of the subsequent measures is due to a definite start placed on that position by the analysis of the previous measure.

**Polska efter Bleckå**  
*Played by Gössa Anders Andersson, Orsa, Dalarna*

OBS! Simplified version, ornamentation and short ties omitted. Only first 8 bars.  
 Primary grouping obtained by the initial analysis showed by beaming  
 original segment start indication values below system

Bar 1: asymmetric grouping (2-4 1 0 0 0 8-4 0)  
 Bar 2: asymmetric grouping (9 0 8-4-4-4-4-4 0)  
 Bar 3: no grouping (9-4 0-4 0-4 0-4 0)  
 Bar 4: asymmetric grouping (2 0 8-4 0-4 8-4-4)  
 Bar 5: prev-grouping (9-4 1 0 0 0 8-4 0)  
 Bar 6: prev-grouping (9 0 8-4 0-4 8-4-4)  
 Bar 7: prev-grouping (9 0 8-4 3-4 2-4 0)  
 Bar 8: 3-grouping (9 8-4 8-4-4-4-4-4 -)

Figure 5-47. Local start-oriented grouping indications in “Polska efter Bleckå” played by Gössa Anders Andersson (1883-1964), Orsa, Sweden (See recordings). Simplified version (ornamentation omitted) derived from the metrical quantization analysis of the audio input.

The groupings occurring in this example is 2+4+3, 2+7 and 3+6 and one ungrouped measure. The first two of these groupings are compatible, and since 2+4+3 conforms to the principle of most similar group length the 2+7 is substituted by 2+4+3 in the standard

grouping analysis. The ungrouped measure inherits the grouping of the previous measure (now transformed into 2+4+3) since the segment start indications conform better to this grouping than to the 3+3+3 which has negative segment start indications for the second group start, even if it is not supported by the segment-start indications.

The final result of the analysis performed by the computer model is displayed below. The metrical grouping is indicated by small hooks at the top of the systems.



*Figure 5-48. Final metrical grouping analysis obtained by the computer model in polska after Bleckå.*

The addition of the original ornamentation gives a little more input to the grouping analysis but does not change it.



*Figure 5-49. Output from computer model, displaying metrical analysis based on the metrical quantization analysis of input from recording of Polska after Bleckå.*

### ***Asymmetrical grouping in South Indian (Carnatic) Classical music***

*Varnams* (composed melodies) of South Indian art music provide excellent examples of very complex grouping patterns at primary level as a result of the *tala* system, the additive rhythmic principles according to which melodies are composed. Many of these compositions can be regarded as an elaborated play with grouping expectations, which requires a great deal of cultural learning to fully comprehend.

Chapter 5

Some aspects of this is, however, possible to trace by this method of analysis. Below is the initial analysis of such a varnam, 'Raag Suddha Danyasi' (see also section 5.2.4.7 and 6.3).

The figure shows three systems of musical notation for Raag Suddha Danyasi in 18/8 time. Each system includes a treble clef, a key signature of one flat, and a tempo marking of 140. The first system is labeled '4-grouping' and 'asymmetrical grouping' with a start valuation of (2-4 0-4 1 0 8-4 1 0 1 1 3 0 8-4) (9 0 0 9 0 0 9 1 1 0 1 0 3 1 3 0). The second system is labeled '3-prev-grouping' and 'prev-grouping' with a start valuation of (2 0 1 3 0 0 1 0 9 0 1 0 9 0 1 0) (9 0 0 9 0 0 9 1 2 0 3 2 3 0 2 0). The third system is labeled '5-4-grouping' and '4-grouping' with a start valuation of (9-4 3-4 1 0 2 1 1 1 1 0 8-4-4-4) (9 0 1 1 3 0 1 0 1 0 1 0 3 1 3 0).

Figure 5-50. Initial metrical grouping analysis performed on Raag Suddha Danyasi. The original grouping analysis performed within the metrical grid analysis indicated by beaming and hooks at top of the systems. Below the systems the local group start valuation is displayed.

The resulting groupings are in this case exactly the same as those obtained by the initial metrical grid analysis:

The figure is a screenshot of a music player window titled 'danyasi.tot.mf'. It displays four staves of musical notation for Raag Suddha Danyasi. The tempo is marked as 140. The notation shows the final analysis of metrical grouping at the central pulse and measure level, obtained by evaluation of asymmetrical grouping indications.

Figure 5-51. Final analysis of metrical grouping at central pulse and measure level obtained by evaluation of asymmetrical grouping indications.

The level of concurrence with the notated groupings provided by the south Indian performer will be covered more thoroughly in the evaluation section (see section 5.2.3.5). As an example, the same section as notated by the performer in the *sa-re-ga* score is displayed below. If we compare the above grouping with the grouping notated by the performer in same section it is obvious that most of the groupings retrieved by the analysis concur with the notated groupings, especially regarding compatibility.

## Raag Suddha Danyasi

Grouping according to original notation

after K Shivakumar

Figure 5-52. Primary grouping (composite beat level) of Raag Suddha Danyasi transcribed from *sa-re-ga* notation

### Indication of alternating metrical grouping at beat level in Norwegian gangar tunes

Yet another example will be provided here. Some Norwegian *gangar* tunes have been perceived by transcribers, scholars and performers to have changing beat-grouping between measures at the melodic level even when the foot tapping is consistently periodic and symmetric, which have been characterized as creating a contrametric or metrically ambiguous experience. (see e.g. Kvitte 1983, Kvitte & Blom 1986, page 510, Levy1983:73ff).

Figure 5-53. From Levy 1983; Motifs from 'Norafjells' as played by Andres Rysstad.

The metrical interpretation of these tunes has been debated (Kvifte & Blom 1986, Levy 1983 et al) and we will discuss the analysis further in the evaluation section (see section 5.2.4.4 *Examples*). But the different points of view can be summarized as whether the grouping on the melodic level can be interpreted as a metrical or not.

For the purpose of demonstrating the method of analysis I will return to the example *Norafjells*, as played by Andres Rysstad, which was used as an example in the metrical grid analysis (see section 5.2.4.7)

Here the initial result of the beginning is exactly the same as the final result (see below). However, in later parts of the tune the sequences of measure grouping support the analysis of ungrouped measures.

**Norafjells**  
after Andres Rysstad, Setesdal, Norge

Segment start indications below system, grouping demonstrated by beaming trskr SA -92

The musical score for *Norafjells* is presented in five systems, each with a treble clef and a 6/8 time signature. The score is annotated with primary groupings and fingerings. The first system (measures 1-3) shows a 2-grouping (measures 1-2) and a 3-grouping (measures 2-3). The second system (measures 4-6) shows a 3-grouping (measures 4-6). The third system (measures 7-9) shows a 2-grouping (measures 7-8) and a 3-grouping (measures 8-9). The fourth system (measures 10-12) shows a 3-grouping (measures 10-12). The fifth system (measures 13-15) shows a 2-grouping (measures 13-14) and a 3-grouping (measures 14-15). Fingerings are indicated by numbers 1-5 and 0 (for natural). Some measures are marked as 'no grouping'.

Figure 5-54. Initial result of primary grouping of simplified version of *Norafjells* after Andres Rysstad (beginning of tune)

norany.t.1-46.mf



Figure 5-55. Final result. Output from analysis by computer model of metrical analysis of Norafjells after Andres Rysstad (original score input).

Outline of the MODUS implementation of metrical analysis

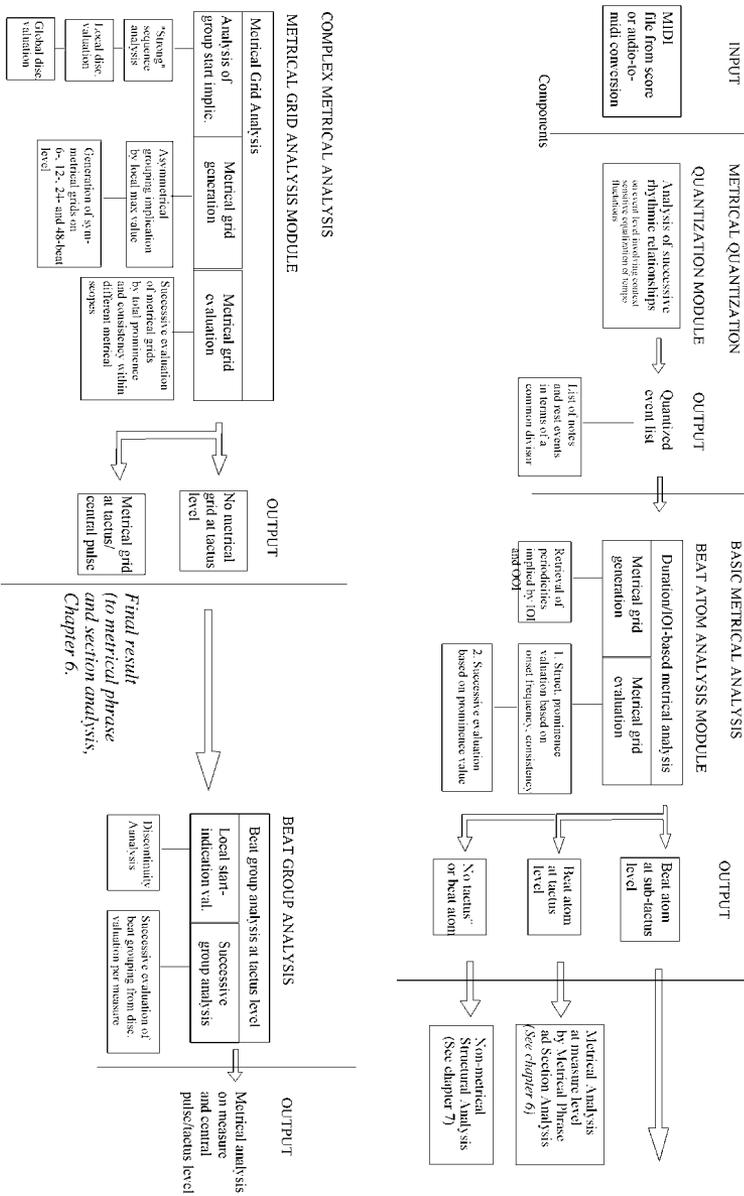


Figure 5-56. Outline of the MODUS implementation of the Metrical Analysis/Evaluation of metrical analysis

### 5.2.5 Means of evaluation

To recapitulate, the metrical analysis is divided into four steps:

- a) **Metrical quantization** at division level (see below)
- b) **Beat atom analysis** based on onset/offset information, gives a result within the range of possible beat levels (see section 2.1)
- c) **Metrical grid analysis**, at about measure and tactus/central pulse level. This part of the analysis only applies when the beat atom analysis do not provide a result at about tactus level. (see section 2.2.1-2.2.7)
- d) **Beat group analysis**, at about central/pulse level. This part of the analysis only applies when the metrical grid analysis do not provide a result at tactus level (see section 2.2.8)
- e) **Metrical analysis by phrase analysis** at measure level. This analysis is obtained as an output of the phrase analysis (see phrase analysis)

The evaluation of the metrical analysis is not an entirely trivial matter. The fundamental claim of the current model is that *a rule-based analysis based on the general assumptions of the nature of metrical perception and gestalt laws from information of information of event duration (IOIs) and relative pitch is sufficient to extract a metrical structure which would conform to peoples experience of metrical structure better than chance in monophonic melodies*. However, this immediately raises the question of how able people are to experience of metricity from such musically reduced information and most important to what extent people agree in their experience. The latter question is addressed above all in ethnomusicological studies, where the experience of meter in traditional sub-Saharan African music has been particularly in focus (see e.g. Nketia 1974). Another example with special significance for the current method is the interpretation of meter in Scandinavian fiddle music, which has been much debated especially in Norway since the late 19<sup>th</sup> century. (see e.g. Ramsten 1982, Blom & Kvifte 1986, Blom 2001, Groven 1982, Ahlbäck 1989)

There seem however, to be very few experimental studies of the variability of perception of metrical structure in melodies (exceptions are studies by Povel 1981, Deliège 1987, Parncutt 1994, Blom 1981, for a summary, see Clarke 1999). The focus of the existing research on meter seems rather have been on finding models that track meter correctly (see e.g. Temperley 2001:52 ff).

The current model, aims to analyze meter without a great deal of the information which is used when perceiving meter in melodies under normal musical circumstances (e.g. articulation, accompaniment, foot tapping etc.). Consequently the questions, as to which extent meter can be extracted from ‘piano roll’-information by people and to what extent they agree in their experience, become crucial.

Therefore, I have conducted a test of how metricity is perceived in some melodic examples, using as restricted input information for the listeners as is given to the model, see section 5.3.4.

The purpose of these tests is not to thoroughly test the validity of the method, but rather to get an indication of the plausibility of the method.

Another means of obtaining information of the plausibility of the method is to compare the analysis of the method with notation of meter in transcriptions in different musical styles and of different cultural origin. (see section 5.3.3)

The above discussion raises yet another question: The method was originally developed to handle 'score input' rather than 'performance input', i.e. quantized input where tempo fluctuations and variation in timing have been omitted. There is rather strong psychological evidence that such a quantization is performed by people while perceiving music (Fraisse 1956, Povel 1981, Lee 1991, Parncutt 1994 et al). The problem of a model that presupposes quantized input is, however, to address the problem of how quantization affects the rest of model of metrical conception.

After testing a few methods of initial quantization I decided make my own attempt at such a method of initial quantization including the adaptation of the rest of the model to account for this. As the purpose is mainly to demonstrate how the model relates to un-timed input, this method is not included in the specific description of the metrical analysis and is not thoroughly tested. The results so far are, however, promising. (see section 5.3.2 below)

## **5.2.6 Evaluation of Metrical Quantization Analysis**

My preliminary attempt to make a tempo-equalized quantization based on the assumption of a common time denominator at division level seems to be quite successful. In the test material, which involves only 20 melodies from the realm of Swedish instrumental folk music performed in traditional style, the ratio of correctly (in relation to existing transcriptions) traced division unit relationships is 0.94.

As have been discussed above, the impact of the quantization process demands greater levels of freedom to be placed on the subsequent metrical analysis. This has demanded some adaptation of the previous version of the method of analysis to account for this. However, the order between the different steps of the analysis is confirmed by the trial, where the quantization is regarded chiefly as a primary process, which provides input to the periodicity analysis. However, the trial has also shown that under some aspects of quantization cannot be regarded without the involvement of the latter parts of the analysis, which confirms the general model of melody conception and metrical conception but demonstrates the limitation of the current implementation, where these levels of analysis are performed in separate steps and not in parallel. The purpose of the implementation here is mainly to demonstrate the validity of the model.

In this respect, the preliminary quantization model has confirmed the assumption of the primacy of event onset and offset as the main source of information at the lowest level of perception of periodicity; the pre-metrical level of successive relationships between duration of events.

## **5.2.7 Beat atom analysis - Basic analysis of pulse levels by onset information**

### **5.2.7.1 General discussion**

This part of the analysis is quite sufficient with regard to many musical examples in providing the necessary prerequisites for the subsequent phrase and section analysis, namely an analysis of central pulse level (tactus level). Previous methods of metrical analysis are predominantly based on onset information (for an overview see e.g. Clarke 1999). The relative

success of these methods regarding common-practice Western popular and Classical music (Temperley 2001), suggests that onset information has special significance for the perception and cognition of metrical structure. The reasons for this suggestion was discussed in section 5.1.3.

There are however obvious limitations to onset information as signifier of metrical structure. Strange as it might seem, the opposite qualities of a musical structure to cause problems at the level of quantization analysis, makes the beat atom analysis easier: The more complex divisions of a given beat level, the more unambiguous the beat level will be in terms of onset/offset periodicity. Conversely, the fewer the number of complex divisions, the more ambiguous the periodicity at higher levels will become since you cannot use onset periodicity to track it. Imagine a **moto perpetuo** consisting only of sixteenth notes at a tempo at the quaver of 140 MM. The central pulse level will most likely be perceived as longer than a 16<sup>th</sup> note, but without any other categories of duration, this would be impossible to guess without analysis of pitch change.

The analysis of more complex metrical levels involving periodicity in pitch change will be discussed in the next section. This section regards only the basic beat analysis, named beat atom analysis, based on onset periodicity. This can either give pulse at central pulse level or metricity at division level as output. If the test file is clearly non-metrical it will not give any result. As have been described above (see section 5.2.3.2), the output typically includes more than one possible set of solutions, a first and a second guess for different starting points of the grid, each of which can contain a list of hierarchically related possible beat durations. They are ranked based on tempo, division, continuity, support and negative support with the highest ranked metrical division being the main return.

I have examined the performance of the method for different material of different musical characteristics. A simple rating of the success of the method says little if is not related to the rhythmical and metrical characteristics of the style/example.

For instance, the method was tested on 75 tunes in the central collection of Swedish folk tunes “Svenska Låtar” (sections Dalarna and Jämtland) with only one tune being interpreted incorrectly, assuming the original interpretation to be correct. (see further Chapter 6, section 6.3, *Evaluation of Metrical Phrase and Section Analysis*) This is, however a relatively easy task for the beat atom analysis due to the diversity of divisions of central beat levels caused by the relative precision in rhythmic notation and the reduction of ties between notes obstructing the beat.

The difficulties regarding the method are inversely related to the level of onset periodicity and can be summarized as follows:

- 1) Pre-onset of beats
- 2) Beat marking by other means than onset, such as articulation by dynamic or timbre change within IOI:s or beat markings by pitch change
- 3) Syncopation
- 4) Polymetricity by phase displacement of pulse (obscuring periodical onset patterns)
- 5) Polymetricity by uneven superimposition of pulse (obscuring periodical onset patterns)
- 6) Quasiperiodical beat patterns, such as periodical shift in pulse period.

The first, second, third and sixth difficulties are common to Swedish folk music, when using input from quantized live performance as exemplified in the previous section. Therefore, I have used mainly that kind of material to evaluate the performance of the

method regarding these aspects. Quasi-periodical beat patterns are extremely common in southeastern European folk music traditions as well as traditional and art music from the Middle East and India. But the problems of finding a central pulse level in these kinds of melodic structures, is mainly a matter of the subsequent complex metrical analysis and will be discussed more thoroughly in the next section. Poly-metricity by phase displacement of pulse is well represented in Norwegian instrumental folk music, and consequently I have used some Norwegian examples to examine this problem. Syncopation is also frequent in early 20<sup>th</sup> century Western popular music, and I have used some examples from such musical styles to study this phenomenon.

### 5.2.7.2 The performance of the beat analysis regarding obscured beat onsets by pre-onsets and syncopation in Swedish folk music transcribed from live performances

The material used to test the performance of the method in this respect, has been restricted to 20 melodies notated by audio-to-midi conversion and quantized through the method described in section 5.2.2 .



Figure 5-57. *Quantized input file*



Figure 5-58. Metrical beat atom analysis. Beat atoms designated by books at top of system. Confirmed as central pulse level.

The placement of the beats at central pulse level coincides with the dominant result from a test with 36 subjects, who had never heard the tune before, (The melody is composed in a style typical of traditional Swedish folk music.) The task was to put in bar-lines, beams and rhythmic relationships in a graphic image consisting only of the relative pitches. This involved the problem of dividing notes into two tied notes to obtain this solution. 64% of the subjects provided exactly this metrical solution, while most of the remaining arrived at a solution, which strongly confirmed to the above. The character of the differences between the dominant solution and the results of this group suggests that this was mainly due to notation problems. A minor number of participants provided very different solutions.

It is important to note that most of the subjects were expected to have at least slight familiarity with the intended musical style of the example.

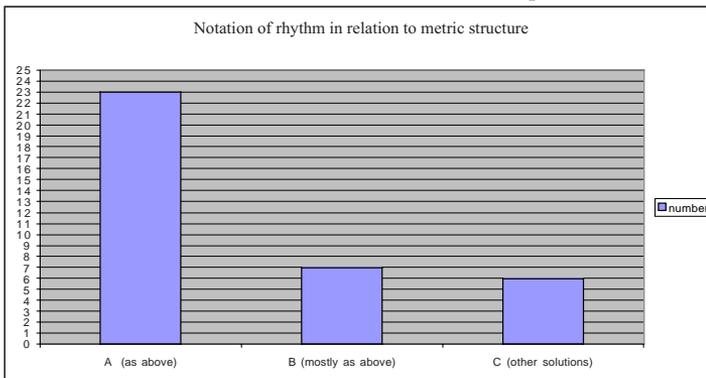


Table 7. Notation of rhythm in relation to metric structure

The problems of obtaining the solution from the onset analysis is mainly due to the obscured onsets at beat positions. However, of the different possible solutions, this has the greatest number of onsets at beat positions in the relation to the possible beat positions at a preferred tempo and relationship between division of beats and events of longer duration than beats.

This supports the assumption that periodicity regarding onset information is of special significance for metrical analysis at lowest levels, from division up to about central pulse levels.

### 5.2.7.3 Syncopation/obscured beat-atom in 20<sup>th</sup> century popular music – “Fascinatin’ rhythm” by G. Gershwin

I will provide yet another musical example, which illustrates the same kind of difficulties due to performance practice as the example above. Below is a quantized transcription of the refrain song line of “*Fascinatin’ rhythm*” by G. Gershwin, as sung by Ella Fitzgerald (Verve records 1962). In this case, also the beats are regularly obscured by pre-onsets and syncopation. The most recurrently marked period by onset at central pulse level is however the beat level marked in the output as dotted eighth notes. The tempo factor relating to preferred pulse rate also the level of rhythmic division of the beat and the number of beats within notes of longer duration also influences the result, while this beat size fits the category requirements of a central pulse level (tactus level).



Figure 5-59. Excerpt from the song line of ‘Fascinatin’ Rhythm’ by George Gershwin (as sung by Ella Fitzgerald)

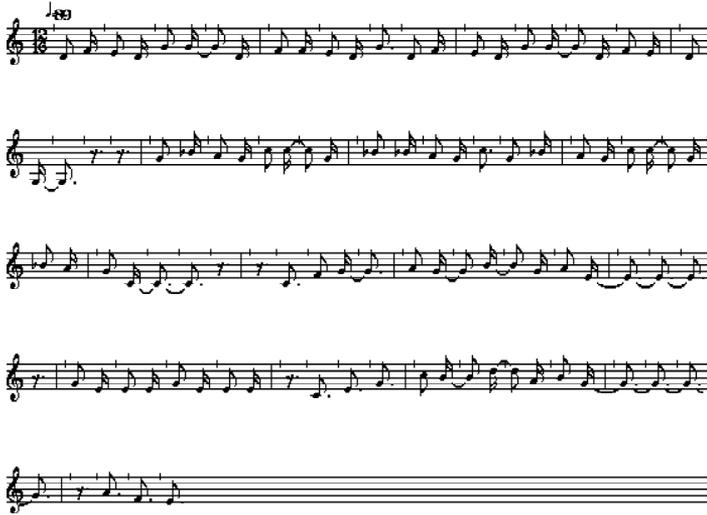


Figure 5-60. 'Fascinating rhythm'. Output of analysis by computer model.

#### 5.2.7.4 Irregular syncopation relating to a metrical organization provided by accompaniment – Bolero by M. Ravel

When a melody is conceived in relation to an accompaniment, the need to provide the information necessary to conceive the central pulse level becomes less obvious. There are many examples in which the method of beat atom analysis would not be able to find the central pulse level. This would be true for a great portion of the 19<sup>th</sup> century Western opera repertoire. In some cases, the melody, however, provides the necessary information, even if the obscuring of the beats are frequent and irregular.

In the example below, the notated central pulse level (the quarter note level) turns out to be the only metrical solution which survives throughout the melody. This requires, however, that the meter in question is very well established, e.g. more than 23 successive onsets occur at beat positions and more than fifty percent beat positions are at onsets. If this applies it herein assumed that we can accept temporary deviations from the pulse within the categorical and temporal limitations of the perceptual present.



Figure 5-61. 'Bolero' by M. Ravel (composed 1928). Output from computer model analysis.

### 5.2.7.5 Phase displaced pulse – Example from Norwegian 'Halling' music

The regular syncopation of the above example (*Fascinating rhythm*) comes close to the phenomena of complex/superimposed meter by phase displacement: If there is generally a certain pre-onset of beats at a level which is above or at division level, this can be conceived as the cooperation between two meters of the same tempo, albeit phase displaced. (see section 5.1.3)



Figure 5-62. Illustration of conception of phase displaced super-imposition of pulses.

In Western popular music styles from the 20<sup>th</sup> century, such as e.g. reggae, the so called 'back-beat' feeling which can be designated as perception/conception of phase displacement of pulse is regarded as an important stylistic feature.

The same can be said of some Norwegian *halling* and *gangar* tunes (which also do occur within traditions in western parts of Sweden), for which the term '*mottak*' (counter-time/beat)

has been used by performers to designate the metrical feeling of some tunes. (Kvifte 1983, Blom & Kvifte 1986).

Contra-metricity is present already in one of earlier notations of a ‘balling’, originating from a personal collection from the late 18<sup>th</sup> century. In this case, the tune is arranged for a keyboard instrument with the central pulse level indicated by the accompaniment, while the melody is devoted to the contra-meter, the off-beats.

## Halling from Inger Aalls collection

From Hroar Dege: C Hammer, Norsk kokebok 1793



Figure 5-63. *Halling from Inger Aalls collection* (Dege 1990)

If the “melody line” alone is analyzed according to the current model, the rest starts, i.e. offsets or OO!s, provide the essential information of the metrical structure:



Figure 5-64. *Computer model output from analysis of Halling from Inger Aalls collection.*

Chapter 5

This example shows why it is essential to the model not only to consider interonsets but also offset-onsets, since periodicity in offsets can be perceived as well as periodicity in onset. In this particular case, the onset solution starting on the first note is almost as likely as the offset solution. It is the negative preference for short-long starts (16<sup>th</sup> notes followed by eighth note rest) that favors the rest-start solution

When neither onsets, nor offsets are found at positions that can be perceived as the central pulse, how can the pulse be traced?

The example below is a constructed melody inspired by ‘balling’ tunes, showing this particular situation. (In real musical examples, ornamentation and foot tapping often indicate the obscured beats at this level, while the articulation often enhances the obscuring of the central pulse level).

The first event considered by the analysis is the rest at the beginning of the first measure. While the beat atom length of a quarter note can apply to the rest in a rhythmic relationship, it is tried as beat atom at the position. Since there is no event start at next possible beat position all events (durations of IOIs and OOI:s) up until next event at beat position are collected and checked for their relationship to the tested beat-length. If, when the phase changing notes in the beginning and at the end are stripped, the remaining events are of the same length as the tested beat length, then the segment is considered to be contra-metrical in a perceivable relationship.



Figure 5-65. ‘Halling’-inspired constructed melody. Input to computer model.

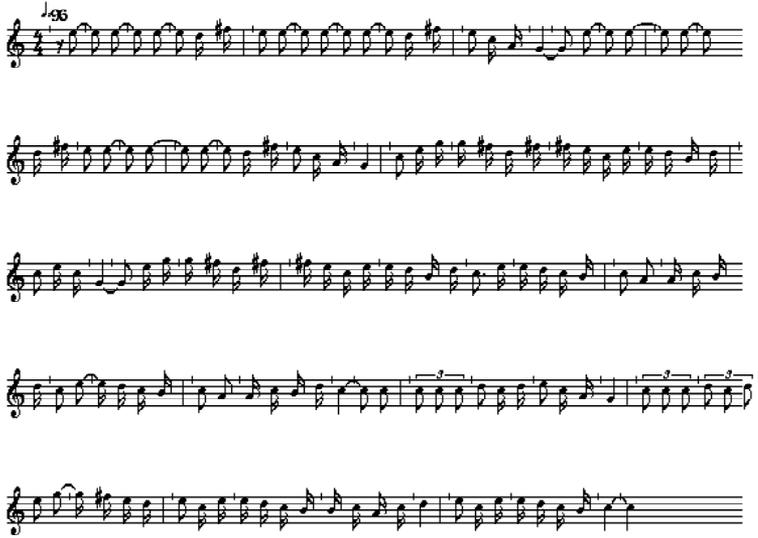


Figure 5-66. Computer model solution for 'Halling'-inspired melody.

The solution, which would start on the first note, is devaluated here, the metrical grid is not by onsets and divisions in the latter part of the melody. If the ornamentation is omitted from the transcriptions, however, there are a number of traditional *halling* and *gangar* melodies, which would lead to a solution, favoring the contra-metrical pulse as central pulse level.



Figure 5-67. Input representation of the beginning of "Skjallmöyslaje after Olav Hegland" (Lande 1983:208). Ornamentation and accompanying notes omitted.

Above (in fig. 67) is an example of the beginning of the *gangar* “*Skjallmöslajé after Olav Hegland*” (Lande 1983:208); the ornamentation and the double stops/accompanying open strings that appear in the original transcription are omitted. (Observe that the time signature changes do not have any significance on this level, the analysis concerns pulse. The input files can have any notation of pulse and time)

Only considering this information the analysis would produce an output that is not in concordance with the original transcription, since the majority of the onsets of the melody indicates the contra-meter as the central pulse:



Figure 5-68. Output of computer model metrical analysis of the beginning of “*Skjallmöslajé after Olav Hegland*”.

If a greater part of the melody, which has dominant support of the originally notated central pulse level, is given as input to the computer model, the dominant solution of the beginning of the melody would be suppressed, and the originally notated pulse will appear in the solution:

The vast majority of *balling* tunes of the traditional *balling* and *gangar* tunes in duple metrical division contained in the “The fiddle tunes volumes” (Norwegian Collection of Folk Music at the University of Oslo) do not exhibit contra-metrical evidence to the extent that it appears in the above example. I have found, however, seven tunes out of a total number of 50 tested tunes in duple metrical division, in which the output of the analysis do not favor the originally notated beat level when ornamentation and accompanying notes are omitted.

This shows clearly the limitations of the method: It indicates that in this particular repertoire, the crucial information for determining the central pulse level is not always present at the level of the brief transcription, i.e. when ornamentation, articulation and accompanying notes are omitted, not to mention the foot tapping. Instead the melody in these relatively rare examples seems to be fully dedicated to the contra-metrical evidence, which requires e.g. the foot tapping to be interpreted as intended. In most cases, the contra-metricity at the level of melody is rather restricted (i.e. when articulation originating from bowing patterns are not taken into consideration). In the example below, contra-metricity appears in the middle section, when the central pulse level is already established:



Figure 5-69. Output from beat-atom analysis of "Skjallmöyslaje after Olav Hegland".



Figure 5-70. Halling after Ole Evju, Eggedal (From "Norwegian Folk music, series I", no 160k (Gurvin 1959), output from computer analysis.

### 5.2.7.6 Uneven metrical complexity

Phase displacement of pulse is one of the fundamental types of *polymetricity* or metrical complexity. Another is *uneven super-imposition of pulse*, i.e. that two periodicities of uneven relationship (3:2, 2:3, 4:3 etc) are conceived to interact in the metrical experience (see section 5.1.3). Even if there is music where the concept of a central pulse level becomes problematic (see above, also e.g. music from sub-Saharan Africa) the most common situation in Western music is that one pulse is perceived to be the central pulse level with a contra-metrical pulse obscuring it. The division between syncopation (i.e. rhythmic figures obstructing pulse) and the perception of super-imposed pulses is conceptually distinct. The latter implies that periodicity is conceived at a contra-metrical level. How to discriminate syncopation and super-imposed pulse in structural terms, is, however, a greater problem; this is because the ability and tendency to relate periodically to events in time is different for different individuals (see section 5.3.4.3 *Discussion*).

In this model, however, I have assumed that when event onsets/offsets at rate relationship 4:3 to the period/beat-length in question, obscure more than one successive beat position, the obscuring events must conform to the limitations of either a phase displaced or uneven relationship.

In the example below, there are three different periodical levels of obscuring the central pulse: In the first measure, just one onset is obscured. In the third measure, two onsets are obscured. In the ninth to eleventh measures, two succeeding onsets at a time are obscured, but here the obscuring event onset periodicity is continued throughout three measures. In the first case, it can hardly be experienced as an obscuring periodicity. In the third case, it does not require a specifically developed sense of periodicity to perceive the repeated contra-metricity. In the second case, it would probably be possible to experience it as either a rhythmic (gestalt) syncopation – polyrhythmic - or as implied contra-metricity – polymetric.

In this case, it is analogously to the phase displacement of pulse (see above) a matter of frequency of occurrence and consistence, which of the two different implied periodicities will be regarded as central pulse/tactus.



Figure 5-71. Input representation of constructed test-melody (Polska style) for *polymetricity vs. syncopation*.



Figure 5-72. Output from computer model metrical analysis of test-melody for poly-metricity.

The limitations of the model are obvious in this respect. It does not accept higher levels of *polymetric* relationships than 4:3, due to the fact that I have not yet encountered more complex relationships within different kinds of traditional ethnic music. I have not managed to find any experimental evidence that more complex relationships between periodicities at pulse level are actually perceived by listeners in e.g. 20<sup>th</sup> century Western Art music where it is frequently used in notation.

## 5.2.8 Complex metrical analysis – metrical grid analysis.

### 5.2.8.1 General description

As have been described above, this part of the metrical analysis is required only when the beat atom analysis does not come up with a beat analysis at central pulse level, but a meter at division level. This means that it either is too fast to fit within the range of preferred central pulse rate or it does not conform to the category requirements of a pulse at central pulse level, i.e. do not have enough divisions of the meter or that it divides longer events into too many metrical units. The purpose of the complex metrical analysis is to find a meter at central pulse level, since this level is regarded the primary level of grouping on which the phrase and section analysis is based.

This typically concerns melodies dominated by events of equal length and melodies with asymmetrical grouping at central pulse level.

The complex metrical analysis, therefore, involves the analysis of pitch information and repeated patterns (sequences) in addition to the event duration (IOI, OOI), which is considered in the beat atom analysis. This reflects the assumption that pitch patterns and sequencing (parallelism) have greater impact at higher levels of metrical organization. (see section 5.2.4.3 *Assumption no 32-34*)

With regard to this level of analysis, it is essential to evaluate the results of the analyses in relation to listener's perception and cognition of meter and not only to notations. This is because organization by pitch patterns and sequences are 'weaker', more ambiguous indicators of grouping than onset and offset periodicity. This is indicated by experiments performed by Deliège and others (Deliège 1987, Bertrand 1994:E-6). Other factors, such as phenomenal accentuation by change of timbre, change of intensity etc. will have a stronger influence on the perceived meter.

I have, therefore, performed a set of tests to find out to what extent people conform in their perception of metrical grouping when given the same information as is used in the computer model analysis.

In addition to the comparison between test results and the performance of the analysis, I have used notations of metrical grouping for evaluating the performance of the analytical method and, thus the model.

### 5.2.8.2 Experiment 1. Sensitivity to grouping by pitch sequencing (1996-97)

The analysis of sequences constitutes a fundamental part of the complex metrical analysis. It is assumed that immediately repeated patterns, i.e. sequences, represent a special case of parallelism, which can overrule for example grouping by discontinuity and grouping by symmetry (see section 5.2.4.3).

The object of the test was twofold: first to investigate the impact of pitch sequences on listeners perception of grouping; and, second in for listeners who chose a sequence related interpretation to investigate their preference for phase of sequence.



symmetrical or to sequence related grouping. However, there was a general reaction to the task as being very difficult. Taking into account that the pre-test subjects were all music students and musicians with good knowledge of notation, I changed the design of the test for the real test. This time I chose two of the more frequent responses provided by the pre-test and another more unlikely solution according to the pre-test results.

Therefore, in the real test I notated the three alternative groupings by different beaming of the notes as can be seen in the notation in fig. 73. The design of the test evidently made it impossible for subjects who could not read music to participate. The participants were music students at graduate level (30), free-lance professional musicians (6) and amateur musicians (3), but with background in different kinds of musical styles. Fifteen of the participants had their background in Swedish folk music while the rest of the participants were mainly acquainted with popular music, Western classical music and jazz. None of the participants had their background mainly in music in which asymmetrical grouping is dominant, although all of them would have heard such music.

The task was to choose the beaming alternative, which would correspond to their own perception of how the notes were grouped. They had to choose one of the alternatives, but were allowed to suggest changes to the grouping in the preferred alternative. In three responses these changes were quite substantial, but were generally through the responses very few. Based on the pre-test indications, my hypothesis was that alternatives 2 and 3 would be chosen with almost equal frequency. The pre-test had emphasized the preference for symmetrical grouping. Some subjects spontaneously uttered that the mechanical performance of the computer *“makes everything feel like 4/4 time”*. Therefore I also let the subjects listen to the sequence four times with breaks between (20”-1”), to be able to concentrate on the content of the test melody for each interpretation and to confirm a final choice. The results can be read from the table below:

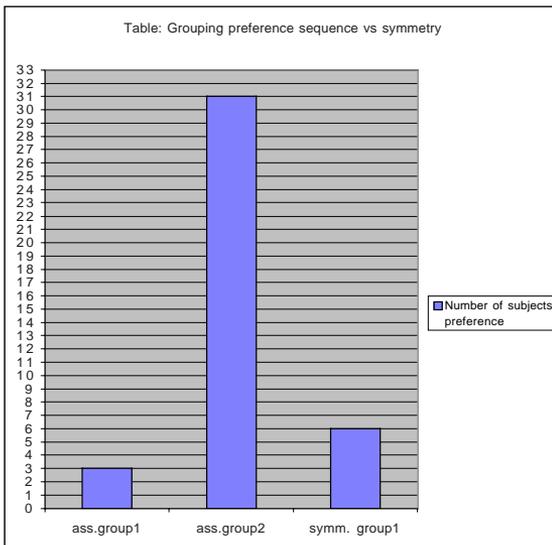


Table 8: Grouping preference sequence vs symmetry in experiment 1a

This result demonstrated higher sensitivity to the asymmetric grouping of sequence than I had anticipated from the pre-test results. This deviation from the expected result can perhaps be explained in that it was easier to draw such a conclusion when the melody was already interpreted.

Still a frequent comment on this test as well as most of the tests using equally articulated MIDI files as sound sources is that it is possible to hear in a number of possible ways.

There were no correlation tests made regarding age or musical background since the results were so clear.

This result indicates that even with a music cultural background in which symmetrical beat grouping is most common (Swedish general musical schooling) people can prefer an asymmetrical grouping when it is perceived to be consistent with the melodic structure. It also indicates that people schooled in a musical environment in which melodies are recognized are able to perceive melodic grouping depending on *acoustically/phenomenally* relatively weak structural factors as pitch sequences, change of melodic direction and relatively moderate change of step-size. Furthermore, this result indicates that the assumption of prominence of the first discrete sequence start among possible different sequence starts is correct.<sup>55</sup>

Accordingly, this result can be interpreted as an indication of the perceptual reality of the general rules of ‘strong contour sequences’ described in section 5.2.4.5. However, each independent rule was not tested on a sufficient number of people to provide perceptual evidence for every single rule; the number of participants in these tests was too limited (see further experiment 1b).

In this particular case, the rules of the strong contour sequences as a part of the complex metrical analysis can account for the dominant interpretation almost entirely. (See example below)<sup>56</sup>

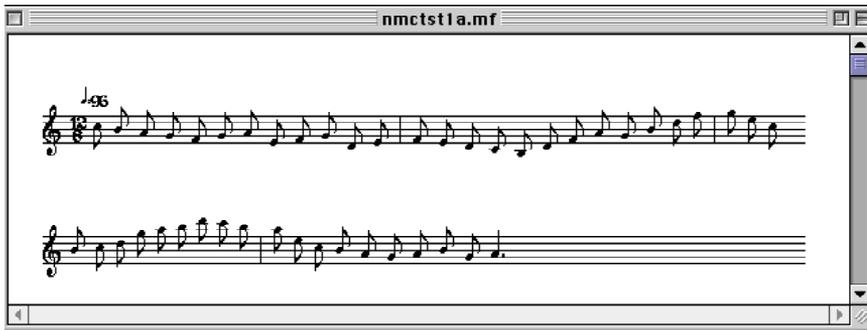


Figure 5-74. Input file of nmctst1a

<sup>55</sup> This is in accordance with the findings of Deliège and others, regarding the prominence of the first presented structural cue. (see e.g. Bertrand 1994)

<sup>56</sup> The extra group marker at the last bar line is due to a programming error

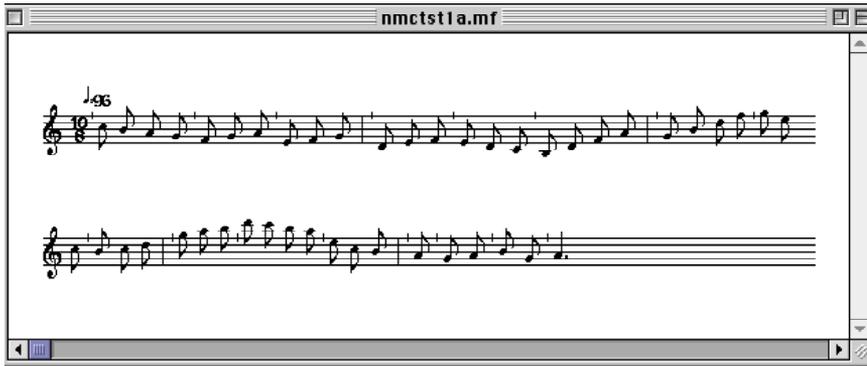


Figure 5-75. Output of complex metrical analysis of *nmctst1a*. 'Hooks' indicate group boundaries

The metrical segmentation is here almost exclusively dependent upon contour sequence analysis. The result of the metrical analysis at measure level would probably not be confirmed by test results. This is, however, of minor importance for the subsequent analysis since the metrical analysis at measure level is not confirmed until the phrase and section analysis is performed. It is important at this point only to the extent that it influences the analysis of primary metrical grouping.

A legitimate question to ask in this case is if this can be considered metrical grouping? Can the groups be referred to as 'beats'? According to my definition, most of them can, since their beginning creates accents in time at certain intervals, which are recurrent and hence at least quasi-periodical during parts of the melody (see section 5.2.4.2). This is evidently not true for all people (according to the test results). Whether or not people would perceive it as regular in any sense would probably be a statistical question. My conclusion here is that where the grouping is perceived as periodical, the same rules of grouping by pitch sequencing applies.

### 5.2.8.3 Experiment 2. Sensitivity to grouping by pitch discontinuity and pitch sequencing at metrical levels

#### Methodological issues regarding quantification of perception of metrical grouping

The main problem with the above experiment – as well as with some other experiments that I have performed to test different grouping preference – was that it required skill in reading music to be able to respond to the question.<sup>57</sup> This involves several problems. One is that it restricts the number of participants possible in a test to those with skills in reading music. One can also question in what way their different skills in reading music influences the result. Yet another problem connected is the influence of the visual image of the notation on

<sup>57</sup> Other methodologies have been successfully used by e.g. Deliège (1987) and Krumhansl (1996) in segmentation tasks

the choice of grouping – the level of familiarity with a certain written structure would perhaps influence how one would group the notes?

Therefore I conducted a set of experiments entirely based on listening to music with written responses not depending on common notation. At first, I tried to have participants tap along with the music. This was, however, difficult as soon as structures became more complex. It turned out to be very hard not to tap on the basic metrical level. To tap just at measures and not at the intermediate beats, to show more than one metrical level at a time or just to tap on phrase boundaries proved difficult for subjects. There were also problems associated with conventions of foot tapping in that responses could be difficult to interpret: In some examples of Swedish folk music, the convention is to tap just on the first and the third beat – which is commonly also regarded as the *first* and *third* beat, not as the first and second beat. Could such a behavior be regarded as corresponding to tapping on each beat? Similar problems occurred in e.g. jazz-related musical examples.

My approach in attempting to manage these problems has been to make choices of different metrical interpretations. I have used percussion accompaniment, accentuation and separation of groups through phrasing, by means of rests between group boundaries and tempo changes to and from group boundaries.

The theoretical basis for the use of these means of indicating metrical structure can be traced back to the fundamental assumptions of the nature of metrical structure: (i) that beats represent prominent periodical points in time, which can be interpreted as accents on a cognitive level, they may be represented by accents also on a phenomenal level; (ii) that groups are constituted by discontinuity on a cognitive level, discontinuity on a phenomenal level may be an evident way of indicating group boundaries. But even more importantly accents made by e.g. foot tapping, drum accompaniment, click track etc., are used by people in connection with metrical experience. As for indicating group boundaries by inserting rests between groups and varying the tempo within groups, these are well documented means of musical performance to indicate melodic structure (Bengtsson & Gabrielsson 1980, Friberg 1995).

In this case in order to retrieve a number of plausible choices, I also made a pre-test. For which skill in reading music was required. In this pre-test, I used MIDI recordings with equalized articulation and without tempo or timbre changes. Subjects were required to notate pulse, division and higher-level meters and phrase structure on a response sheet with equally spaced notes without any indication of rhythm or meter.

## **Purpose**

The fundamental object of this experiment was to investigate if metrical structure could be extracted from the limited information used in the current method of analysis described above; Further, to examine whether the participants could provide concordant responses or not. If the results did not show significant concordance, it could be interpreted as evidence that the stimulus material did not contain sufficient information for metrical interpretation. But it could also be interpreted as if people strongly differ in their metrical perception under the circumstances of the experiment – e.g. when not being involved in any social activity demanding a common metrical interpretation.

Thirdly, I wanted to investigate to what extent culturally learned patterns influenced the result, by using melodies with different cultural origin as well as using people with different musical background as participants.

Further, I wanted to test the strength of pitch repetition and pitch discontinuities as a cue to metrical experience. This was obtained by using test melodies, which did not generally contain different interonset intervals, but mostly notes of the same duration.

## **Participants**

The participants were all students at Royal College of Music in Stockholm. All of them, therefore, had some formal musical training. Their background in this respect, however, varied considerably in length. While the average of formal musical training was 8-12 years, some of the subjects had considerably lower degree ranging between 1 and 3 years. This was because some of the students were picked from a class for musicians with their profile within non-Scandinavian folk music and non-western art music. These participants (5 out of a total of 20) had different cultural background: Senegal, Chile, Iran, Turkey and India and have moved to Sweden during the last ten years. They all have a strong musical identity in the traditional musics of their native countries.

Most of the participants had a general Swedish musical background, which can be classified as generally Western. In the Swedish society of today, it is virtually impossible to avoid being exposed to music, which is presented in television, radio, films, compulsory school etc. and can be said to represent common-practice Western music of today. For this reason I have disregarded the differences between the subjects concerning musical stylistic direction. It must be noted, however, that the participants were all students of the folk music program at the Royal College of Music in Stockholm.

Two of the participants were brought up in other European countries, having a musical background which can be characterized as generally Western.

The participants were payed a small sum in compensation for the participation in the test.

## **Apparatus**

The sound examples were generated by a Roland JV-1010 sound module and recorded as mp3 files on an Apple computer (44.1 Khz, 16 bits). The MIDI performance of examples 1-7 and 10 was controlled by the notation program Finale (versions 2001 and 2002). The examples 8 and 9 were recorded by me on a MIDI keyboard in a sequencer program (Logic). The patches used were Grand Piano for the melody part and General Midi Percussion patch for the percussion parts. (For the individual percussion timbres see notation of examples below). The choices of sounds were meant to be stylistically neutral though still not hard to regard as musically intended. The reason for using piano sounds for melody parts was also that an articulation of each tone, the lack of slurs, would sound natural.

## **Procedure**

The participants were given a form (see example below and appendix for the full form) to fill in their preferences of metrical interpretation for each example:

Lyssningstest - Grupper (kopier)

## Lyssningstest

GÅ IN PÅ FOLKMUSIKINSTITUTIONENS HEMSIDA [WWW.KMH.SE/SF](http://WWW.KMH.SE/SF).  
 KLICKA PÅ LÄNKEN TILL LYSSNINGSTEST SVEN.

### Uppgift 1

**1a**

Lyssna först till denna melodi spelad på synthesizer utan ackompanjemang

**Ex 01 01.** Klicka på dokumentnamnet, så öppnas det i programmet **quicktime**. Tryck på **'play'-ikon**en.

Lyssna igenom den en gång. Stäng den sedan.

**1b**

Öppna både EX 01 02 och EX 01 03 genom att klicka på dem en i taget.

Här är uppgiften att avgöra vilken av dessa båda versioner som överensstämmer bäst med melodin? Kryssa för det alternativ som passar för respektive exempel. Du kan lyssna mer än en gång om du behöver. Stäng den sedan.

**ex 01 02**

stämmer inte alls   
  stämmer illa   
  ...stämmer med bra   
  stämmer bra   
  stämmer nästan helt

**ex 01 03**

stämmer inte alls   
  stämmer illa   
  stämmer   
  stämmer bra   
  stämmer nästan helt

**1c**

Nu följer fyra exempel där melodin följs av en klang. **ex 01 04- ex 01 07**. Vilken av dessa klanger låter mest avslutande? Rangordna dem från 1 till 4, där 1 är den som känns mest avslutande. Skriv siffrorna i den högra kolumnen.

ex 01 04	
ex 01 05	

Table 9. The participants form, in Swedish, (see appendix for the full form).

Thus, the participants of the test had no notation of the melodies in the experiment to look at.

The procedure of the test was generally as follows:

1. The subject was instructed to listen to a melodic excerpt. This excerpt had equalized articulation and tempo with piano timbre. The subject was to make note if he/she recognized or knew the melody.
2. The subject was then instructed to listen to between two to four different versions of the same melodic excerpt with either different percussion accompaniment, melodic articulation or melodic phrasing (ex. 8 and 9) and was instructed to rate how well the different versions fitted with the melody

*(överensstämmer bäst med/passar ihop med'* were the expressions used in Swedish) on a 5-grade scale from very bad to very good. No further instructions or explanation of about the task were provided.

The test was performed in the computer rooms at the college, the participants listening in earphones. The participants performed the test in their individual speed (for an average of 2 hours, varying between approximately 1 and 3 hours).

They were allowed to listen repeatedly to the sound examples of each task, but not to go back to a previous task and compare the example of the different tasks. Thus, the number of times a participant listened to the examples was not controlled in the experiments and can be regarded as a source of variance in the results. However, the reason for this choice of procedure is that there is strong reason to believe that people vary in how quickly they develop a structural conception of a piece of music, i.e. that the conception of a melody is a learning process. If this is true, the impact of conception time factor is likewise an unknown source of variance in a test with a fixed number of stimuli presentations.

In this case, subjects were to continue on to the next task when they had formed an opinion about their preference of the current task. The benefit of this procedure is that the subjects' own decisions of their conception forms a basis of the testdata.

## **Stimulus material**

### ***Trial 1 and 2 – Same pitch content different order***

The stimulus material consisted of melodies and melodic excerpts which were either composed by the author or were traditional melodies of different cultural origin.

Test examples 1 and 2 (see below) were composed by the author and consisted of the same set of notes put in different order in order to invoke different conception of metrical grouping.

## Test ex 01

Melody composed to indicate triple(3, 6, 12) grouping

Test melody

Metrical interpretation: Alternative 1

Metrical interpretation: Alternative 2

Figure 5-76. Test example 1 (tempo: ♩ = 104)

## Test ex 02

Melody constructed to indicate duple/quadruple grouping

Test melody

Metrical interpretation: Alternative 1

Metrical interpretation: Alternative 2

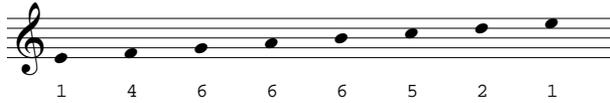
Figure 5-77. Test example 02

In spite of the naming (test ex 01 and 02) these were not presented in immediate succession in the test, but instead as the first and seventh of the examples.

Chapter 5

The occurrences of notes in these two melodies were identical so as to test the influence of ordering – melodic contour – as a cue to metrical grouping.

Number of occurrences of notes in ex 01 and 0

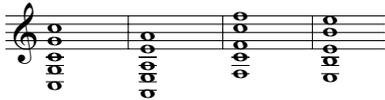


Here I used the same metrical indication as the default setting in Finale as accompanying ‘click’ for MIDI recording. This setting is possibly chosen by the developers of Finale from experience of conventional drum accompaniment. The lower pitched, longer sounding sounds (snare drum) indicate downbeats at about measure level while the shorter higher pitched sounds (side stick) indicate central pulse (tactus) level. I have chosen this way of indicating metrical interpretation in this example, since it is in accordance with the general assumption of metrical accentuation (see section 5.1.4, *Assumption no 22-23*), which states that grave accents have greater metrical prominence than acute accents.

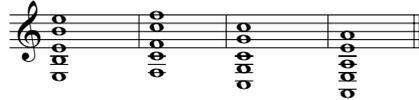
In this case, a subsequent test, generally designed according to the probe tone method (Krumhansl & Shepard 1979, Krumhansl 1990), was also performed. In this instance subjects were to rank four different chords following 2’ after the sequence in regards which of the chords sounded most finite with regards to the melody. The chords consisted of a fifth and a fourth in spaced over two octaves. The reason for adding a fifth and not use just octave-spaced single notes was to enhance the sense of tonic and suppress the possibility of perceiving melodic finality.

The chords and order of presentation were:

Test ex 01



Test ex 02



The chords were chosen from the most frequent choices of perceived tonic of the pre-test. These sequences were also tested in an earlier test regarding perceived finality (Ahlbäck 1996). This task was added to the test in order to test the extent to which perception of metrical grouping related to the perception of tonality.

**Trial 3 – indicated asymmetrical grouping****Test ex 03**

Test melody composed to indicate asymmetrical grouping

Melody example

Alternative accentuation 1

Alternative accentuation 2

Alternative accentuation 3

Alternative accentuation 4

The figure displays five staves of musical notation in treble clef. The first staff is labeled 'Melody example' and shows a sequence of notes with varying durations and accents. The following four staves are labeled 'Alternative accentuation 1' through 'Alternative accentuation 4'. Each of these four staves shows the same sequence of notes as the first staff, but with different accents (indicated by a greater-than sign > above the note) placed on different notes to illustrate alternative groupings. The notes are primarily eighth and sixteenth notes, with some quarter notes.

Figure 5-78. Test example 3 (tempo: ♩ = 120)

In test ex 03 I used acute accentuation through increased sound level (increased MIDI note velocity resulting in relatively higher initial sound level) as the cue of metrical interpretation. This is in agreement with the general assumption of the punctuate/impulse nature of meter and the interrelationship between phenomenal and structural accents (see section 5.1.4, *Assumption no 21-23*).

The test melody was constructed to indicate asymmetrical grouping, which was supposed to be culturally distant to the majority of participants in the experiment. Therefore, two examples with consistent symmetrical grouping were included as well as an interpretation which was phase displaced to the other examples.

The main hypothesis was that Westerners would be sensitive to indicated asymmetrical grouping.

**Trial 4 – metrical ambiguity by sequencing at measure level**

This test melody was composed to indicate ambiguity between the metrical structure at measure level indicated through interonset structure in the beginning and mid-end of the melody and the indication of metrical structure through melodic sequencing in the middle part of the melody. The former was composed to indicate grouping in 2 or 4 beats at central pulse level while the latter was composed to indicate periodical grouping in 3 beats.

Chapter 5

Yet another ambiguity was involved here. The beginning of the melody (the part with indication of 2 or 4-beat grouping) was composed to suggest ambiguity between grouping by discontinuity and repetition. More specifically between relatively long duration as group ending indication (the segmentation by interonset distance principle, see section 5.2.4.3, *Assumption no 28*) and sequence start as start indication. The start was intended to sound as a typical ending in 18<sup>th</sup> century common-practice Western music, while the first identical and continued sequence starts on the note of longest duration.

The melody was, as a whole, meant to remind the Western listener of 18<sup>th</sup> century common-practice Western music to invoke expectations of symmetrical and hierarchical structure by a listener with experience of Western classical music in general.

These different interpretations are briefly covered through four different metrical interpretations indicated by percussion accompaniment. Two of these are sensitive to the melodic sequence period in the middle while the two others are not. Two of the interpretations likewise favor the four-beat period indicated by discontinuity while the others favor sequencing.

### testmelody 04

composed to indicate ambiguity regarding start- and end-rhythm and with melodic sequence athwart established meter

Alternative 1

Musical score for Alternative 1. It consists of two staves: Piano (top) and conga/tom (bottom). The Piano staff has a treble clef and a key signature of one sharp (F#). The conga/tom staff has a bass clef and a 2/4 time signature. The score is divided into two systems. The first system has 8 measures. The second system has 8 measures, with the final measure ending with a double bar line and a fermata.

Alternative 2

Musical score for Alternative 2. It consists of two staves: Piano (top) and conga/tom (bottom). The Piano staff has a treble clef and a key signature of one sharp (F#). The conga/tom staff has a bass clef and a 2/4 time signature. The score is divided into two systems. The first system has 8 measures. The second system has 8 measures, with the final measure ending with a double bar line and a fermata.

## Alternative 3

Musical score for Alternative 3. The top staff is a melody line in treble clef, and the bottom staff is a bass line in bass clef. The melody consists of eighth and sixteenth notes. The bass line consists of quarter notes. The time signature changes from 2/4 to 3/4 to 3/8.

## Alternative 4

Musical score for Alternative 4. The top staff is a melody line in treble clef, and the bottom staff is a bass line in bass clef. The melody consists of eighth and sixteenth notes. The bass line consists of quarter notes. The time signature changes from 2/4 to 3/4 to 3/8.

Figure 5-79. Test example 4 (tempo: ♩ = 96)

Regarding the change of metrical grouping from 2/4 to 3-beat grouping, which was indicated the main hypothesis was that there would be a tendency towards this interpretation, however weak, among the listeners. The pre-test as well as other previous experiments have evidently shown the firmness of the listeners' first assumption, especially when the interpretation turns out to be generally dominant. Such deviation from the main metrical interpretation is often designated *hemiola* syncopation against the prevalent pulse.

Regarding the conflict between discontinuity and sequence-grouping in the beginning, I expected no significant tendency because of the possible solution of adapting to a persistent grouping in two beats and leaving the confusing higher levels ungrouped!

### **Trial 5 – metrical ambiguity/indication of complex meter**

The next test melody was a Norwegian *gangar* melody, Kjempe-Jo, from the repertoire of the fiddler Salve Austenå from Tovdal, Norway. The use of an existing melody makes it difficult to measure the influence of prior knowledge of the melody – including a prior opinion of the metrical interpretation - on the results. Nevertheless I chose to use this melody partly because it is an interesting example of metrical ambiguity on a structural level and partly because it has been used as an example in discussion about metrical conception in Norwegian *gangar* and *halling* melodies (Kvifte 1983, Kvifte & Blom 1986). The melody is also not among the most well-known Norwegian tunes in Sweden. Only one of the participants, a performer of Norwegian folk music, commented that the melody reminded her of a tune she knew.

## *Chapter 5*

I have used a version of the tune recorded in 1991 (Austenå 1991) and since it is quite long only about half of it is used. I have also made a radical reduction of the performance in the transcription below, omitting chords, accompanying drones, ornamentation, articulation/bowing and last but not least rhythmic changes below eighth note level. In addition some extra repetitions were omitted, that is more than one repetition of the same sequence. The reason for this general reduction of the melodic structure was twofold: (i) To obtain a stimulus at the same level of specification as the other examples in the test which would concur with the object of the test and (ii) to adapt the melody to the level of specification at the rhythmic level used mostly in traditional transcriptions of this kind of music.

The structure of the melody, begins with an interonset structure that can be interpreted as indicating duple division, i.e. a pulse duration of  $2/8$ , while the next section has an interonset structure that indicates triple division, i.e. a pulse duration of  $3/8$  (see below). The indication of change of pulse level is especially emphasized in the performance of Austenå through articulation (bowing patterns), through grave articulations and uneven rhythmical division of the eighth notes.

The alternatives were in this case (i) change of pulse duration (ii) consistent duple division (iii) consistent triple division (the alternative suggested by the foot tapping of the performer) (iv) complex meter with superimposition of duple division on triple division. These different alternatives were indicated by percussion accompaniment.

The tempo was based on the initial tempo of the original performance.

The hypothesis was that the participants would be sensitive to the change of indication of central pulse level favoring either alternative a) or d). An alternative hypothesis was that the consistent triple division would be favored because it represented the least conflict with interonset structure.



## Kjempe-Jo

played by Salve Austenå, Tovdal, Norway

simplified notation of 55 measures based on chain of motifs represented in recording from 1991 (Austenå 1991)

Chords, acc. strings, ornamentation and changes of duration below eighth note level omitted as well as identical extra repetitions of repeated phrases (1st round)

The image displays a musical score for the piece 'Kjempe-Jo'. It begins with a piano melody in G major (one sharp) and 6/8 time, consisting of six measures. Below the piano part are four alternative conga accompaniment interpretations, each labeled 'alt.' and numbered 1 through 4. Each alternative is shown for six measures, with measure numbers 2 through 6 indicated above the staves. The first alternative is labeled 'changing beat 3/2'. The second is '3 beats per measure'. The third is '2 beats per measure'. The fourth is 'super-imposed pulse - 3 against 2'. The piano part continues with measures 7 through 11, which are repeated in each of the four alternative accompaniment systems. The piano part consists of eighth and sixteenth notes, while the conga parts use various rhythmic patterns including eighth and sixteenth notes, rests, and accents.

Figure 5-81. The different alternative interpretations of test melody 5. Each system represents an alternative accompaniment.

A subsequent test was also performed, based on the hypothesis of changing indication of pulse level, the object of which was to test whether the participants favored sequence start on a moderately longer duration or start after longer duration, i.e. longer duration as metrical accent vs. longer duration as distance and group ending.

This was performed through different accentuation of percussion accompaniment of the first 11 measures of the melody.

Test of start accentuation preference in relation to sequence

test melody  
piano

### Tst ex 06

Raag Suddha Danyasi (after K Shivakumar, Bomba)  
Beginning of varnam. Based on indian brief notation by performer (transl  
Repetitions and ornamentation omitted.

Figure 5-83. The melody used in the experiment. (tempo: ♩ = 110)

### Tst ex 06

#### Raag Suddha Danyasi (after K Shivakumar)

Based on indian brief notation by performer (translated into western nota  
Repetitions and ornamentation omitted

Figure 5-84. Principles of alternative interpretations displayed on different systems.

The main hypothesis was that the subjects would be sensitive to the change of group size indicated by pitch sequencing, discontinuity and interonset structure. Alternatively the subjects would prefer a consistent grouping indication with four atomic beats per group starting on the first event.

**Trial 7 – Periodic asymmetric beat – Swedish polska**

In this test the object was to examine whether participants would be sensitive to asymmetrical beat grouping as indicated by the interonset and melodic contour structure of the melody. The melody used here was a rather well-known polska melody within Swedish folk music, transcribed by me from a performance of Gössa Anders Andersson, Orsa, Dalarna (Andersson 1949) (see also section 5.2.4.8). Again I omitted all ornamentation, articulation/slurs and use of sympathetic strings as well as rhythmic variation below eighth note level used in the original performance. Only the first half of the melody is used and the tempo was consistent with the mean tempo of the original recording. Alternative interpretations were indicated by the percussion accompaniment as can be read from the notation below. The alternatives were (i) periodic asymmetrical grouping (2+4+3 /8) with grave accent on first event of the melody, (ii) symmetrical grouping (3/8) and (iii) periodic asymmetrical grouping (4+3+2/8) starting on the second event of the melody (preference for start on the longest beat).

**Test ex 07**  
*Polska after Gössa Anders Andersson (1st part)*

Orig. played on the fiddle. Ornamentation, articulation, chords and accomp. open strings omitted.  
 Rhythmic variation quantized below eighth note level

test melody  
 piano

alt. 1. periodic asymmetric grouping, grave accent at start

bongo  
 conga mt

alt. 2. periodic symmetric grouping, grave accent at start

bongo  
 conga mt

alt. 3. periodic asymmetric grouping, acute accent at start

bongo  
 conga mt

5 6 7 8

5 6 7 8

5 6 7 8

Figure 5-85. *Polska after Gössa Anders Andersson* (tempo:  $\text{♩} = 180$ ). Alternative interpretations are displayed on different systems.

**Trial 8 and 9 – Melodic parallelism. Nested asymmetrical melodic sequencing versus group size consistency**

Test examples 8 and 9 consisted of two artificial melody excerpts composed by me in order to examine the strength of melodic pitch sequences versus tendency towards consistent symmetrical grouping. In example 8, the sequences were identical in pitch content, while in example 9 the sequences were transposed to test the impact of transposition on the perceived primary grouping. The overall style of the melodies was intended to indicate a symmetrical structure typical of common-practice Western music. The beginning of the melody was constructed of a sequence with a slight change at the end of the second repetition. At the point of this change, another sequence started with an exact repetition. The question was whether the symmetry indicated by the first repetition would be preferred to the grouping which consistently was based on the sequences or if it did not matter at all.

Since these melodic examples are rather ambiguous in terms of structure with nested sequences, I chose to use melodic phrasing as a means to indicate grouping. This was obtained by the insertion of rests between indicated groups together with change of tempo within the indicated groups. This artificial melody was unknown to all subjects. The alternatives in both examples were intended to regard primarily grouping at measure level; the central pulse level has, in a previous experiment (see section 6.3, *experiment 3*), been generally considered as 2/8. Therefore, the alternatives were (i) consistent grouping in six elements, (ii) consistent grouping in eight elements (concurrent with 4/4 time or 8/4 time) (iii) asymmetrical grouping according to adjacent sequences and (iv) grouping by all major discontinuities, i.e. pitch distance.

**Test ex 08**

*"Melody" composed to test the influence of parallelism on metrical grouping. Composed to indicate ambiguity regarding grouping by parallelism, grouping by discontinuity and symmetrical grouping*

Grouping indicated by means of a general accelerando-rallentando pattern and rests between indicated groups. (This is indicated in the notation below through beams, breath marks and slurs).

Testmelody



alt. 1 (consistent grouping in six elements)



alt. 2 (consistent grouping in eight (4+4))



alt. 3 (grouping by adjacent sequences)



alt. 4 (grouping by all discontinuities)



*Figure 5-86. Parallelism. Asymmetrical sequencing (tempo varied around: ♩ = 104). Alternative interpretations displayed on different systems.*

### Test ex 09

"Melody" composed to test the influence of parallelism on grouping. Composed to indicate ambiguity regarding grouping by parallelism, grouping by discontinuity and grouping consistency. Here including transposition of segments

Grouping indicated by means of a general accelerando-rallentando pattern for beginning and end of groups and rests between indicated groups. (This is indicated in the notation below through beams, breath marks and slurs).

Test melody

alt. 1 (group of six (2+2+2))  
 accel.:... rall. accel.:...rall. accel.:...rall. accel.:...rall. accel.:...rall. accel.:...rall.

alt. 2 (adjacent sequences)

alt. 3 (every discontinuity)

alt. 4 (consistent grouping in eight (4+4))

Figure 5-87. Asymmetrical sequencing (tempo varied around ♩ = 104). The alternative stimuli displayed on different systems.

### Results

The test results are presented in the tables below. The mean of the results, are presented in a table for each trial. This table shows the mean result, a computation of significance by means of analysis of variance (ANOVA) or t-test, depending on the number of stimuli and factors involved.

#### Trial 1 and 2a metrical interpretation in relation to pitch content and melodic contour

Even if these tests were not presented in immediate succession in the actual test, they are presented together since they are connected in the design of the stimulus material (see above).

The main hypothesis was that the different ordering of the pitches of the two melodies would result in different preference of metrical interpretation.

	3/8 beat	4/8 beat
Test melody 1	3.328	1.5
Test melody 2	1.143	3.3

Table 10: Average result of preference for metrical interpretation of test melodies 1 and 2

The interaction between test melody and metrical interpretation was highly significant  $p < .0001$  (2-way related ANOVA  $F_{1,83}=74.55$   $p < .0001$ ), while the test-melody and metrical preference alone was not significant  $F_{1,83}=0.40$   $p=0.5293$  and  $F_{1,83}=0.71$   $p=0.4021$  respectively.

The preference regarding each example taken separately was also significant: For test melody 1  $p = 0.0002$  (t: -4.22) and for test melody 2  $p < .0001$  (t: -5.85)

**Trial 1 and 2b finality rating in relation to perceived metrical structure**

		final chord alternative			
		C	A	F	E
mean	melody 1	1.857	1.667	3.571	3.048
	melody 2	2.095	2.667	2.095	3.143

Table 11: Mean values of rating of chord finality for test melody 1 and 2, rated from 1 to 4, where 1 is most final and 4 is least final.

As can be seen in the above table the predicted difference in ratings between the two different test melodies did occur in terms of mean. The interaction term between test-melody and final chord alternative was significant (2-way within subjects ANOVA)  $F_{3,167} = 12.94$   $p < 0.0001$ , as was final chord factor alone  $F_{3,167} = 13.58$   $p < 0.0001$ , while the test-melody factor alone was not significant  $F_{3,167} = 0.06$   $p < 0.8050$ .

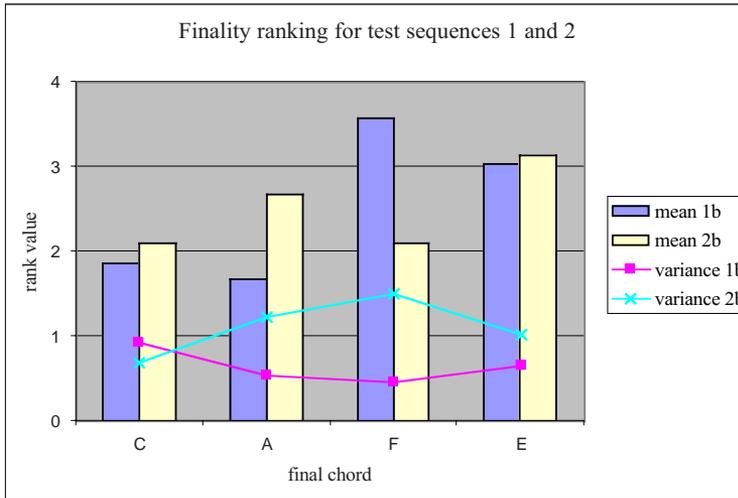


Table 12: Comparison between finality ranking for test melodies 1 and 2. Note that lowest number represents highest rank.

Note the high variance of the F chord alternative for melody 2, considering the ratings are made on a four-grade scale. This reflects a grouping of the responses of the subjects around

two opposing strategies, one rating F as the most final chord with regard to the second melody and the other group rating F among the least final chord alternatives in both alternatives.

As can be seen in the diagram above, the A chord alternative received the highest rank for the first melody, while it was regarded as only the third best alternative for the second melody. In this case, the variance increased also in the rankings of the second melody.

The C chord alternative received a relatively high rank for both test melodies (second and first rank) while the E chord alternative was relatively low ranked for both melodies.

**Trial 3 Indicated asymmetrical grouping**

In this trial a preference for asymmetrical grouping was hypothesized.

	3/8-grouping	4/8-grouping	asym. grouping	3/8-group. displ.
Mean	2.762	1.000	3.619	1.524

Table 13: Results (mean values) for trial 3.

This was a highly significant result supporting the current hypothesis,  $F_{3,83} = 43.80$   $p < 0.0001$ , which was a bit surprising since I thought the largely symmetrical musical structure, which is dominant in the musical background of the participants, would more strongly influence their preference for grouping by accentuation.

All the differences between alternatives were significant (Tukey 95% CI), except for the difference between 4/8-grouping and 3/8-grouping.

**Trial 4 Metrical ambiguity at measure level by sequence contra-metrical to the established meter**

The experimental hypothesis was preference for change of metrical interpretation (however weak). This was not supported by the data.

	cont. meter	varied meter
mean	2.952	3.000
variance	0.973	1.000

Table 14: Mean and variance for trial 4.

The mean values (based on highest values for each category) were slightly higher for varied meter, but the difference was very small and the variance was considerable. Thus, also in this case, it appeared that the participants could be divided into two groups, one with preference for metrical change and the other for metrical constancy. Both these groups were substantially large, while none presumed strategies could be rejected.

The difference was not significant  $t_{1,20} = 0.14$   $p = 0.4452$

Comparing all of the alternatives, the division between the two alternative strategies shows even more clearly:

	4-group feminine	4-group masculine	4/3-group feminine	4/3-group masculine
mean	2.857	1.905	1.571	2.571

Table 15: Mean results for trial 4, with regards to consistent 4/8 accompaniment vs. varied accompaniment and feminine and masculine interpretation respectively.

The factor of metrical interpretation was significant  $F_{3,83} = 6.21$   $p=0.0008$ , but the only significant differences were between 4-group feminine and 4-group masculine and 4/3-group feminine respectively and between 4/3-group masculine and 4/3-group feminine respectively.

The second hypothesis predicted that the differences between preference for start on longer duration (*masculine* start preference) or end after longer duration (*feminine* start preference) would not be significant. This turned out to be true. The mean values based on highest rating for each category show only a small difference between the groups.

	feminine start	masculine start
mean	3.25	3.13
variance	0.21	0.98

Table 16. Mean and variance values for trial 4, with regards to preference for feminine and masculine start.

The differences between choice of feminine or masculine start was not significant for this group of subjects ( $p = .7513$ ), in this case also a lower level of variance within groups.

**Trial 5a – metrical ambiguity at beat level**

In this trial, the experimental hypothesis was that the metrical change would be recognized resulting either in preference for the alternative signifying changed central pulse level or the alternative signifying consistent complex meter. This experimental hypothesis was supported by the data.

	2/8 beat	3/8 beat	changing or complex
mean	1.857	2.048	3.571

Table 17: Mean values for trial 5a, with changed or complex metrical interpretation grouped.

$F_{2,62} = 22.86$ ,  $p<.00001$ , where the differences between changing or complex and the continuous/simple alternatives were significant (Tukey 95% CI). Means based on highest ratings for the combined data.

When changing or complex meter are measured separately there is no significant preference for changed or complex meter in the data.

	changing	3/8-beat	2/8-beat	3 against 2
Mean	2.905	1.857	2.048	3.143

Table 18: Mean values with regards to preference for changing or consistent metrical interpretation.

$F_{3,83} = 8.17$   $p<0.0001$ , showing significant differences between all pairs but changed and 3 against 2 and 3/8 and 2/8 beat grouping.

**Trial 5b – feminine start vs masculine start preference**

In this trial the experimental hypothesis was that there should be a preference for feminine start, i.e. start after long duration. This hypothesis was supported by the data, however weak, and not on a <0.01 level of significance.

	feminine start	masculine start
mean	2.952	2.048

*Table 19. Mean values for feminine and masculine preference in trial 5b*

$t_{1,20} = 1.98$   $p=0,0309$ .

***Trial 6 – changed asymmetric grouping on beat level***

The experimental hypothesis here was that participants would be sensitive to the indicated asymmetrical grouping and prefer the accentuation pattern which concurred with such a grouping.

This hypothesis was supported by the data, however weakly:

	4/8 beat	changed	phase displaced 4/8
mean	2.2	3.0	1.3

*Table 20: Mean results for trial 6.*

$F_{2,59} = 12.23$ ,  $p < 0.0001$  for the impact of metrical interpretation factor. The pair-wise differences between the alternatives were significant except for the difference between 4/8-beat and changed, which was not significant (Tukey 95% CI)

***Trial 7 – consistent asymmetrical grouping on beat level***

The experimental hypothesis was that asymmetrical beat grouping starting on the first beat (short-long-medium) would be preferred. The data strongly supported this hypothesis:

	2+4+3/8 beat grouping	3/8 beat grouping	4+3+2/8 beat grouping
mean	3.8	1.9	1.6

*Table 21: Mean results for trial 7.*

$F_{2,59} = 23.76$ ,  $p = 0.0001$ . The differences between the 2+4+3/8 beat grouping and the other grouping alternatives being significant (Tukey 95% CI).

***Trial 8 – asymmetrical sequencing vs consistent group size on measure level a) identical pitch sequences***

The experimental hypothesis was that asymmetrical grouping by sequencing would be favored before a continuous group size. I expected this tendency to be weak. The data supported the experimental hypothesis, however, more strongly than was expected.

	6/8 group	8/8 group	changed by sequences	changed by discontinuity
Mean	1.5	2.5	3.5	2.0

Table 22: Mean results for trial 8.

$F_{3,79} = 27.25$   $p < 0.0001$ , in which all mean differences were pair-wise significant (Tukey 95% CI) except for the difference between the 8/8-grouping and changed by discontinuity alternatives.

**Trial 9 – asymmetrical sequencing vs. consistent group size on measure level b) transposed pitch sequences**

The experimental hypothesis was the same as above (trial 8).

The data supported the hypothesis, but not as consistently as in trial 8.

	changed by discontinuity	changed by sequences	8/8 group	6/8 group
Mean	1.8	3.1	2.3	1.6

Table 23. Mean results for trial 9.

$F_{3,79} = 14.50$   $p < 0.0001$ , in which the differences between pair-wise conditions were significant for the changed by sequence alternative in relation to the other alternatives (Tukey 95% CI).

**Known error variables**

According to a study on the relative impact of different grouping factors performed by Deliège (Deliège 1987), the order in which segmenting factors are presented in a melodic excerpt influences the result. It indicates that we choose the segmenting factor first presented and disregard subsequent contradictory information. Since this mechanism could, in a transferred sense, be an important factor also in the above trials, I deliberately varied the position of the hypothesized response in the order of alternatives.

If the order in which the alternatives were presented had impact on the result, it would presumably be apparent by favoring of either the first or last alternative in each trial.

Another factor, which may have influenced the choice of preference, was cultural background, assuming that metrical preference is related to metrical patterns within culturally delineated musical styles. The number of people of different musical background participating in the tests was however, too small to be able to measure the influence of this factor properly.

Yet another factor, which could have an influence on the outcome of the tests, is the number of years of formal music education. In these experiment, the group consisted of performers at a high level (conservatory level music students), of whom the majority had a considerable amount of years of formal music education. The influence of this factor was thus not possible to evaluate properly.

Other possible factors such as age and sex are also hard to evaluate properly because of the limited number of subjects and regarding age relatively low variability. There were exactly the same number of women and men participating in the experiment.

However, when measured in the current data, none of the above factors could be shown to be significant.

<b>Error factor</b>	<b>Mean Difference</b>	<b>Significance measure</b>
Trial vs. Rank	Significant pair-wise (95% confidence)	p < 0.0001
Order of stimulus presentation vs. Rank		p = 0.7028
Cultural background vs. Rank (non-Western/Western)	1.833 (non-Western) 2.111 (Western)	p = 0.2524
Age vs. Rank		p = 0.6074
Years of formal musical training vs. Rank		p = 0.7898
Sex vs. Rank	2.1 (feminine) 1.938 (masculine)	p = 0.3220

*Table 24. Error factors in relation to Mean Difference and Significance measure*

For the two factors which have  $p < 0.5$ , I have added the mean values for the groups. It shows that the found mean differences were very low for these variables, which, together with the low significance measure, cannot be interpreted as a support for the influence of either factor on the results.

However, the only error factor, which can be rejected on the basis of the above test, is the order of stimulus presentation in relation to trial. This factor had no significant impact on the results.

The influence of gender can also be considered very low and considerably lower than the influence of the individual. Since data is based of group of equal number of men and women the data here provides more reliable measures.

With respect to the other factors, the group of participants was either too small, not symmetrical enough regarding categories, not involving enough variability regarding years of musical training or age to rule out the possible influence of these factors.

## **Discussion**

### ***General discussion***

The results of the set of trials within experiment 2 can generally be said to strongly support the hypothesis that, within a cultural context where the concept of melody is significant (at least with a predominantly Western background, at least musically trained etc.) people use features of melodic contour, i.e. grouping of melodic events, as a basis for metrical

perception and conception. Thus, the result of the experiment 2 generally supports the design of the complex metrical analysis, which is based on the above hypothesis.

What may seem questionable is the assumed linkage between the indication of metrical structure in the stimulus material and metrical perception. Does the preference of a metrical accompaniment, accentuation or phrasing, regarding the best fit with the melody, conform to the subject's metrical perception? I consider this linkage valid of a number of reasons. First, the means of indications metrical structure with percussion accompaniment, accentuation patterns and phrasing are used widely in musical performance in connection with metrical structure. Further, this linkage is supported by the notation of metrical structure in identical stimulus material of another test, which functioned as a pre-test for the current experiment. (See Chapter 6, section 6.3 *Experiment 3*). The hypothesized alternatives used in the current test were derived from the results of that test. The current test results showed for most trials, significant choices of preference, which strongly concur with the notation of metrical structure in the earlier test.

The results showed no trend towards symmetrical and asymmetrical grouping respectively that could be traced to which means of indication of grouping used.

There is a notion, which is often mentioned among musicians, that meter can be inferred from a melodic (or more generally musical) structure in many ways, resulting in a number of metrical interpretations. Two of the participants (no 3 and 10) in the current experiment did comment that the stimulus melodies were possible to hear in different ways regarding grouping/meter. It is obvious from a number of pieces in which meter is subject to variation on a given melody that this is perfectly possible.

This notion is reflected in a number of different ways in the literature. In an interesting article entitled 'Inferential Ambivalence in Musical Meter' (Blom & Kvifte 1986) concerning the perception of meter in "Norwegian *gangar/balling* melodies the two authors base their discussion of metrical perception on the following assumption:

*"The discussion and musical analysis assumes that peoples' ability to produce, recognize and experience patterns of sound depends on the acquired perceptual structures of performers and listeners (Blacking 1973:10) and that important cues to structure and meaning are only partly, if at all, transmitted through the auditory channel. (p 491)*

Furthermore, Kvifte points out some general aspects of meter as phenomenon:

*That is: meter is something one uses: a way of ordering sounds in a musically meaningful fashion. Meter is not "in the music," but in the mind. That is, by the way, also obvious from the fact that metrical signs in written music - time signatures and barlines - are not represented by distinctly audible features in the music. One can pick out the pitch, say, "a-flat" from the sound of the music; barlines can only be inferred, not directly perceived. (p.495)<sup>58</sup>*

From a point of view of cognitive psychology or psychoacoustics one would want to add that also categorization of pitch in a musical context, e.g. the distinguishing of an a flat used in the example from a g sharp or the perception of pitch quality of a sound in relation to timbre quality, is as much as the perception of musical meter something which is inferred by the mind of the listener and performer, and could this be due to cultural learning. One can also ask if Kvifte believes that a strict interonset periodicity of musical sounds – meter as an

---

<sup>58</sup> Note that Kvifte does not rule out the possibility that qualities of musical sound could influence the perception of musical meter "The first approach could be to argue that even if meter is not present in the sound, as a listener one will nevertheless have to rely on information that is present in the sound." (p. 495)

evident structural quality of the sound of music - would not be audible and recognized by a listener?

There is also a general problem revealed in the fundamental assumption cited above – if the cues to structure are not transmitted through the auditory channel – how do people learn when to apply which structure? Not every musical activity is strictly regulated as in e.g. a rite procedure where you collectively learn how to move to a certain musical stimulus. Even in such situations, – how do you confirm that people have the same perception of the stimulus?<sup>59</sup>

The basic assumption of metrical perception, which is tested in the current experiment, is that people in general are sensitive to pitch and interonset periodicity in melodic stimulus material. The results of the experiment support this claim. The influence of cultural learning on this sensitivity has yet to be investigated. However, the data does not support that cultural learning should be the only determining factor of the perceived meter, ruling out every aspect of structural properties of the sound stimulus.

### ***Trial 1 and 2 – Same pitch content, different order***

The result of this test was highly significant with regards to the interaction between melodic contour and the choice of metrical interpretation as was predicted by the experimental hypothesis. What can be derived from this result?

The means of indicating a 3/8 - beat structure were rather modest in the first stimulus melody: The important clues were meant to be change of melodic direction (global and local) and step-size at positions which would conform to a 3/8 metrical grid together with rather weak parallelism (weak contour sequence of 12/8 period) indication. A more subtle cue was the 3/8 sequence in the very beginning, supported by the continuous motion broken at the twelfth position. The changes of step-size and melodic direction conformed to possible Western cultural cues towards the end of the melody by the forming triads of thirds delineated by change of melodic direction and step-size. In summary: discontinuity as main cue to grouping, parallelism as a secondary.

The means of indicating 4/8 - beat structure in the second stimulus melody was change of general (global) melodic direction at the fourth position which was meant to confirm the ambiguous sequence of broken thirds of the four first notes indicating two levels of primary grouping at 2/8 and 4/8 - level. Furthermore, changes of melodic direction and step size at positions fitting in with the metrical grid. The 4/8 beat period was also thought to be supported by parallelism through a sequence of 16/8. In summary: discontinuity as main cue to grouping, parallelism as a secondary.

These rather subtle means of indicating a metrical structure were obviously successful, although they involved neither true parallelism in terms of full or exact sequences nor extreme changes of step sizes such as register changes. It is possible that this success can be attributed to cultural learning since I have used some common patterns such as broken triads. The data from the participants with background in foreign musical culture, however, did not support this notion. I would expect the result to be less evident with participants outside the Western environment. Three of the participants reported that the melodies sounded a bit artificial.

---

<sup>59</sup> See also comment by Blacking (1973)

In the subsequent task of ranking finality, the general hypothesis that the melody contour would influence the rank of finality seem to be valid for a substantial number of the participants. I interpret these differences in finality ranking as an indication of a difference in perceived tonality for the two melodies. The melodies were composed so as to indicate A-minor/aeolian and F Lydian respectively, basically by the placement of assumed hierarchically prominent melodic pitch categories / scale degrees in the modes on first event of implied metrical groups.

The implied tonality difference between the two melodies can be said to be weakly supported by the results. For the first melody, the A-chord received the highest rank, and, for the second melody, the F chord got the highest mean rank together with the C chord. In both cases, however, the C chord was highly ranked which indicate that pitch set matching with a culturally learned pattern (the notes of C major) should not be disregarded.

The result of the test can be interpreted as support for the notion that perceived grouping structure affects tonality perception for some individuals in a predominantly Western cultural context.

### ***Trial 3 – Indication of asymmetrical grouping***

The result of this trial confirmed the main experimental hypothesis that the asymmetrical grouping would be preferred by the subjects, even if it did not conform to the culturally dominant pattern of symmetrical grouping patterns.

How can this be explained? My interpretation is that persistent parallelism both on beat and measure level confirmed a consistent grouping pattern in an evident way. On beat level, the repeated rising third pattern delineated by an ascending stepwise figure was meant to indicate 3 groups of 2/8 and one group of 3/8. On measure level, this was repeated in the second measure forming a perfect sequence. The following implied measures which were also sequenced conformed to the first pattern.

It is perfectly possible that the use of the “à la Turk” theme as a metrical model for this melody could give a clue to the metrical grouping by melodic resemblance. (None of the participants commented, however, on this resemblance). But the melodic content (dominated by thirds and seconds) is general enough to make associations with a number of symmetrically grouped melodies (e.g. Irish jigs), which can also be considered to belong to the musical styles known by the participants.

The result of this experiment therefore indicates the strength of melodic parallelism as an input to metrical grouping.

### ***Trial 4 – Metrical ambiguity on measure level / Strength of melodic parallelism in relation to grouping established by interonset patterns***

In this trial, the data did not support the experimental hypothesis that the majority of the participants would change prefer a changing metrical grouping related to melodic pitch sequences obstructing the established metrical pattern.

The results rather indicated that the melody was ambiguous in relation to meter. It could go either way: Close to half the group of participants chose the two possible types of interpretation. The high level of difference between the ratings of the alternatives indicated that the choice was not due to insignificance or chance (the mean difference between highest and lowest rated alternative was  $> 1$ ).

This result can be interpreted as indicating the existence of different principles influencing metrical grouping: The influence of the established pattern and the influence of melodic pitch sequence as a cue to metrical grouping.

The ambiguity of the rhythmical structure of the example was confirmed by the result, but, interestingly enough, the preferred continuous alternative indicated preference for the first presented pattern, while the preferred changed alternative indicated sensitivity to the sequence structure analogously with the preference of sequence as a cue to metrical grouping.

**Trial 5 – Metrical ambiguity at beat level – complex and changing meter**

In the paper cited above ('Inferential Ambivalence in Musical Meter' (Blom & Kvifte 1986)) Kvifte makes another interesting statement regarding musical meter:

*“One more point: One can use only one single meter at a time. If one goes back to the rhythm in Example 2, one can choose to perceive it in 6/8 or in 3/4, but it is impossible to perceive both meters simultaneously. One might be able to perceive the common pulse shared by the two meters and thus form a bridge between them when changing from one meter to the other, but that is not the same as using both meters at the same time” (page 495)*



Figure 5-88. Adapted from Kvifte & Blom (1986:495)

That Kvifte makes such a rigid statement that is not supported by any reference or data is interesting.

In trial number 5, one of the alternatives indicated by a metrical structure which Kvifte excludes, which would form a third alternative to the metrical interpretations offered by Kvifte:



This illustrates the conception of two simultaneous pulses at about central pulse level in cooperation - complex meter - which results in the above interonset structure (ex *a*). This illustration does not include a conception of a central pulse (tactus) but if this is included there are yet two more interpretations possible, 2 beats to 3 or 3 beats to 2. In the sound example, the two layers are emphasized by different timbre, loudness and articulation.

This alternative had the highest mean rating, compared with single pulses corresponding to Kvifte’s alternatives *b*) and *c*) and a changing meter alternative corresponding to a change between Kvifte’s alternatives *b*) and *c*).

This result can be interpreted as indicating the perceptual reality of conception of super-imposed pulses forming a complex meter and thus seem to be in striking conflict with Kvifte’s statement. One can object to this interpretation in that it is possible for the subjects to perceive it as a one-layered rhythmical gestalt (interpretation *a*). But why would it be strange or impossible to perceive a meter as two pulses at different paces in cooperation? This particular case is only a special case regarding the well-documented phenomena of metrical

complexity to which the simultaneous conception of meter at measure level and at pulse level belong (see e.g. London 2002:541ff).

A complex meter implies the conception of simultaneous metrical layers in cooperation. The foot tapping at 3/8-level, together with the articulation and melodic structure of a *gangar* melody like ex. 5 which frequently emphasizes the 2/8-level, indicates such complexity quite convincingly. It can even be traced in the movements of traditional dance patterns of *gangar* and *balling* (Blom & Kvinfos 1986:508). The question, which is the core of Kvinfos's argument, whether or not the melodic meter is in conflict with the dance meter (1986:491), becomes immaterial if one accepts the concept of complex meter. According to this interpretation, the meter indicated by the melody and the foot tapping forms a complex meter. If one interprets Kvinfos's statement as concerning the need for a referential central pulse level – a *tactus* – it would be easier to accept in the light of the current results. It remains to be shown, however, that the experience of a *tactus* is always present in metrical conception (see also Meyer & Palmer 2001 quoted in London 2002:541).

The result of this trial indicates that subjects could infer the fundamental features of the metrical complexity articulated in the original performance of this melody from the very reduced stimulus material, from periodicity in interonset and pitch structures. This indicates that musical sound can convey cues about metrical structure to people without *emic* knowledge, which could be relevant also in a cultural context.<sup>60</sup>

#### ***Trial 5b – Feminine vs. masculine rhythmic grouping - Interpretation of interonset distance as a cue to group start***

The underlying question behind this trial was if a moderate difference in interonset interval / duration (2:1) in the beginning of this melody would signify ending, according to the distance principle implying a segment start after the long note, or if it would signify start, according to the prominent event principle, implying start at the long note (see *Assumption no. 3f*)?

The experimental hypothesis was that the melodic structure would make the subjects favor the long note as end-interpretation, because of the relatively high number of shorter notes of the repeated sequence.

This hypothesis was not supported by the data at a significant level. Even if there was a small mean emphasis of the feminine (end-oriented) interpretation the number of participants which preferred the first sequence start /masculine (start-oriented) interpretation was considerable.

This indicates that this melody is quite ambiguous in this respect and that 2:1 relationship can function as start indication as well as end indication on this grouping level.

#### ***Trial 6 – Changed asymmetrical grouping on beat level in a South Indian melodic excerpt***

In this trial, the test melody originated from a South Indian raga, *Raag Suddha Danyasi* transcribed from an Indian notation without ornamentation and rhythmic and melodic elaboration and without the repetition of the major sections.

---

<sup>60</sup> This interpretation of the results conflicts with Blom's general statement: "sound structure probably does not offer sufficient clues for making a definite metrical choice, even among known alternatives." (p514)

The research question was: Would the participants prefer the asymmetrical possibility which briefly resembles the Indian original melodic grouping, which is derived from melodic discontinuities and parallelism, or would they prefer the symmetrical versions more in concordance with common-practice Western music?

The result showed that the subjects actually preferred the asymmetrically grouped version in concordance with the experimental hypothesis. This is interesting because the symmetrical version is structurally supported in numerous measures. Compared with the evident sequencing in trial 4 at measure level, the grouping indications are quite ambiguous in trial 6, and more so, trial 4 is composed in a style which is well known to most of the participants which could make the change in grouping easier to track. In spite of this, the subjects were in general more open to an asymmetrical grouping in trial 6. One possible explanation is that the lack of culturally known cues in trial 6 made the participants more open to asymmetrical grouping. Another possibility is that the lack of a persistent, and thus established, primary grouping at beat and small-scale measure level increases the participants sensitivity to finer cues of grouping, such as pitch sequences and pitch distance. Note however that the first symmetrical grouping alternative received a considerable amount of general preference and that the difference between the two alternatives was not significant at a 95% level (See *Results, Trial 6*, above).

The result is interesting in the light of the above discussion (trial 5) regarding the possibility to infer a metrical conception from the sound of music which could be relevant from a culture-specific perspective.

Even though the grouping indicated by the accent structure in trial 6b, which was most commonly chosen as fitting best with the melody, did not fully concur with the grouping indicated by the Indian musician, still the main features of the grouping structure coincided, the segment borders in common being about 90% (cf. section 5.2.4.8 *Asymmetrical grouping in South Indian Classical Music*). However, making such a comparison one has to consider that the Indian notated grouping does not necessarily reflect the perceived grouping by an experienced Indian listener.

To conclude, it seems possible to track the main features of metrical structure in this example in spite of lack of cultural learning.

### ***Trial 7 – asymmetrical grouping at beat level – Swedish polska meter***

The experimental question in this trial also concerned the strength of structural cues to asymmetrical grouping in relation to expectancy of symmetrical grouping. In this case the asymmetry was at beat level and the melody was chosen to indicate recurrent asymmetrical beat length in a context of strict periodicity at measure level.

The results were extremely clear in favor of asymmetrical grouping. In interpreting this result, one has to consider that most Swedish *polska* melodies have symmetrical primary grouping - equal beat length. Nevertheless, asymmetrical beat length within the measure is not an uncommon feature within Swedish folk music. It is, therefore, possible that the result is influenced by the experience of this less common style by the participants.

But the first guess from a culturally learned perspective would probably be a symmetrical grouping at beat level. Thus, the cues to an asymmetrical grouping must have been evident to

the participants, and in this case, it concerns also interonset structure. The result can thus be interpreted as indicating the strength of interonset structure as a cue to metrical grouping.

Another interesting result is the strong preference of regarding the first melodic event as the downbeat of the measure. Two different alternatives of asymmetrical grouping were among the alternative choices, one with the lowest percussive sound on the first event and the other with the lowest percussive sound on the beginning of the second group, which was the group of the longest duration. The latter metrical interpretation occurs within Scandinavian folk music, but is not as common in Sweden as in Norway.

This alternative was evidently disregarded by the participants, which can be interpreted as a preference for start on the first event given the reduced stimulus information in the current trial.

### ***Trial 8 and 9 – Strength of parallelism as a cue to primary grouping***

The test melodies in example 8 and 9 consisted of overlapping repeated sequences of different length creating an ambiguous structure. In this case, the pre-test (when the subjects had to notate the perceived grouping in a notation of the sequence) showed a preference for symmetrical grouping for most of the subjects, i.e. using the first repeated sequence as a cue to overall grouping, while a smaller group of the subjects gave responses that corresponded to the repeated sequence pattern.

The result of the current test gave the converse result with a strong preference for the grouping which concurred with the repeated sequence structure throughout the melody, disregarding consistent periodical grouping at measure level.

How can this be explained? In this case, I suspect that the means of indication of grouping structure influenced the result. Here I used articulative phrasing by means of short pauses between segments and an *accelerando-ritardando* pattern within the segments to show the implied grouping.

These means of phenomenal/structural segmentation enhance the gestalt group perception - gestalt rhythm perception rather than metrical grouping - since the lack of a constant tempo makes it harder to perceive impulse points at constant time intervals, true periodicity. One could, therefore, argue that this test does not answer the question of metrical grouping, but of gestalt rhythm grouping.

However, two of the alternatives were constructed by a segmentation of the whole sequence into equal parts and could thus easily be perceived to reflect a constant periodicity at measure level. These were considered less consistent with the melodic structure by most of the subjects in the present test.

The methodology of the pre-test, in which the subjects were to notate the grouping in a score, may also have been in favor of a symmetrical periodicity at measure level, since the sequence structure was quite hard to recognize in the written score. The influence of the initial grouping indicated by the first repeated sequence may thus have been enhanced.

Therefore, in such an ambiguous structure, one cannot rule out the possibility of the metrical grouping at measure level being concurrent with the gestalt grouping. Nor can one rule out the possibility that subjects would chose a initial grouping being withheld throughout the test melody as a conceived constant metrical grouping at measure level, while the gestalt groups are conceived to be asymmetrical. The same ambiguity between concurring symmetrical grouping and metrical grouping, on the one hand, and a perceived obstructing

asymmetrical gestalt grouping perceived to have more or less metrical significance is also apparent in trial 4.

My interpretation of this result is that there is a clear linkage between perceived gestalt grouping and metrical grouping since the favored grouping structure at measure level in both tests is concurrent with the primary sequence structure.

## **Results of the computer model of complex metrical analysis in comparison with the results of experiment 2**

### ***General discussion***

The results of the computer model regarding the test melodies in experiment 2 are listed below and are commented upon in detail, when they differ substantially from the result of the experiment.

In general, the result of the computer model and the generally most favored alternatives in the experiment strongly coincide. There are, however, some test melodies where the computer model of complex metrical analysis does not result in a metrical solution at measure level that coincides with the listener test.

This has to do with the design of the model. From the general model of melody cognition (see Chapter 3) it follows that knowledge of meter at pulse level is prior to the phrase and section analysis, because global organization is derived from local information. The general assumption regarding the nature of meter (*Assumption* no. 1) states that whenever meter is present it functions as the basic temporal grid of the music, which makes the structural prominence of events related to the metrical structure. Thus central pulse levels have to be determined before phrase and section analysis takes place.

But the general model of melody cognition also states that the conception and perception of the local structure is dependent on the conception of the global structure - the whole. It follows that the more complex a segment is, the more it is determined by the the global conception of a melody, which is fundamentally hierarchical (due to the singularity of the melody as a *gestalt*). The boundaries of more local segments thus have to conform to the boundaries of more global segments.

But where does the borderline go between local structure, which determines the global structures, and local structures, which are determined by global structures regarding meter? This is, according to this model, precisely between pulse/beats and time/measure. Pulse periodicity operates within the level of the fundamental perceptual present, the fundamental *now*, (see section 3.1) while time typically operates across fundamental perceptual presents, involving more *nows*<sup>61</sup>. There is evidently not always a clear division between pulse and time, but at the conceptual level of distinction between beats, and groups of beats this dichotomy is unambiguous. That is, whenever a metrical period is conceived as a grouping of beats it has to conform to the fundamentally hierarchical structure of grouping structure.

From this it follows that analysis of meter at measure level is dependent on phrase and section analysis and cannot be finally determined without global analysis, or more precisely, can be successively reconsidered on basis of global melodic structure.

---

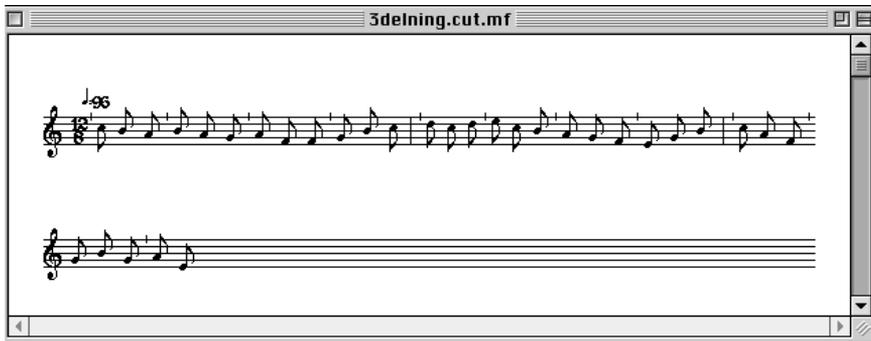
<sup>61</sup> Periodicity at measure level can thus be said to reflect composite perceptual presents. (see section 3.1)

Therefore, the metrical analysis at measure level will be preliminary in the output of the complex metrical analysis and have to be reconsidered in the light of the final analysis of phrase and section structure. The general model states that this interplay between local structure and global structure is performed continuously, while I have separated these levels of analysis in the computer model for analytical and practical reasons. This is the major drawback of the computer model at present, which though can be interpreted as demonstrating the validity of the assumptions of the general model.

Below, the graphical output of the computer model is accompanied by the numerical output of the three possible levels of metrical analysis (since this concerns quantized input); beat atom analysis, metrical grid analysis and (asymmetrical) beat group analysis.

As can be read from the results below, the metrical grid analysis sometimes suggests a basic grid at about central pulse level, which is not recognized as pulse level since it is not supported strongly enough by the beat impulse values. This level of acceptance is based on a mean value but can be altered in the computer implementation to either favor symmetrical or asymmetrical beat grouping. The results of the experiments show that some of the participants consistently preferred a continuous symmetrical primary beat grouping. This tendency would probably be even stronger if the task had been to tap the pulse.

#### **Test melody 1 and 2 - computer solutions**



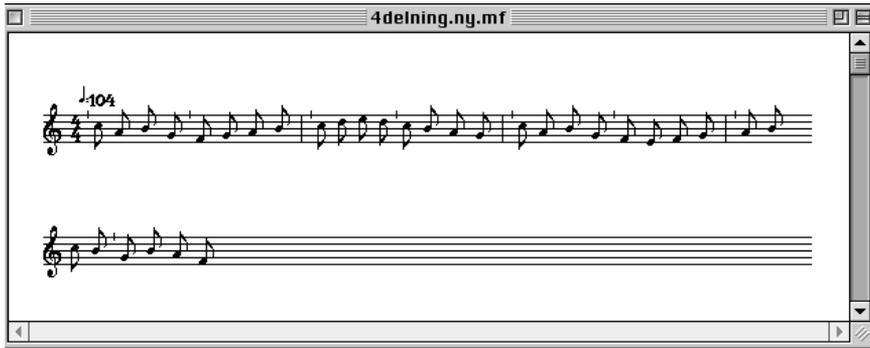


Figure 5-90. Computer model output from metrical analysis of test melody 2.

The result was also, in this case, concurring with the most favored alternative by the subjects. The alternative percussion accompaniment, however, indicated a central pulse level of 2/8.

*Beat atom analysis:* (0 512), beat atom – division at the level of 1/8

*Metrical grid analysis:* 24-beat-scope (0 8), equals 8/8 time, basic grid (0 4), equals 4/8 pulse.

### Test melody 3 – computer solution

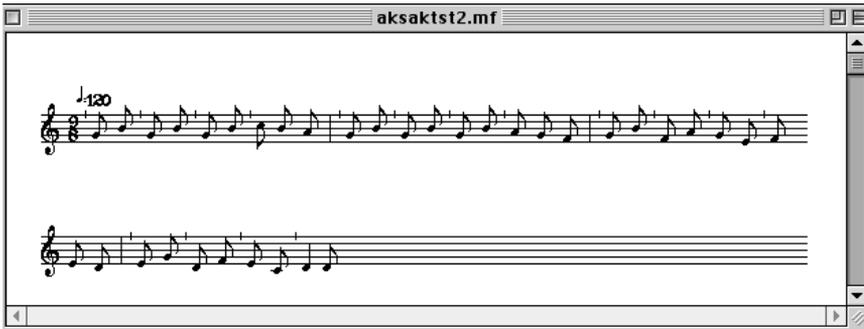


Figure 5-91. Computer model output from metrical analysis of test melody 3.

This result also reflects the most chosen alternative by the subjects.

*Beat atom analysis:* (0 512), beat atom – division at the level of 1/8, not recognized as central pulse level

*Metrical grid analysis:* 24-beat-scope (0 9), equals 9/8 time, basic grid (0 3), equals 3/8 pulse, but too inconsistent with asymmetrical grouping to be recognized.

*Asymmetrical beat group analysis:* consistent (2+2+2+3) grouping

### Test melody 4 – computer solution



Figure 5-92. Computer model output from metrical analysis of test melody 4

In this case, the metrical analysis did not come up with a result at measure level, which was in fact the object of the trial 3.

*Beat atom analysis:* (0 1024), beat atom – central pulse level at the level of 1/4. Recognized as central pulse level; further metrical analyzes were not performed. (The 4/4 timesignature was present in the input file; the measure level was not analyzed and not altered by the computer model) Here is an example in which the measure level is not analyzed by the program at this point, since it is dependent upon the phrase and section analysis. (For phrase level analysis see Chapter 6. *Metrical phrase and section analysis*, section 6.3)

### Test melody 5 – computer solution



Figure 5-93. Computer model output from metrical analysis of test melody 5.

Here the result coincided with the most chosen alternative by the subjects in trial 5. However, common notation does not allow superimposed meter to be recognized, while the computer model cannot distinguish between the alternative represented by superimposition of 2/8 and 3/8 pulse and the changed between 2/8 and 3/8 beat grouping, which is the solution which comes by default.

*Beat atom analysis:* (0 512), beat atom – division at the level of 1/8, not recognized as central pulse level

*Metrical grid analysis:* 12-beat-scope (0 6), equals 6/8 time, basic grid, changing and dominated by (0 2), equals 2/8 pulse, but too inconsistent with asymmetrical grouping to be recognized.

*Asymmetrical beat group analysis:* changing between (2+2+2) grouping and (3+3)

### Test melody 6 – computer solution

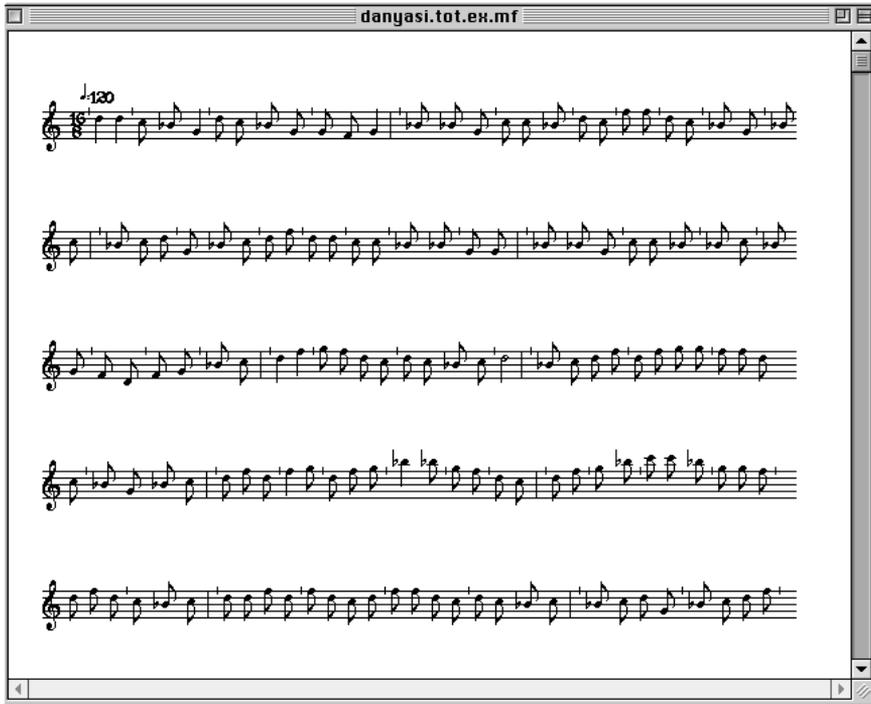


Figure 5-94. Computer model output from metrical analysis of test melody 6.

Here, the computer solution coincided with the most favored alternative regarding accentuation structure.

*Beat atom analysis:* (0 512), beat atom – division at the level of 1/8, not recognized as central pulse level

Chapter 5

*Metrical grid analysis:* 48-beat-scope (0 16), equals 16/8 time, basic grid (0 2), equals 2/8 pulse, but too inconsistent with asymmetrical grouping to be recognized.

*Asymmetrical beat group analysis:* Changing between (4+4+4+4) dominant, (3+3+2+2+2+2) dominant, (3+3+3+3+2+2), (2+2+3++3+3+3).

The favoring of fewer groups to a measure within the asymmetrical beat group analysis causes the 4 x 4/8 solution to be chosen instead of the 8 x 2/8 solution, which is probably more consistent with people's experience, since 2-grouping becomes evident in other measures.

**Test melody 7 – computer solution**



Figure 5-95. Computer model output from metrical analysis of test melody 7.

In this case, the most favored alternative of percussion accompaniment and the computer solution coincide.

*Beat atom analysis:* (0 512), beat atom – division at the level of 1/8, not recognized as central pulse level

*Metrical grid analysis:* 48-beat-scope (0 9), equals 9/8 time, basic grid (0 3), equals 3/8 pulse, but too inconsistent with asymmetrical grouping to be recognized.

*Asymmetrical beat group analysis:* Changing between (2+4+3) dominant, and (3+3+3) exception

**Test melody 8 and 9 – computer solution**



Figure 5-96. Computer model output from metrical analysis of test melody 8.

*Beat atom analysis:* (0 512), beat atom – division at the level of 1/8

Metrical grid analysis: 24-beat-scope (0 6), equals 6/8 time, basic grid (0 2), equals 2/8 pulse.

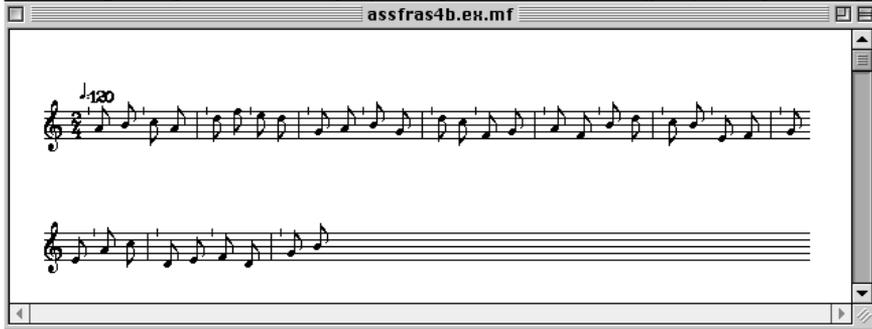


Figure 5-97. Computer model output from metrical analysis of test melody 9.

Beat atom analysis: (0 512), beat atom – division at the level of 1/8

Metrical grid analysis: 12-beat-scope (0 4), equals 4/8 time, basic grid (0 2), equals 2/8 pulse.

As can be seen, when comparing these two solutions with the alternatives in trial 8 and 9 the computer solution does not coincide with the most chosen alternatives on measure level. Observe that all alternatives in trials 8 and 9 have segments that concur with the central pulse levels determined by the computer model.

This is yet an example of the need for the phrase and section analysis to finally evaluate the metrical structure at measure level. As can be read from the outputs of the phrase and section analyzes below, they actually do concur with the grouping structure most favored by the subjects in these tests.



Figure 5-98. Test melody 8 – phrase analysis (see also section 6.3 results)

Please note that the phrase border A4 in system 2 is misplaced due to problems with the graphics. It should be placed at the next beat.

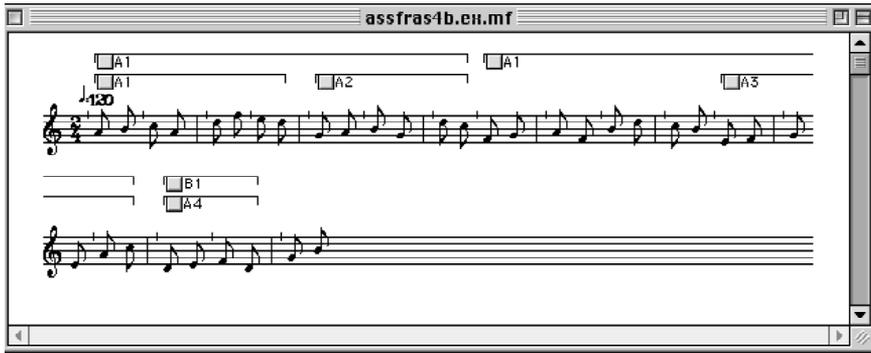


Figure 5-99. Test melody 9 – phrase analysis

The last phrase border is also in this example misplaced due to the deficiencies in the graphical interface. It should be placed after the last notes.

As can be read from these examples, these solutions of phrase structure concur with the most favored phrasings chosen by subjects in trial 8 and 9.

#### 5.2.8.4 Evaluation of complex metrical analysis – metrical grid analysis and asymmetrical beat group analysis through comparison with culturally informed notations

##### Purpose of the evaluation

The final evaluation of the complex metrical analysis involves comparisons between the results of the computer model analysis and culturally informed notations of melodies. The listener tests can be criticized for not taking into account the influence of cultural learning<sup>62</sup> or the individual acquaintance with a certain musical style within a musical culture. Therefore, as a complement to the experiments 1 and 2 referred above, I have tested the computer model against culturally informed notations.

This methodology is not without problems. The most crucial problem is the relationship between notation and conception of musical structure. To what extent does the notation of a metrical structure tell us anything about the transcribers or composers conception of the metrical structure of the melody? It is definitely not always a one-to-one relationship, mainly because notation practice is very much a matter of conventions developed to function as a guideline for performance or as a note for the memory. In some of today's Western art music styles it is common to notate a metrical structure for the purpose of readability, implying that the notated meter will not be heard or conceived, neither by listeners, musicians or the composer himself. In addition, there are problems concerning how detailed a transcription is regarding metrical structure and which conventions are used. What one transcriber would notate as regular beat length with ties between beats, can be notated by another transcriber as change of beat length (See e.g. Gurvin 1959). The very definition and meaning of

<sup>62</sup> If not a true cross-cultural test with known material is performed

orthographic symbols do vary considerably between transcribers, not to mention the conventions of orthography, that can result in e.g. the practice of placing bar lines at positions which do not conform to the perceived meter in 18<sup>th</sup> century Western music.

Still, I have found it useful to make this test mainly for two different reasons. First, what sources are better than notations (considering all problems involved) for finding well considered conceptions of metrical structure made by people who want to communicate their conception. Secondly, since the model is based on gestalt principles and cognitive premises which are assumed to be cross-cultural, it should be able to come up with results tested on different stylistic material that can be regarded as plausible.

## The test material

As mentioned above, the complex metrical analysis involving metrical grid analysis and asymmetrical beat group analysis is needed when the interonset/offset structure does not provide enough information for the beat atom analysis to come up with periodicity at central pulse level. This happens, for instance, when a melody either does not contain enough different duration categories or the interonset structure at central beat level does not support the conception of a constant beat length at central pulse level.

A typical example of the first case is the *moto perpetuo* known in Western art music and some Western folk music styles, such as some traditional dance music of the British Isles. Probably the most well known examples of the second case to Western listeners are the composite meters of Balkan music and Oriental music, where beat grouping at central pulse level regularly change within a constant or regularly changing meter at measure level. But, there are also examples from other parts of Europe such as Scandinavian instrumental folk music with regular change of beat length (e.g. *Polska* and *Springlek* styles of western Sweden and *Springar* and *Pols* styles in Norway) and some older Norwegian dance music styles (such as some *Gangar*, *Halling* and *Rull* melodies) in which the metrical beat grouping can be perceived to change as well as meter at measure level. Yet other examples are found in e.g. Indian music where beat grouping can be perceived to change within the framework of regular meter at measure level.

I have chosen a number of melodies of the styles mentioned above. I have used six movements composed by Johann Sebastian Bach, taken from his “*Violin Sonatas and Partitas*” (BWV 1001) and the *obbligato* melody from the cantata piece ‘*Jesu Bleibet meine Freude*’ (Cantata no 147, BWV 147). These are all examples of the *moto perpetuo* type, with little or no changes of interonset categories.

I have also used seven Swedish polska melodies taken from the repertoires of Gössa Anders Andersson and Spak Olof Svensson, with regularly changing beat length transcribed by myself via audio-to-MIDI conversion and quantized using the quantization method described above. These all belong to the type with regular change of beat length within a consistent meter at measure level.

Ten Norwegian *Halling/Gangar* and *Rull* melodies have also been taken from transcriptions made by e.g. Arne Björndahl, Morten Levy and myself. These are commonly known as *tredelt* (the beat divided in three) and are often notated in changing 6/8 and 9/8 or 3/8 time in these transcriptions.

Thirteen popular Balkan melodies have been randomly chosen from common notations of very different origin (Bulgaria, Macedonia and Serbia) spread amongst Balkan musicians in

Sweden.<sup>63</sup> This category is generally regarded to have regularly changing length of primary beat groups composed of a general meter at beat atom level. The metrical structure at measure level is generally persistent throughout the melody.

In addition, are also five South Indian ragas, noted by an South Indian performer (K. Shivakumar 1998, 2002), so as to reflect the changed grouping of beats within a general meter at measure level, in most cases encompassing 16 beats (*Adi tala*). The sa-re-ga notation allows the grouping conceived by the performer to appear through grouping at primary group level.

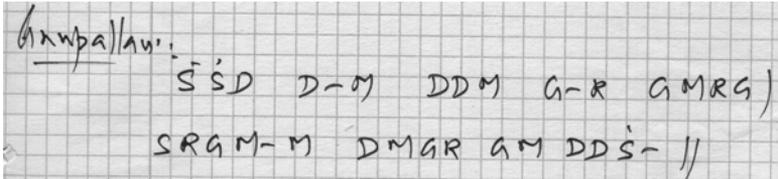


Figure 5-100. Sa-re-ga notation by K. Shivakumar. (Excerpt from “Abhogi”, Carnatic raga composition, notation for practice use)

## Results

### Summary

In the table below, the total number of beat atoms (divisions of beats to be formed into beat groups at central pulse/tactus level) and total number of notated beat groups and measures are listed. To obtain a weighted measure of the result, I have compared the result of the model with the original notation using the F score measure (Bod 2001, Höthker et al 2002).

This measure is based on the number of correctly identified positions (True Positives), the number of not identified positions (False Negatives) and the number of not correctly assigned positions (False Positives).

“In the literature, a melodic segmentation algorithm’s performance has been systematically measured using F score (Bod 2001, Höthker et al 2002)

$$F(s, s^*) = \frac{1}{1 + \frac{FN+FP}{2TP}} \in [0, 1],$$

where  $s$  and  $s^*$  are two segmentations of the same note list. Whereas TP, the number of true positives, records how many boundaries are identical in both segmentations, FP and FN, the number of false positives and false negatives, records how many boundaries are inserted in only one of the two solutions. F score is an appropriate evaluation measure because it excludes true negatives, and keeps them from misleadingly dominating the assessment. On the other hand, it makes the questionable assumption that errors at different boundary locations are independent.” (Spevak, Thom and Höthker 2002:5)

---

<sup>63</sup> Provided by Petter Landin 2003

Table 25. Results of the evaluation of metrical analysis in relation to score interpretations

Category	Notated beat groupings	Tot no of beat atoms	Ratio of correctly identified beat groups	Ratio of correctly identified measures	F score (Bod 2001, Höthker et al 2002)
J.S. Bach (7 mel.)	2+2,2+2+2, 3+3+3,4+4, 4+4+4 (4/4,6/16,3/4,9/8)	2007	92.2 %	92.7 %	0.92
Swedish <i>polskas</i> (7 mel.)	2+4+3,3+3+3 (9/8)	1161	93.5 %	99.2 %	0.93
Norwegian <i>gangar</i> etc. (10 mel.)	2+2+2, 3, 3+3 (3/8,6/8,9/8)	2397	82.8 %	89.8 %	0.84
Balkan tunes (13 mel.)	3+3,2+2+3,3+2+2,2+2+3+2+2 (6/16,6/8,7/8,11/16,11/8)	6011	93.0 %	96.9 %	0.94
South Indian ragas (5 mel.)	4+4+4+4,3+3+2+4+4,3+3+3+3+4 and num.other (16/8, 6/8)	1596	70.4 %	100 %	0.69
<b>TOTAL</b>		<b>13172</b>	<b>88.9 %</b>	<b>94.9 %</b>	<b>0.90</b>

Note that five melodies had to be removed from the test either because of erroneous results of the beat atom analysis or because the computer analysis did not go through, due to programming errors.

### Examples

The different character of the different melody types can be read from the following examples of output from the analysis.

The image shows a screenshot of a music player window titled "gminorpresto2.mf". It displays a musical score for the Presto movement from the Violin Sonata in G minor, BWV 1001. The score is written on eight staves. The first staff begins with a tempo marking of "♩ = 120". The music is in 6/16 time and G minor. The score is annotated with vertical lines and brackets, representing a metrical grid analysis. These markings indicate common 6/16 groupings on the measure level and common 2/16 groupings at the central beat level. The notation includes various rhythmic values such as eighth and sixteenth notes, and rests.

Figure 5-101. Presto from Violin Sonata G minor (BWV 1001) Computer analysis (excerpt).

In this example the metrical grid analysis provides the entire solution (common 6/16 grouping on measure level, and common 2/16 grouping at central beat level).



Figure 5-102. Polska after Gössa Anders Andersson, Orsa, Sweden. Computer analysis (excerpt)

Here the metrical grid analysis provides the result at measure level (common 9/8 grouping on measure-level, common 3/8-grouping on central beat level and an asymmetrical interpretation), while the grouping level is performed by the beat group analysis, and results in (2+4+3) and (3+3+3) grouping interpretations.

The image shows a musical score for a piece titled 'gangar' by Andres K Rysstad, Hylestad, Setesdal. The score is written on eight staves of music. The first staff is marked with a tempo of 120. The music is in 6/8 time, with various rhythmic patterns and melodic lines. The score includes a variety of note values, including eighth and sixteenth notes, and rests. The key signature is one sharp (F#). The score is transcribed by the author, and ornamentation and accompanying notes are omitted.

Figure 5-103. Excerpt of Norafjells, *gangar* played by Andres K Rysstad, Hylestad, Setesdal (transcription by the author). Ornamentation and accompanying notes omitted.

Here the output of the metrical grid analysis is changing metrical grouping on measure level. 12-beat-scope: 6/8 and 3/8 grouping and common 3/8 grouping on beat level, however too weak to be considered dominant by the program. Therefore the beat group analysis is performed, which results in alternation between (2+2+2) and (3+3) grouping. In this case, it concurs closely with the bowing pattern, which is regarded by Levy (1983/1989,III:5) to be metrically significant.

The *gangar* and *rull* examples are more complex on measure level than the previous style examples, which can be read from the model's poorer average performance on measure level when applied to this stylistic category (see table 25 above). Many of the *gangar* and *rull* tunes can be designated as **heterometric** on measure level, i.e. having irregular changes of measure length – or, interpreted in another way, weak indication of metrical grouping at measure level, but rather strong on pulse level. This quality has made it more common nowadays in Norway to designate the meter of these tunes as 3/8 meter, i.e. congruent with the beat level that equals the foot tapping, dance steps. (See Blom & Kvifte 1986, Kvifte 1983, Blom 1989)

Why then bother to find the weak indications of metrical grouping at measure level? There are two interconnected important reasons for this: 1) If we consider the metrical grouping on beat level to be able to change between 3, 2+2+2 and 3+3 grouping (as suggested by the notations I have used) the measures are the superordinate metrical level to the grouping; 2) Since the metrical structure is a presupposition for the phrase and section structural analysis, similar melodic segments must have identical metrical structure with respect to beat length and position in relation to beats or they will not be recognized



Figure 5-104. Cetvorno. Bulgarian tune. (Landin 2003) (excerpt)

In short, as soon as we start recognising the melodic content, the measure level becomes important, and this is obviously the purpose of the subsequent analysis. It is hardly even useful to try and generalize to which extent the melodic content/structure is experienced as a determinant of metric structure at measure level among musicians and dancers, though it is very obvious that the pulse levels are much articulated in this musical style.

Here, (see fig. 104), the output of the metrical grid analysis is a common 7/8 grouping on measure level, no common metrical grouping on central beat level, but an asymmetrically grouped sequence list. The beat group analysis then provides the common 3+2+2 grouping at central beat level. In this style, the grouping at measure level, which of course governs the beat grouping, is often indicated by melodic sequence structure in addition to interonset structure. A substantial part of the melodies also largely consist of tones of equal duration (*moto perpetuo* style), which makes the sequence structure even more important as an indicator of periodicity at measure level.

## *Chapter 5*

The metrical grid analysis provides common 16/8 metrical grouping at measure level, common 2/8 grouping at central beat level, but strong asymmetrical sequence indication. The beat group analysis provides a changing grouping consisting mainly of 2,3 and 4 beat atoms per composite beat group.

In this case the common measure level is in reality indicated by repetitions of long segments – sequence structure, while the metrical grouping at beat level is much more heterogeneous where almost every measure has a different duration and melodic contour pattern. It can be questioned if the primary groups are at all conceived by culturally learned listeners to be metrical groupings, or if they are conceived entirely as rhythmic gestalts without metrical implication? Is, for instance, the 3+3+2 grouping, which is so frequent (indicated even by interonset structure) in the above melody, perceived as a metrical beat grouping (with metrical qualities as in e.g. the rumba or as in a modern oriental dance music), or is it perceived as an gestalt level completely without any metrical implication?

It is, however, interesting to see that the analysis made by the beat group analysis of the computer model resembles the notated phrasing of the Indian original transcription considerably better than chance. If the computer model would have used an even more gestalt and distance oriented evaluation of segment borders, it would have performed even better. This indicates that the same structural forces apply for the asymmetrical beat grouping of e.g. the Balkan dance tunes as the Indian Classical music composition.

The image shows a screenshot of a music notation software window titled "danyasi.tot.mf". The window contains eight staves of musical notation. The first staff begins with a tempo marking of  $\text{♩} = 140$ . The notation is in a single melodic line, featuring a mix of eighth and sixteenth notes, with some accidentals (sharps and flats) and a key signature of one flat. The software interface includes a standard window title bar and a vertical scrollbar on the right side.

Figure 5-105. Raag Suddha Danyasi (after K Shivakumar). Excerpt

**Structural indications related to stylistic material**

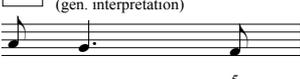
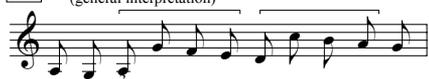
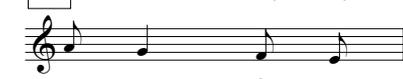
One question that can be raised concerning the result of the method on the different stylistic test material is to what extent the rules are bound to a certain style, to what extent the rules are style dependent?

One part of the complex metrical analysis, which would be assumed to return considerably different result for the different stylistic material, is the analysis of local start indications (local discontinuities and parallelism), where the different structure, would be suspected to show.

The ten most frequent local start indications for the different stylistic materials respectively as shown below. All together, there were 20 different local start indications and they are ordered in the table from the most frequent to the least frequent. The indications are labeled according to the labeling in Table 1, section 5.2.4.6.

Table 26. Frequency of occurrence (mean) of structural grouping indications in relation to different melodic styles.

Indication	Occurrence Perc. of total	Among ten most frequent in
<div style="border: 1px solid black; padding: 2px; display: inline-block;">B2C</div> NOT ONSET (REST/TIE), ONSET ON NEXT BEAT $\Rightarrow -2$ 	14.2 %	Norwegian Indian Swedish Balkan
<div style="border: 1px solid black; padding: 2px; display: inline-block;">D3G</div> ON NOTE LONGER THAN BEAT ( <i>xlong duration-class</i> ) $\Rightarrow 4$ (gen. interpretation) 	11.5 %	Norwegian Indian Swedish Balkan
<div style="border: 1px solid black; padding: 2px; display: inline-block;">Ä</div> NO CHANGE $\Rightarrow 0$ 	9.2 %	Norwegian Indian Swedish Balkan Bach
<div style="border: 1px solid black; padding: 2px; display: inline-block;">SSC</div> CHANGE OF MELODIC DIRECTION WITH ON SEQUENCE BORDER 	4.9 %	Norwegian Indian Swedish Balkan Bach

<p><b>S6D</b> CHANGE OF MELODIC DIRECTION, NOT SIGNIFICANT <math>\Rightarrow 0</math></p>  <p style="text-align: center;">0</p>	4.8 %	Norwegian Balkan Bach
<p><b>B2D</b> NOT ONSET (REST/TIE), ONSET ON NEXT BEAT <math>\Rightarrow -2</math></p>  <p style="text-align: center;">-2</p>	4.1 %	Norwegian Swedish Balkan
<p><b>N3</b> REPETITION/RETURN TO NEXT BEAT WITH OTHER INDICATION (as e.g. "leap to") <math>\Rightarrow 4</math></p>  <p style="text-align: center;">4</p>	3.7 %	Norwegian Indian Balkan
<p><b>D12</b> FIRST AFTER TIE <math>\Rightarrow 5</math> (gen. interpretation)</p>  <p style="text-align: center;">5</p>	3.4 %	Norwegian Indian Swedish Balkan
<p><b>S3B4</b> PLAIN CHANGE OF MELODIC DIRECTION OVER MORE BEATS <math>\Rightarrow 2</math> (dir-turn)</p>  <p style="text-align: center;">2</p>	3.3 %	Indian Bach Balkan
<p><b>W4</b> NOT OVERLAPPED SEQUENCE-BORDER <math>\Rightarrow 3</math> (general interpretation)</p>  <p style="text-align: center;">3</p>	3.2 %	Norwegian Swedish Balkan
<p><b>P4D</b></p>  <p style="text-align: center;">3</p>	3.1 %	Indian Bach
<p><b>D7B</b> FIRST AFTER XLONG/TIE (2 X BEAT) <math>\Rightarrow 2</math></p>  <p style="text-align: center;">2</p>	2.7 %	Balkan Bach

Chapter 5

<div style="border: 1px solid black; padding: 2px; display: inline-block;">W3D</div> WEAK SEQUENCE WITHOUT START-INDICATION AT BORDER $\Rightarrow$ 0 	2.2 %	Norwegian Indian
<div style="border: 1px solid black; padding: 2px; display: inline-block;">T5B</div> CHANGE OF STEP SIZE ON SEQUENCE BORDER $\Rightarrow$ 3 	1.5 %	Indian Bach
<div style="border: 1px solid black; padding: 2px; display: inline-block;">B2E</div> 	1.2 %	Swedish
<div style="border: 1px solid black; padding: 2px; display: inline-block;">D3F</div> 	1.2 %	Swedish
<div style="border: 1px solid black; padding: 2px; display: inline-block;">T9</div> CHANGE OF STEP SIZE MINOR $\Rightarrow$ 1 	1.1 %	Bach
<div style="border: 1px solid black; padding: 2px; display: inline-block;">T1C</div> CHANGE OF STEP SIZE, MAJOR LEAP $\Rightarrow$ 3 	0.5 %	Bach
<div style="border: 1px solid black; padding: 2px; display: inline-block;">P4A3</div> 	0.4 %	Bach

Of the total of twenty indications, fourteen are among the ten most frequent for more than one repertoire type. These fourteen indications represent by themselves ca 72 % of all indications in the entire material. They represent 96.3 % of the ten most frequent indications. Thus, one can definitely consider these fourteen as the main types of indications for the tested material.

Considering this small number in relation to the entire list of indications which have been tested, it would be interesting to test how few types of indications would actually return a result better than chance?

The nine of these, which are among the ten most frequent in more than two repertoires, represent 59 % of the total of indications. On this basis we can conclude that the vast majority of local start indications are common to the different repertoires.

Of the twenty local segment-start indications, eight consider mainly *rhythmic* qualities, while nine consider mainly *pitch*. The relative ratio of rhythmic and pitch indications respectively demonstrate the priority of rhythmic structure over pitch structure: Rhythmic structural indications (only considering the 10 most frequent) represents 38.4 % of the total of indications, while the pitch indications represents 23.4 %. Parallelism – represented by repeated sequences – represents a total of 10.7 %. (NB among the top ten indications). About 1/10 of the beat events (9.2 %) does not get any start indication value.

The most important (among top ten) start indications that apply only to one repertoire category can be related mainly to the rhythmic characteristics of the test melodies. The Bach samples differ from the others in that only melodies with almost no rhythmic differences were selected for the study (*moto perpetuo* type). That is why the rhythmic indications, which dominate regarding the other repertoires, are not as prevalent in the Bach melodies. That also explains why certain local start indications by pitch structure are among the ten most frequent in the Bach melodies while it do not display the same significance in the other repertoires.

The Swedish *polska* category differs from the others in the respect that it includes example melodies with unquantized input from live performance, with e.g. the ornamentation preserved. This creates an interonset structure with a great many events of shorter duration than a typical beat atom, which naturally will be reflected in the start indications.

Since the selection of test melodies and wealth of details in the transcriptions is very different for these two repertoires, the differences of the results in this respect can hardly be interpreted as stylistically significant.

There are certain differences regarding frequency of local segment start indications which may reflect stylistic differences that may be typical of the repertoire, such as the relatively high share of start indications based on large leaps between steps (pitch distance) in the Bach repertoire. But one would need a comparable material to make such deductions from this rating.

The overall result does instead point to the great structural commonalities between the repertoires. The same ten most frequent local segment start indications seem to be most frequent overall when the rhythmic structure is comparable.

## **Discussion**

### ***General success of the model***

The results presented in the above section strongly support the experimental hypothesis, namely that a rule-based method of analysis, based on the general assumptions of the nature of metrical perception and gestalt laws from information of event duration (IOIs) and relative pitch, is sufficient to extract a metrical structure which conforms to people's experience of metrical structure better than chance in monophonic melodies.

The relatively high score for the model (F-score 0.90), when compared with the notated meter in the different melody repertoires, indicates that the measured structural qualities are significant. Moreover, the similarity regarding the most frequent structural qualities measured

between the different repertoires (72 %) indicates that the most important structural qualities for the cognition of meter may be style-independent.

This conclusion can seem to conflict strongly with the often mentioned metrical ambiguity of melodic structures (see e.g. Cooper & Meyer 1960, Blom & Kivifte 1986, London 2002:541ff). In its most radical form, this notion implies that meter is applied on a melody (or any other musical structure) without any necessary connection to the structural qualities of the melody. One basis for this view is the common experience that it is possible for the same person to give the same melody different metrical interpretation.<sup>64</sup>

But it is obvious that this is not normally the case. Mostly people agree in their metrical conception in a cultural context, which is, for instance, evident from mostly common and metrical compatible behavior of people when they are dancing. The result of the listener experiment referred in section 5.3.4.3 demonstrates that people within a culture do agree, to a certain extent, even when confronted with culturally unfamiliar melodies. The result is shown to be in concordance with the result of the current model, which uses the structural qualities of the melody to predict a metrical structure probable to be conceived.

Both this result and that of the current evaluation based on notation of melodies, points to an adjusted version of the notion of metrical ambiguity. Even if it is possible to apply any metrical structure on a melody, people usually normally infer a metrical structure based on the structural qualities of the melody. The tested material indicates that melodies often contain sufficient information to give a metrical interpretation of limited ambiguity (with high probability to be conceived by listeners), and, interestingly enough, the same main structural qualities seem to work as signifiers for very different monophonic melody styles.

But it is important to stress that the results of the listener tests clearly demonstrate that, based on information of duration and relative pitch only, one cannot predict how every individual will conceive the metric structure. Even within a culture, among people of similar musical experience and for the same individual over time, the perception and cognition of a metrical structure of a certain piece of music can shift. Hence, the goal of the model, is to make a qualified guess about what will be the conceived as the dominant metrical structure among people.

### ***What structural elements/parameters determines the metrical structure?***

The success of the model supports the experimental hypothesis that duration / interonset structure and relative pitch structure generally is sufficient for obtaining a useful metrical structure from a monophonic melody.

This indicates that locally *strong* perceptual signifiers such as phenomenal articulation or change of timbre or culturally acknowledged signifiers such as harmony and tonality conception are *not* generally primary prerequisites for metric conception. While the melodic contour, duration and tempo qualities of the melodic structure are generally *primary*, the

---

<sup>64</sup> Another reason for the special subjectification of meter in relation to e.g. rhythmic and pitch grouping is, perhaps, the relatively general treatment of meter in common Western notation. A common notion is that bar lines and time signatures are a matter of notation and not of musical perception.

phenomenal and tonality based qualities are generally *secondary* (relating to the tested repertoires).<sup>65</sup>

### ***The ambiguity of meter indication***

As have been mentioned above, it is not possible to predict how each individual will conceive the meter of a certain melody - it is a matter of probability. Apart from the differences between individuals regarding cultural background, personality, musical experience and skills etc. the structural indications of impulse points are almost always ambiguous. How is, for example, a change of melodic direction to be interpreted in terms of impulse point? Does the top note belong to the previous direction or the new direction? Is an impulse strongest at the event from which the change takes place or at the event, which confirms that a change has taken place? Which is most defining - proximity or distance, similarity or change?

Regarding meter, it is obviously the impulse quality of the different structural features that is most important. This implies that while a long note among short notes generally can be conceived either as marking the end of something - (group border after) - or the start of something new (group border at the onset of the long note), the latter interpretation is generally favored in the metrical analysis. Analogically, the turn of the change of melodic direction is favored in relation to the after turn interpretation. The gestalt – distance driven qualities are generally suppressed in favor of the impulse qualities of structural changes.

But the fact remains that all single structural start indications have at least two interpretations and thus the total indication involving more elements is the basis of the analysis. This implies that the degree of ambiguity is inversely related to the sum of coinciding structural indications, which implies a relationship not only to the proportion of events considered but also to consistency and salience of indication. Salience of a metrical pattern relates to amount of information and consistency of information. It is time-dependent.

### ***Relative structural significance of interonset structure vs. pitch structure***

The results confirm the assumption that locally, at primary metrical levels such as beat division and beat atom level, the interonset or rhythmic structure is more important than pitch structure as a determinant of metrical structure. The general rule of about two times the importance of pitch structure (which is expressed in the values given by the local discontinuity valuation in the model, see sections 5.4.2.6, *Analysis of local discontinuities* and 5.2.4.8 *Primary grouping conditions*), seems to work quite well as an approximation.

On the other hand, the assumption that pitch dimension is dominant on a global level – the longer the period, the more events are involved – is also confirmed by the results. This is expressed by the dominance of long pitch sequences as determinant of metrical structure through melodic parallelism.

### ***The structural significance of parallelism in relation to discontinuities***

---

<sup>65</sup> Considering the notation systems of e.g. Indian and Western musics this seem to be logical, since these are focused mainly on pitch and duration while performance information such as articulation is generally notated to a much lesser degree.

Without the indication based on melodic parallelism, incorporated in the model through the analysis of sequences – repeated segments, the complex metrical analysis will perform considerably worse for e.g. the Balkan repertoire in the test, F score 0.94 with sequence valuation and 0.82 without. This means that for about one third of the melodies in the test the result is around or just above chance.

This demonstrates the significance of the sequence parameter for the performance of the method. The result of the model applied to the Balkan repertoire is very dependent on sequence analysis and the same is true for e.g. the Bach repertoire in which the analysis has to rely almost entirely on pitch information.

The result also confirms the assumed hierarchy between sequences and discontinuities, where sequences generally are superordinate to discontinuities as markers of segment starts, while discontinuities determine the phase of the sequence at the level of complex metrical analysis.

### ***The significance of parallel ‘now’s – beat scopes***

The method of evaluating the periodicities is based on the concept of a perceptual present, which can be defined in terms of duration as well as in terms of category perception. I assume that we can handle a *present* in relation to the next *present*, obtain a higher structural level where the individual *present* is generalized and combined into *composite presents* or ‘*nows*’.

The result seems to indicate that the hypothesis of parallel ‘now’ perspectives has relevance. It would not have been possible to trace changing periodicities within the range of the smallest beat scope (12 beats) – which is characteristic of the Norwegian *gangar* melodies – with the perspective of the largest beat scope (48 beats). And the large-scale periodicities of the *Camatic* melodies with 16 beat periods would not be possible to trace within a shorter beat scope than the 48 beat scope.

The need for handling different levels of periodicity simultaneously seems to indicate that we are able to trace periodicities at different parallel levels, which can be regarded as compounds of the basic ‘*now*’, here quantified as the basic 12 beat scope. The size of the basic beat scope is determined by the size of primary beat groups, which is limited to a maximum of four beat atoms and a minimum of two beat atoms depending on tempo (multiples of two and three), and the requirements of periodicity which imply that it should be possible to experience the periodicity as regulative within the beat scope. This further implies, that for a two-group to be established as a recurrent grouping, there needs to be at least one repetition and continuation implied, which requires five beats; and, for a four-beat group periodicity to be indicated, there needs to be a minimum of nine beats. Twelve equals the number of three times a four beat group, which means that eventual 2- or 3-beat based grouping can be determined at that beat scope.

### ***The relationship between different metrical levels***

The results can also be interpreted as to confirm the assumptions of structural relationship between different metrical levels, which means that more local periodicities determine the more global, providing the temporal grid for events to happen. Periodicity at division level (division and beat atom level) seems to be primary and superordinate to higher metrical levels. Periodicity at central pulse level is likewise superordinate to periodicity at more complex levels, such as measure level. But, on the other hand, e.g. the Balkan and Indian

repertoire clearly show that, when there is no isochronous beat at central pulse level, periodicity at measure level can delineate the boundaries of quasi-periodical composite beat levels.

Thus, true periodicity is superordinate to *quasi-periodicity* as metrical determinant.

The composite beat patterns of e.g. the Balkan repertoire would not have been possible to determine without tracing the periodicity at measure level. Conversely, the existence of a regular symmetrical pulse at central pulse level, provides the fundament of measure level periodicity, such as in the Bach melodies.

### ***The significance of periodicity persistence***

The assumption that the first established periodicity will disguise later periodicities seems also to be confirmed by the results. When this feature of the model is removed, the result quickly drops radically regardless of repertoire.

The strength of a periodicity does, however, seem to stand in a simple relationship to the number of periods. Even a well established periodical pattern can be changed, if there is enough contradictory indication for another periodicity. Change to a compatible periodicity or to periodicity based on another division (beat-atom period) is more easily managed than change to an incompatible period based on the same division as the previous periodicity. Accordingly, change of periodicity on a higher level is easier if the new periodicity conforms to the same basic periodicity at a lower level.

The chosen limits, which are based on the beat scope concept (the concept of parallel *noms*, see above), seem to work reasonably well.

The general model of melody perception (see Chapter 3, section 3.2) predicts that it should be possible to reconsider the metric structure of a melody during the melody. This feature of the model has been proven to be essential for the success of the model. There are numerous of examples where the most marked periodicity in the beginning of the melody is replaced by a periodicity which was obscured in the beginning.

### ***The significance of absolute tempo and pulse salience***

The model performs considerably worse when constants regarding typical tempo values are removed.

The most important constants used are the 100 ms ornament-limit, the 250 ms limit for typical beat atom value, the preferred central pulse level not below 500 ms and a maximum metrical period limit of 5,625 ´´, the latter value being a constant based on the estimated period (5-6 ´) of the perceptual present (Fraisse 1982). All of these are adapted from the experimental literature (for a summary, see e.g. London 2002:529ff).

One can interpret the tempo dependency of the model as yet another indication of the existence of typical tempo ranges for different metrical levels.

### ***Periodicity preference***

The result of the listener tests clearly demonstrates the need for a periodicity preference choice of the model. It is obvious that, even within a cultural sphere, people seem to have different levels of periodicity sensitivity and tendency towards perception of periodicity. It is also obvious from the results of the current test that the model's weighted choice sometimes

does not conform to the reference of the notation, even if the alternative represented by the notation is identified and considered by the model.

A typical example is the level of acceptance of syncopation. A ‘pulse-oriented’ person searches actively for a common, all-pervading periodicity at pulse level and can accept quite a lot of syncopation, i.e. obscured beat onsets with interfering neighboring onsets, while a person which is very *metrically ‘gestalt’ oriented*, who assumes note onsets to be on beats, may be perceiving a composite meter or no meter at all from the same stimuli.

Because of this, the model is designed to give alternative interpretations when more than one metrical choice is identified, regarding the choice between common pulse level (at tactus level) and division level regarding beat atom analysis and the choice between common pulse (at tactus level) and asymmetrical composite pulse level in the complex metrical analysis.

A typical example where the default interpretation made by the model is in conflict with the originally notated interpretation is the J.S. Bach *Courante* from the *B minor Partita* (BWV 1001), which is treated as asymmetrically grouped within a common periodicity at measure level, rather than by the common beat alternative, which is also identified by the metrical grid analysis.



Figure 5-106. *Metrical Analysis output of the Opening section of Courante from “Partita in B minor” by J.S. Bach.*

It is not entirely unlikely that someone could experience or interpret composite grouping in the manner of the analysis from this melody. However, the example reflects one of the basic limitations of the model: It cannot make a prediction that encompasses every individual or cultural preference. But it can provide plausible alternatives. In the current example the metrical grid analysis provides the notated metrical structure of the original composition, which opens the possibility for a second guess.

The result of the test presented here demonstrates that it is possible to guess better than chance for different stylistic and structurally different materials with the use of the same gestalt psychologically based rule system. In other words, what unites the styles is greater than what separates them. This could be interpreted as an indication that we possess the same basic

sensitivity to structural qualities of a melody and possesses a common sensitivity to periodicity in melodic structure, at least in cultures where the concept of melody is significant.

### ***Limitations and failures of the model***

The most important limitation of the model, regarded as a method of analysis, lies in the limitation of the predictability of metrical experience for individuals mentioned above. The individually and culturally determined preference for metricity and metrical consistency cannot be predicted by the model based on the structural qualities of melodies alone. The model is limited to identifying metrical structures that might be recognized by listeners.

Another obvious limitation is the small scope of the melody concept that is used; articulation, instrumentation and other performance factors as well as polyphony, harmony accompaniment and other structural factors are not taken into account. The reason for this is the aim to use material that lacks all of the above information and, thus, to study the extent to which the structural qualities of relative pitch and duration influence the conception of meter. But it would most certainly be possible to find a collection of melodies for which other musical parameters determine the metrical structure.

Erroneous results of the tests are generally due to the limitations regarding the basic design of the implementation of the model. The most important limitation regarding the implementation is that the different methods of analysis are run sequentially and not in parallel. According to the general model of melody perception, the perception of local structure gives rise to a conception of global structure which in turn affects the perception of local structure and so forth.

This implies that periodicity on division level, beat atom level, central pulse level and on measure level, is in reality, conceived in parallel, not analyzed sequentially. The identification of a division level periodicity affects the analysis of central pulse level, but the identification of a melodic sequence and phrase structure can also influence the analysis of central pulse level or the significance of a certain periodicity in the listeners mind (see also discussion in section 5.1.3 and Chapter 6). The step-wise design of the implementation is a matter of evaluation of the different steps in the analysis and the reciprocal relationship between those.

The typical errors of this analysis are the following:

- The beat atom analysis either gives an erroneous result or no result at all because the beat atom analysis does not use pitch information, parallelism and phrase and section structure;
- The complex metrical analysis provides erroneous results due to lack of elaborate sequence identification and phrase analysis (which is implemented only in the phrase and section analysis);

Since the periodicity analysis is highly circumstantial, an error or lack of information generally produces an output that is very different from the correct result.

A possible development of the current method of analysis, therefore, would be to change the design of the implementation as to reflect the design of the general model, involving phrase and section analysis to be performed in parallel with metrical analysis.

# 6 Metrical Phrase and Section Analysis

---

## 6.1 What are we looking for? General concepts

*“Music is now so foolish that I am amazed. Everything that is wrong is permitted, and no attention is paid to what the old generation wrote as composition.” (Samuel Scheidt 1587-1654).*

### 6.1.1 The syntactic and hierarchic structure of melody

In this study a melody is viewed as a comprehensible *gestalt* involving pitch change over time. This implies that substructures and individual events in a melody are conceived to be related to each other to form a meaningful ‘whole’, a melodic *gestalt*. The converse is a series of pitch changes over time conceived merely as a series of random tones lacking intelligible, conceivable inner structure.

The substructure of a melody is thus here regarded as syntactic, that substructures relate to each other to form a meaningful whole. The process of analyzing a melody must involve the identification of substructures of structural significance, i.e. that can be regarded as the structural components of a melody.

In this part of the analysis we are examining grouping above the level of primary grouping, which was the level in focus in the metrical analysis. The higher level of substructures that is the focus of this chapter is often described in terms of phrases, periods and sections, more large-scale structure. The term segment will be applied globally to refer to all of these substructures. The term phrase applies to levels above primary grouping (beat level) and section regards global substructures, which by themselves contain contrasting substructures of the same complexity as a melody. Section typically refers to the major elements of a movement of Western Classical composition or similar large-scale substructures of other repertoires. Phrase (and sometimes turn) applies to all other substructural levels above the primary grouping/beat level.

The general theory of melody conception (see Chapter 3, *General model*) states that the process of conceiving a melody can be characterized as a bottom-up – top-down recurring process, i.e. the perception and conception of local structure gives rise to conceptions about global structure which in turn affects the perception and conception of new information at local level and so forth. New information provides input to the global conception of the melody throughout both the listening process and afterwards, the conception of the melody being revisable all throughout the process of conceiving a melody.

This theoretical concept has five basic implications for this part of the analysis:

- (1) The substructure of a melody is syntactic; substructures relate to each other to form a meaningful whole.
- (2) Melody is thought to be conceived in local segments, where segment size is determined basically by the limitations of the perceptual present. Thus, there is always at least one substructural level to a melody.
- (3) Since the conception of the melody takes place in time and can be revised throughout the experiencing process, the method of analysis must to consider the whole melody.
- (4) For the same reason as in (2), i.e. global structure is revealed in retrospect, the global substructures determines the local substructures, which implies that the identification of more global substructures, e.g. sections, will override more local sub-structures, e.g. phrases.
- (5) The identification of global structures is determined by the analysis of local structure, which implies that local structure must be input to the analysis of global structure.

### 6.1.2 Start- and end-oriented grouping

One crucial problem concerning structural analysis of melodies is that conception of melodic structure is not formalized and socially controlled in the same way as the conception of language structure, where the function as a communicative tool is entirely dependent on common structural understanding between sender and receiver.

In music, there is usually no external, referential or precise meaning connected to the conceived structure<sup>66</sup>. In many musical activities, such as concert events, there is no need for a common experience of the structure as long as both the sender and the receiver experience the melodic structure as intelligible and syntactically meaningful, usually no control or confirmation of a common experience occurs.

Therefore, it is no wonder that people experience melodic structure in different ways. It is common that a person can experience a melodic structure as ambiguous, and further choose how to conceive the melody. (see e.g. Temperley 2001:65). Such ambiguity is perhaps even a typical feature of melodic structure in contrast to language.

But ambiguity in the experience of melodic structure does not seem to be arbitrary. One example, which is frequently used in the literature of cognition of musical structure, is the theme of Mozarts Piano Sonata in A major, (K 331).

---

<sup>66</sup> There are, however, exceptions to this general rule, such as talking drum traditions in Africa and signal systems of some herding cultures.



oriented, since similarity in a temporal context is recognized through recurrence; repetition of what is already heard which promotes identification by start.

The common music terminology includes terms such as upbeat, pickup, anacrusis etc. and reflects the structural position as something that happens before the *real* start of the segment. This means that generally a segment with or without upbeat or pickup is considered structurally equal; the upbeat/pickup is not considered to have structural significance. That is, the structural identity of a segment does not generally seem to be affected by the existence or absence of an upbeat.

For the structural analysis of melodies this is, in fact, of crucial importance, since the existence of a pickup, upbeat etc. can be subject to variation even within a melody.

The start- and end-orientated interpretation of structure can be interpreted as two complementary forces of structuring that can interact within the conception of melodic structure. Therefore both a start- and end-oriented interpretation of the melodic structure as output of the analysis is provided.

The start-oriented interpretation is, however, used as the basis for the hierarchical and syntactic analysis of the melody.

- (6) **Segment boundaries are conceptually determined either by the start or the end of the segment, start- or end oriented. They can be considered as complementary conceptions; the first focuses on impulse quality, the latter on gestalt quality and may be combined in forming a conception of segment structure.**
- (7) **Segmentation by similarity is generally start-oriented, since the similarity is recognized by the recurrence of a pattern.**
- (8) **Segmentation by similarity is generally structurally significant in metrical melodies since similarity between segments is a powerful factor for determining syntactic relationships between segments.**

### 6.1.3 The concept of sequence

In the chapter of metrical analysis (see section 5.2.4.3) I have already introduced the concept of *sequences*, which here designates a case of melodic parallelism with special significance to the current model. A sequence is defined here as two contiguous series of melodic events conceived to be similar, thereby creating a duple segmentation, which can be described briefly as segmentation by repetition.

The special status of *sequence* in relation to parallelism in general is that a *sequence* defines a temporal relationship creating contiguous segments and is hereby a much stronger organizing structure than similarity in general between non-adjacent segments. *Sequences* seem to play an important role as structural determinant in e.g. Western music, but also in e.g. different Asian musical traditions.

The relationship between grouping by similarity and by discontinuity is also discussed in the previous chapter (see section 5.2.4.3) concerning low-level grouping but it also holds for large scale grouping.

The central assumption is that because similarity involves the clustering of events, it has greater significance as a structurally determining factor the more events are involved. Discontinuity on the other hand is more influential on a local level.

- (9) Similarity has greater significance as a structurally determining factor on a global level, while discontinuity is more influential on a local level.
- (10) Sequence in the sense of segmentation by contiguous repetition is a stronger structural factor than melodic parallelism in general.

The relationship between discontinuity and sequence can generally be characterized as follows: Sequences rules over discontinuities above local level, but discontinuities determine the phase of the sequence.

This relationship has implications for the design of the method. As explained in section 6.2.3, the segmentation based on sequence and similarity structure is performed before segmentation by discontinuities.

However, from the description of melodic general similarity rating (section 6.2.3.3 *C-, X- and I-sequences*), it will be evident that discontinuity evaluation must be included in the sequence analysis because of the possibility of identification of similarity between segments is greater the more evident the segment is.

- (11) Similarity between segments is easier to recognize when segment boundaries are prominent, thus allowing more general similarity to be recognized

### 6.1.4 Implications of metrical structure for melodic structure on phrase and section level

According to the fundamental definition of meter (see section 5.1.1) the metrical structure (whenever present) functions as a basic temporal grid to which all events are perceived to relate. This has many important implications for the conception of melodic grouping structure.

Most important, melodic segments are assumed to be congruent with the beat structure, which implies that large-scale melodic segments in general can be regarded as groupings of beats.

In Chapter 5, section 5.2.4.2 *Analysis of low-level grouping*, the relationship between metrical grouping and gestural grouping is discussed. The prominence of metrical grouping as structural determinant on a primary level can be deduced from the observation that the gestural grouping of the example generally has an equal period but different phase than the metrical grouping, implying the prominence of periodicity over *gestalt*.

Consider this excerpt from “*Jesu Meine Freude Bleibet*”

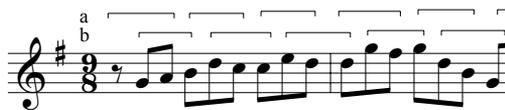


Figure 6-2. Excerpt from “*Jesu Meine Freude Bleibet*” (J.S. Bach, BWV 147, Cantata no 147, Obligato melody)

Both of the indicated groupings are likely to be conceived by listeners as primary groupings. Judging from the results of the hand selected grouping test, both interpretations exist among listeners (see section 6.3.6.3, *Experiment 3k*). While grouping *a* is based on beats,

periodic impulse points determined by e.g. sequence structure, grouping *b* is based on e.g. pitch proximity and distance. Yet grouping *b* is congruent with grouping *a*, only phase displaced.<sup>69</sup> In terms of articulation we may think of the relationship as strong – weak – weak as opposed to weak – weak – strong grouping.<sup>70</sup>

On a primary level both these interpretations may be considered as equally probable. When moving up to a higher level grouping, problems arise when considering symmetrical relationships, since judging from gestalt qualities alone would make it impossible to reveal intrinsic symmetrical and similarity relationships between segments. On the other hand the general congruence between the two groupings does not exclude the mirroring gestalt grouping on a primary level.

Thus, it can be argued that large-scale melodic segments are congruent with the beat structure in melodies with a conceived metrical structure. This means that segmentation at phrase and section level can generally be considered as a grouping of beats.

- (12) In a metrical context, melodic segments of structural significance will be congruent to the conceived central pulse level/tactus. Thus, structurally determinant groupings will primarily begin on beats.**

The prominence of metrical grouping at primary level can also be viewed from another perspective.

Consider the following constructed melody (excerpt):



The brackets mark a perfect sequence, a true repetition by pitch, which was not recognized by some people who were asked to segment this melody from a dead-pan performance (see section 6.3.6.3 Experiment 3a). The perfect sequence did not form the basis of the segmentation and structural conception of the melody.

When the melody was slightly altered, so that the sequence became concurrent with the beat structure, the sequence was recognized by the listeners (see section 6.3.6.3).



The recognition of the sequence implied a segmentation which was matched with the sequence structure. This example demonstrates that even a perfect sequence can be disguised and be without structural significance when it does not conform to the beat structure at tactus level. Conversely, a sequence which is congruent to the beat structure is likely to be recognized.

This leads to yet another consequence of the existence of metrical structure. Since periodicity at the beat and/or measure level creates a time grid for the melody, the duration of a large scale segment can be estimated and hence act as a regulating factor.

<sup>69</sup> NB, this is actually a polyphonic composition where the bass line unambiguously determines the metrical structure. However, there are many examples of similar structure in solo pieces by the same composer.

<sup>70</sup> See e.g. Cooper & Meyer (1960:6) who describes grouping in this manner, integrating metrical articulation and gestural grouping.

- (13) **The existence of a metrical grid makes it possible to conceive time spans of melodic segments, which makes duration of segments a possible structural factor. Symmetry and periodicity/good continuation thus become more forceful as determinants of grouping structure and makes structural hierarchy easier to conceive.**

Generally speaking, the strength of a time span as structural determinant is inversely related to its length and is assumed to apply for melodic segments as well as metrical units. It is assumed that the conception of segment duration is related to metrical units and the possibility of estimating segment period/duration will be determined by the limitations of the categorical and perceptual present, since longer periods of time with more elements are more difficult to estimate.

The cognitive reality of metrical units as the measure of melodic segments is denoted in common music terminology by the frequent use of bars and beats to designate structurally significant segmentations, such as *12-bar blues*, *adi tala* etc.

- (14) **As a structurally determining factor the strength of segment duration is inversely related to the length of duration in terms of metrical units. The longer the segment, the less determining its duration relationship to other segments will be.**
- (15) **And as a consequence of (12), there is a limit to the degree which symmetry between successive segments can be perceived as a structurally determining factor. This is tentatively in the proposed model set to 48 beats<sup>71</sup> based on the limitations of category conception**

This implies that, given a hierarchical structure composed of measurable units of segment duration, it is possible to estimate the duration of longer time spans by the creation of supra-metrical levels.

## 6.1.5 The means of melodic structure in the view of melodic similarity

### 6.1.5.1 Pitch structure

Since structure is based on the conception of same and different. Does something change or does it persist? We can thus regard structural significance in the view of similarity and change. There are two basic aspects of pitch similarity at the level of melodic segments: (1) similarity regarding pitch set, i.e. same pitch in both segments, and (2) similarity regarding pitch change. While similarity with regard to pitch sets can be more specific - true repetition -, similarity with respect to pitch change holds the greatest structural significance. This follows from the general concept of melody, since melodic identity is generally thought to be preserved regardless of transposition. Melodic identity is thus generally not defined by absolute pitch<sup>72</sup>.

---

<sup>71</sup> In concordance with the beat scope model, which sets the maximum periodicity perspective to eight successive basic psychological/categorical presents of 6 beats.

<sup>72</sup> This quality of melody to preserve its identity through transposition was observed by the early gestalt psychologists and is experimentally well documented in western music (see e.g. Krumhansl 1990:139)

- (16) **There are two different basic aspects of pitch structure when regarding similarity between segments, pitch set similarity and pitch change similarity, the former being based on absolute local pitch memory and the latter being based on relative pitch change.**
- (17) **While pitch set similarity is assumed to be more specific pitch change is assumed to be of greater structural significance, being a stronger determinant of melodic identity.**

But what is pitch change? And of what are the structurally significant qualities of melodic pitch structure? Pitch height is the structurally most determining quality of pitch. This consequently means that changes to a higher or lower pitch category respectively, will be generally recognized and considered different, hence structurally significant.

- (18) **The main structurally significant quality of pitch in melody is assumed to be pitch height, which implies that change between melodic pitch categories are perceived in terms of pitch height, patterns of ups and downs.**

Melodic pitch categories (MPC), (see chapter 4, section 4.1), represents the structural components that determine melodic motion and melodic contour and are assumed to be the primary structurally significant level of pitch in melodies. Consequently, two segments, which are identical with regard to melodic pitch categories, e.g. a repetition of a phrase in minor mode in major mode, are regarded identical in the analysis of melodic similarity (Cf. Dowling & Harwood 1986).

- (19) **Pitch change in melody is assumed to be perceived categorically. The level of melodic pitch categories (MPC) is hence assumed to be the primary structural level of pitch in melodies. Two segments which are identical with regard to melodic pitch categories will therefore be regarded structurally identical.**

The MPC structure also displays an intrinsic structural dichotomy between change between neighboring pitch categories and distant pitch categories. These are designated as *steps* and *leaps* respectively. The categorization in steps and leaps is considered to be structurally significant in the current model. The cognitive basis for this dichotomy is revealed in the concept of melodic pitch categories, which is determined by transition between neighboring categories. Indeed, it is the basis of the concept of scale. To produce a leap with the voice also requires a greater effort than to produce a step. It is reasonable, therefore, to assume that the categorical difference between step and leap will be recognized.

- (20) **Melodic motion, in terms of pitch-change between melodic pitch categories (MPC), is cognitively grouped into two basic classes, steps and leaps. The former type refers to motion between neighboring MPC categories. A step is assumed to be of the maximum size of a major third in this model, based on MPC systems in different traditional musical styles.**

I have found useful to make yet another distinction within the leap class, to separate between small leaps and large leaps, where large leaps generally exceed the perfect fifth, which is considered here as a limit for local melodic grouping.

- (21) **Leaps are categorized as small or large, a difference which is considered to be structurally significant. This categorization is fundamentally contextual, and also has general limits. Leaps exceeding a perfect fifth are assumed to be**

**conceived as large in this model. However, the influence of this categorization is context dependent.**

The influence of metrical structure as the primary structural level of a metrical melody has a decisive significance for the conception of pitch change in melodies. The assumption that the central pulse level is the basic temporal mapping of metrical melodies, implies that change of pitch must be conceivable between beats and even between groups of beats. In fact, the pitch change between the impulse points/starts of beats is considered as the most structurally significant level of pitch change in metrical melodies. It concerns pitch change between the articulated, structurally important events, whether or not the primary grouping is conceived to coincide with beat grouping.

Pitch change is thus regarded as possible to conceive between groups of tones, where the central beat level determines the most structurally significant level of pitch change in metrically conceived melodies.

- (22) **Pitch change is regarded as possible to conceive between groups of tones, and between articulated tones.**
- (23) **Pitch change between the initial notes of beat-groups (primary groups determined by beat period), is, along with successive tone-to-tone pitch change, the most structurally significant level of pitch change in metrically conceived melodies.**

In the metrical analysis, two global levels of pitch change were presented; global register change and global direction change (see section 5.2.4.6). These levels of melodic pitch change apply to the change of pitch between groups of elements, typically within the limits of perceptual present.

- (24) **Global pitch changes regard the changes between groups of beats and are categorized into global change of register and global change of melodic direction, typically defined within the limits of the categorial present (7±2 beats)**

The different types of pitch structure described above can be regarded as representing different levels of specificity of pitch structure, from pitch identity which is very specific to more general structural features as changes in register between groups of pitches.

- (25) **The more precise structural quality of pitch becomes, the more salient it will be in the determination of melodic structure. However, it need not reflect a difference of structural significance.**

### 6.1.5.2 Rhythmic structure

Rhythmic similarity between segments like pitch similarity may be viewed on different levels. In its most precise form, it concerns identical patterns of interonset/duration of events. In this case, however, change of duration is more structurally significant than absolute duration. This implies that, if a pattern of interonset/duration changes between events is repeated at another tempo, it will be regarded as structurally identical if the relationship to the metrical structure remains the same. In longer musical pieces, most of us have difficulties in recognizing or controlling moderate, gradual tempo changes.<sup>73</sup>

---

<sup>73</sup> The cognitive reality of this statement is evident from e.g. the frequent use of quantization in the recording industry, which reflects the low significance of perfect tempo in most musical performances.

- (26) **Patterns of event duration are the most specific level of rhythmic structure<sup>74</sup>**  
(27) **Patterns of duration changes are generally more structurally significant than patterns of absolute duration**

The fundamental rhythmic structural quality of which these patterns are comprised is the successive relationship between events of different duration/interonsets. These are typically categorized as ratios of small integers. The maximum complexity of rhythmic relationships is generally assumed to be 3:4, 4:3 in the current model when the relationships can be determined by a common divisor (see section 5.2.31, *Beat atom analysis*). The number of duration categories that we can manage is determined by the capacity of category perception.

- (28) **Patterns of interonset/duration are basically determined by successive relationships between pairs of events, the categorization into long and short durations. On a more specific level, the successive temporal relationships are conceived in terms of small integer ratios, the number of categories of duration relationship being limited to the capacity of the categorial present.**

The relationship to the metric structure is possibly even more evident with regard to rhythmic structure than to pitch structure, since both rhythmic and metric structure concern the temporal dimension of melody.

Consequently, it is assumed here that a structurally significant quality of event duration is the relationship to the beat at the central pulse/tactus level. This relationship can be expressed as a categorization of durations in durations shorter than the beat, durations at beat length and durations longer than the beat period.

- (29) **It is assumed that the categorial relationship between interonsets/ duration of events and beat length (at central pulse/tactus level) can be structurally significant.**

The use of the term rhythmical *figures* does sometimes represent this relationship. In this signification a rhythmical *figure* denotes the duration/interonset pattern within a beat, such as long-short-short pattern in one beat and no division of another. To conceive the rhythmic structure of a metrical melody as based on duration/interonset qualities of beat-groups (primary groups coinciding with beat interval) seems reasonable given the fundamental assumption of metrical structure as the primary level of structural organization.

Thus, it is assumed that rhythmical structure in metrical melodies can be conceived as patterns of durational qualities of primary groups coinciding with beat interval, as patterns of divided, undivided and tied beats.

- (30) **On the basis of (26) it is assumed that rhythmic structure in metrical melodies can be conceived as patterns of beat categories based on interonset qualities, as patterns of divided, undivided and tied beats.**  
(31) **At a more specific level, this rhythmic categorization of beats also considers the relationship within divided beats as being structurally significant, involving a categorization into the two basic categories short and long durations within the beat.**

---

<sup>74</sup>This concerns interonsets as well as offset-onsets

In analogy with pitch structure, changes of duration can also apply to groups of beats and events, labelled global rhythmic changes. This is basically described in the chapter of *Metrical structure*, section 5.2.4.6, concerning global rhythmic structural breaks. Regarding melodic segmentation, global rhythmical breaks typically involve groups of beats. They can be divided into two categories, general changes of rhythmic density and typical global rhythmic breaks. The former refers to the change of rhythmic density over a period of beat-groups, e.g. from divided beats to undivided beats. The latter typically consider interonsets which encompasses more than two beats at central pulse level or consider a series of events which disguises the pulse, causing a node of melodic motion at a global level.

**(32) Global rhythmical changes, based on rhythmical qualities of groups of beats, can be structurally significant**

The same general rule of prominence of specificity is assumed to apply for rhythmical structure as well as pitch structure. This implies that more specific structural means are more significant than less specific ones given that they concur with the primary metrical structure. But prominence is also due to the global influence of the structural change, i.e. the more beats are involved in a rhythmical change, the more prominent.

**(33) More specific rhythmic structural means rule takes precedence over more general means, given that they concur with the primary metrical structure**

**(34) Prominence of rhythmic structural means relates to the metrical impact of the change, i.e. the number of beats involved**

### 6.1.5.3 The relationship between pitch and rhythmic structure as structural determinants

Which is most structurally significant, similarity regarding pitch or regarding rhythm? This question cannot be answered without considering the context. The general assumption is that while rhythmic similarity is most structurally significant locally (within the perceptual present – temporal/categorical) pitch similarity is more structurally significant globally.

The theoretical basis for this assumption is that the fundamental two-dimensionality<sup>75</sup> regarding pitch change in time (involving both position in time and in pitch space) makes it possible to address and cognitively handle more unique events than the fundamentally one-dimensional rhythmic change (involving only position in time space).

The primacy of rhythmic sequences on the local level is a consequence of the primacy of interonset structure as structural determinant within the perceptual present (see *Metrical Analysis, assumption no 18*, section 5.1.3 and *General model*, section 3.1). This argument is based on the assumption that event onsets are more salient than event offsets and no onset and that this principle is more influential on a lower level (less contextually dependent).

**(35) Rhythmic similarity is more structurally significant on a local level (within the perceptual present), while pitch similarity is more structurally significant on a global level (the more events involved)**

---

<sup>75</sup> Pitch is in itself conceived in three (or more) dimensions, involving dimensions of absolute pitch memory octave equivalence and tonality (See e.g. Krumhansl 1990) In this case dimensions refer to pitch and time treated as singular dimensions.

## 6.1.6 Hierarchy of melodic segments

The fundamental assumption (*Assumption no 1*) of the general model of melody conception implies that there exist at least two hierarchical levels of sub-structures in a melody, the primary grouping structure and the melodic grouping at phrase level.

The possibility to conceive the duration of segments as structurally determinative, is a consequence of metrical structure and makes it possible to create and conceive melodic structures of considerably higher hierarchical complexity. Even if the existence of multiple hierarchical levels seems to be most prevalent in polyphonic music, it is not uncommon with more than one hierarchical level even in monophonic melodies.

The congruence of metrical structure, which constitutes the basis of the phrase and section structure, makes it reasonable to assume that hierarchy of segments is congruent, i.e. segment boundaries at a higher level are always found at a lower levels.

- (36) **Melodic structure in melodies with a metrical structure is typically hierarchical. It is assumed that it is possible to conceive multiple hierarchical levels of melodic segments even in a monophonic melody when metrical structure exists.**

## 6.1.7 General grouping principles

The model is generally built upon the principles of low-level grouping, described in section 5.2.4.3.

Even though these principles of grouping, are described in relation to primary grouping they are assumed to apply generally for melodic grouping. They are in summary:

### Primary grouping principles

- (37) **Grouping by similarity, continuity and proximity – 'sameness':  
Similar and close events tend to be grouped together**
- (38) **Grouping by discontinuity, dissimilarity and distance – 'difference'. Change / difference indicates group boundaries**

### Secondary grouping principles

- (39) **Grouping by good continuation / constancy. Group size, group start etc., which is coherent with previous grouping is prominent and can be implicative**
- (40) **Grouping by symmetry. Symmetrical grouping is more prominent than asymmetrical grouping and can be implicative**

### Tertiary grouping principles

- (41) **Grouping by perceptual prominence / foreground, i.e. articulation, impulse and gravity: Grouping between most articulated events is prominent**
- (42) **Grouping by integrity/'prägnanz' / contrast:  
Discrete grouping is prominent. Groups with a discernible, salient, unique inner structure or groups which are discernible/salient/discrete in the context are prominent.**

The primary grouping principles relate chiefly to the creation of groupings; the secondary concerns both the choice between different possible groupings and implied grouping. The ternary grouping principles have significance for the choice between different possible groupings.

## *Chapter 6*

This implies that a hierarchical structure assembled from measurable units of segment duration, will make it possible to estimate the duration of longer time spans, by the creation of supra metrical levels.

Outline of the MODUS implementation of metricral phrase and section analysis

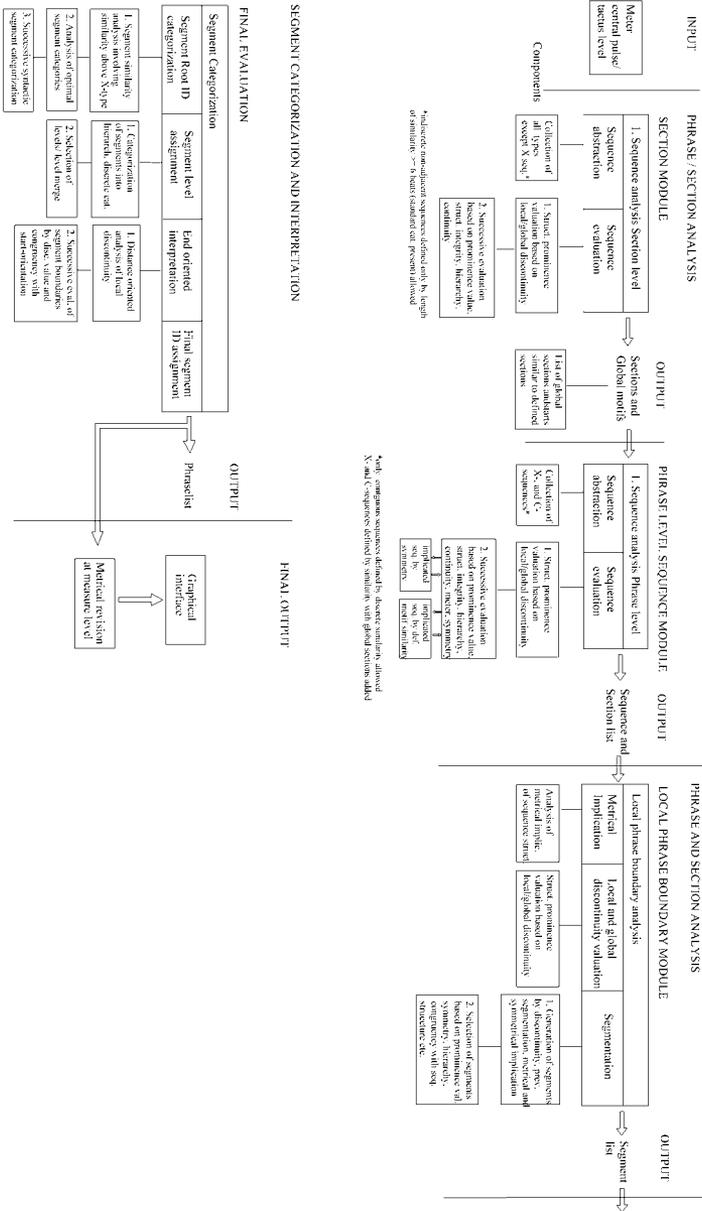


Figure 6-3. Outline of the MODUS implementation of the Metricral Phrase and Section Analysis

## 6.2 The design of the metrical phrase and section analysis

### 6.2.1 Overview

The metrical phrase and section analysis is generally performed in three steps (see fig. 3 above):

(1) First, segments distinguished by similar content – repeated sections and sequences – are identified and evaluated. This analysis is first performed on the section level and then on the phrase level, since global similar segments override local according to the general model (see Chapter 3).

(2) Then, segmentation based on structural discontinuity is performed

(3) Finally, hierarchical levels of segments are evaluated and the syntactic relationships between the segments are determined mainly based on melodic content.

Secondary grouping based on symmetry and good continuation is involved both in the sequence analysis and the discontinuity analysis as well as in the final evaluation.

The output of the model is a segmentation of a melody with up to nine simultaneous hierarchical levels<sup>76</sup>, with a start- and end-oriented interpretation of segment boundaries respectively.

The different stages of the analysis will be classified below.

### 6.2.2 The relationship to beat structure – superimposed-meter-alert

The metrical phrase and section analysis is, as mentioned above, based on the beat groups at the central pulse/tactus level. The basis of the model is that metrical structure, when it exists, is the fundamental level of organization since melodic segments primarily concern the grouping of primary groups at beat level – beat-groups (see section 6.1.5.1). Primary grouping is assumed to be congruent with beat grouping.

*Beat-groups* are thus generally undividable, atomic units, from a structural point of view, on which higher order grouping is based. However, the central pulse/tactus level is not always unambiguous. As previously noted regarding the complexity of pulse perception (Chapter 5, *Assumption* no. 12), the regulative pulses can sometimes be experienced as a complex of superimposed pulses. This has led to the need for a classification of metrical structures into two different types.

A typical example of the two different metrical types, taken from the Swedish instrumental folk music repertoire, is the metrical difference between a typical *polska* and a typical waltz:

---

<sup>76</sup> The graphical interface of the implementation of the model allows currently three levels to be displayed.

The figure consists of two musical staves. The upper staff is in 3/4 time, marked with a tempo of 108. It features a melody with eighth and sixteenth notes. Dots are placed below the notes to indicate the central pulse level. The lower staff is in 3/4 time, marked with a tempo of 152. It features a melody with quarter and eighth notes. Dots are placed below the notes to indicate the central pulse level.

Figure 6-4. Typical *polska* metrical structure (displayed on the upper system) compared to typical waltz regarding rhythmical and metrical structure. Beats at central pulse level marked by dots below the notation.

The central pulse/tactus level in the *polska* example is quite unambiguous. It meets the categorical and temporal conditions of a regulative central pulse in which the beat incorporates a small number of events of shorter duration and divides longer events into a small number of beats, hence falling in to the center of the category spectrum. In addition the tempo is close to the tempo range of greatest pulse salience. (see section 3.1 and Chapter 5).

In the waltz, the designated central pulse (quarter notes) does not meet all these conditions. First, it incorporates very few events of shorter duration per beat. Instead the events of longest duration are divided into a few too many beats to be perfect as a regulative pulse: a maximum division of events in 6 beats and of the beat in 2 is very common in a waltz of this type. In addition, the tempo is a little bit too fast to fit with the tempo range at the peak of pulse salience. Thus, the quarter note level seems to be at one end of the area of central pulse level. On the other hand, the pulse level which does meet the conditions of central pulse level (the dotted half note level) incorporates too many events of shorter duration and divides the majority of longer events into too few beats; a maximum division of events in 3 beats and of the beat in 6 events is quite common. The tempo of this pulse (dotted half note level) is 50 beats per minute (1200 ms), which is a bit too slow to fit in with the peak of pulse salience. At the same time this metrical layer does not always meet the grouping conditions of a composite metrical level – the measure level – since it does not always group any events, which makes the notation in 3/4 a bit dubious.

Still, none of them can be omitted in the typical metrical experience of such a melody. For example some performers tap their feet on the slow pulse, other on the faster pulse, still others change between the one and the other. In some traditions, as in some French folk music traditions<sup>77</sup>, two levels can be simultaneously indicated by foot tapping.

My interpretation views this as an example of a complex pulse structure, with two complementary pulses acting together as regulative pulse complex. This concept has implications for the melodic segmentation analysis, since the measure level (3/4) which meets the condition of a subgroup as super beat level in the case of the *polska* will not be sufficient

<sup>77</sup> Traditional performance of the Bourré

as supra beat level in the *waltz* example. Hence, we will need at least two measures in the waltz to form a phrase.

Consequently, metrical structures such as the waltz, where the notated beat level is close to division level, are classified as *super-imposed-meter-alert* and handled differently from more unambiguous metrical structures, assuming the possibility of a super-imposed pulse.

## 6.2.3 The sequence analysis – Segmentation by melodic parallelism

### 6.2.3.1 Typology of melodic similarity

The sequence analysis is divided into two different steps: Global sections of similar melodic content are evaluated first and are followed by melodic segments on phrase level using the global sections as input.

Sequences are identified through valuation of melodic similarity according to different structural properties, described in section 6.15, *the means of melodic structure*. This implies a categorization of different types of melodic similarity, which are reflected in a classification of different types of sequences, each with different level of similarity.

Similarity is basically valued regarding pitch content, pitch contour and interonset/duration patterns. Similarity within every dimension is valued from very general to more specific similarity.

Observe that the similarity evaluation concerns the similarity between *beats*, since segments such as sequences are considered to be conceived as groups of beats, given that the metrical structure at beat level is structurally fundamental.

This evaluation of similarity is a central concept in the model. Therefore, the typology of sequence similarity will be described below. The designations of the sequence types, NIL- T-, R-, C-, O- etc., which originate from the implementation of the model, are employed in the description to make the explanation of the classification legible and easier to follow.

### 6.2.3.2 Classes of sequences

The different classes of sequences are presented briefly below. A more elaborate explanation of the classification of similarity will follow.<sup>78</sup>

1. **NIL-sequences** are defined by successive pitch similarity, including similarity regarding pitch content (absolute pitch) and similarity regarding pitch contour in terms of melodic pitch categories. The identification is based on a classification of each beat.

NIL-sequences are defined by successive similarity from the first event in the group. The similarity of a NIL-sequence is categorized into different types and levels representing structural specificity of the sequence. Note that level 4 and 5 similarity generally are treated as being of equal structural prominence in the analysis.

---

<sup>78</sup> N.B. The sequence types are designated by a code, e.g. NIL 3 or R 4, where the first item refers to type of sequence, while the second item refers to the level of structural implication within the sequence type (see explanations.)

- **General melodic pitch contour similarity, level 3.** Designates similarity in terms of pitch distance (MPC distance) and melodic direction (melodic contour) between single beats and between pairs of beats, with considerable flexibility regarding MPC distance and melodic direction within beats.
  - **Specific melodic pitch contour similarity, level 4.** Designates similarity in terms of MPC distance and melodic direction (melodic contour), with certain flexibility regarding MPC distance.
  - **General pitch set similarity, level 3.** Designates similarity regarding pitch content (pitch set) in terms of highest/lowest pitches of successive beat groups
  - **Specific pitch set similarity, level 5.** Designates more specific similarity in terms of pitch content (pitch set) between individual beats, partly including successive order constraints.
2. **R-sequences** (reversed similarity) and **O-sequences** (omitted first beat) are like NIL sequences based on successive pitch-set and/or contour similarity, and concern the same levels of similarity. While R-sequences are based on similarity at the end of the sequence O-sequences are based on successive similarity from the second beat of the sequence, allowing start variation
  3. **I-sequences** (imperative similarity) and **C-sequences** (general pitch contour similarity) designate sequences with the most general similarity regarding melodic contour. These do not require successive similarity on the same level between individual beats, and can be regarded as representing the more general pitch similarity on levels 1 and 2. Successive similarity is required, but exceptions are allowed, and similarity can vary between very specific to more general. Sequences are hence identified based on total similarity value, global similarity and based on how discrete the sequence boundaries are in relation to the content of the sequence. I-sequences require less successive similarity but more marked sequence boundaries than C-sequences.
  4. **X-sequences:** (global similarity) are based on similarity with regards to global discrete changes such as global change of melodic direction, of register and global rhythmic breaks. They require discrete sequence boundaries in relation to the content of the sequence. They represent the most general similarity, of the least structural significance, including e.g. sequences based on inversion of melodic direction etc.
  5. **T-sequences** are based entirely on successive rhythmic similarity between beats. Like NIL-sequences the type of similarity is classified into different 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 beats into divided, undivided, beats with onsets that exceed the beat length and beats with no onset at beat start. This classification, called *dur-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 as different as from beats divided into three events etc., but similar within the category regardless of actual pattern of duration within the beat.

- **level 3** concern actual rhythmic similarity in terms of patterns of duration between *beat-groups*<sup>79</sup>
- **level 6** is employed only when there are beats of different length in the structure and concerns similarity in terms of beat length

### 6.2.3.3 General classification of pitch similarity – NIL –sequences

#### Melodic pitch contour similarity – NIL-sequences

Similarity, regarding melodic pitch contour, is measured in terms of changes between melodic pitch categories (MPC, See Chapter 4), based on beat-groups. The structural factors which determine melodic pitch contour similarity involve the MPC shift between the initial MPC of one beat to the next, the MPC shifts within the beat with regards to rhythmic position within the beat, and a classification of these MPC shifts into repetition, steps, small leaps and large leaps.

The assumptions which govern the similarity rating are implications of the fundamental assumptions of the determinants of pitch structure presented in section 6.1.5.1. These are:

- The most structurally significant of pitch changes on the local level are the pitch changes between the initial tones of beats
- The relationship of pitch change to metrical position is structurally significant, but pitch changes within beats are less structurally significant and can be generalized. Thus pitch changes within beat-groups where similar pitch change are employed at the concurrent rhythmic positions, can be regarded as similar, in spite of differences regarding pitch change. If the beat is divided in the same number of pitch events, the precise rhythmic position of the pitch shift can be disregarded.
- Pitch shifts that differ only with regards to steps or within the category of large leaps can be generalized as similar.

The similarity rating thus involves a generalization of successive melodic pitch contour similarity, which can be viewed in the following examples.



Figure 6-5. The beat-groups a1-a2 exhibit perfect MPC similarity. The pitch distances differ only within melodic pitch categories (MPCs) and the inter-beat distance is equal.

---

<sup>79</sup> N.B. The concept *beat-group* may seem confusing, but refers to a primary group delimited by a beat period as opposed to *grouping of beats* which refers to grouping above beat level (See section 6.1.5.1)



Figure 6-6. The beat-groups a1– a2 are considered similar (structurally equal), because of minor differences regarding MPC distance within categories leaps and steps in relation to the pitch change between beats.



Figure 6-7. The beat-groups a3-a4 are considered similar (structurally equal), due to equal successive pitch shifts at concurrent rhythmical positions.

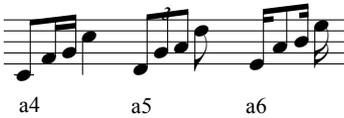


Figure 6-8. The beat-groups a4-a5-a6 are considered similar (structurally equal), because of same number of MPCs at the beats and equal change of step-size and melodic direction to the start of the next beat.

The following example demonstrates the level of generalization of the pitch similarity rating which is the fundament of the NIL-sequence analysis. Each and every of the a-labelled beats and b-labelled beats are regarded pair-wise successively similar/equal.



It is assumed that listeners will conceive the similarity regarding melodic pitch contour between the a - b pairs to be structurally significant and, thus create sub-structural segment level of two beats per segment.<sup>80</sup> Melodic pitch contour similarity is measured on two levels, specific and general (level 4 and 3).

<sup>80</sup> This assumption is supported e.g. by traditional means of variation through *diminution* or *extemporation*. An interesting example can be found in Quantz (1752/1966:141 quoted in Narmour 1999:447) which strongly supports the concept of the beat-group integrity and might well have served as an example of the similarity rules given above.

Chapter 6

Specific (level 4) similarity requires, in addition, perfect MPC similarity regarding the initial pitch events of the first and the second beat. (inter-beat distance)

General similarity (level 3), on the other hand, accepts even more general similarity regarding pitch contour than what is demonstrated in the above examples. This concerns mainly two aspects, similarity between pairs of beats and variation with regard to the first beat of the sequence. (Even more general similarity regarding melodic pitch contour is actually allowed, but will be covered under type C).

Similarity between pairs of beats is implemented by treating two beats as one. This allows greater variation with regard to the MPC change between the initial MPCs of the individual beats, since more variation regarding distance and direction is allowed within beats.

Variation with respect to the first beat of sequence, allows two different aspects of start variation: 1) difference regarding the initial MPC of the first beat, when the MPC changes from the second event of the first beat and 2) all to the next beat is equal and different order but correspondent set of MPCs on the first beat of the sequence in relation to the first MPC of the next beat.

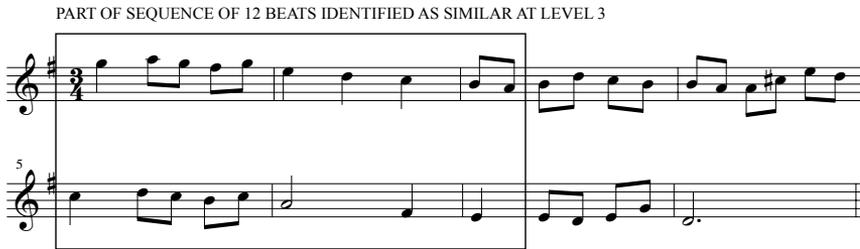


Figure 6-9. General similarity regarding melodic pitch contour (within NIL-sequences) – level 3 similarity

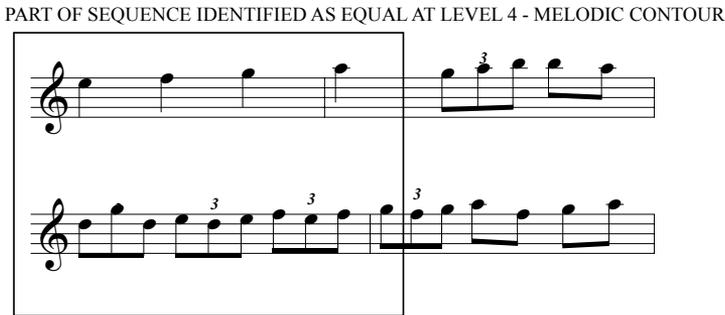


Figure 6-10. Specific similarity regarding melodic pitch contour – level 4 similarity (inter-beat distance and direction preserved)

**Pitch set similarity – specific and general – NIL-sequences**

The level 5 similarity classification apply to pitch set similarity between beat-groups. The same structural hierarchy applies for pitch set similarity as regarding melodic pitch contour, that the initial pitches of every beat are more structurally significant than changes within the beat. One difference in relation to melodic pitch contour similarity is that order between pitches is less significant regarding pitch set similarity (see below). However, the assumption that the position of pitches in relation to the metric structure is significant applies for pitch set similarity as well.

The similarity rating thus involves a generalization of successive melodic pitch similarity, which can be viewed in the following examples, these exhibit the general exceptions from perfect successive pitch similarity on level 5.



*Figure 6-11. Equal pitches on rhythmical correspondent positions and equal or only minor difference between the initial pitches of the next beat. (level 5)*



*Figure 6-12. Equal initial pitches on both the present beat and the next. (level 5)*



*Figure 6-13. Equal pitch set, different order, but identical initial pitch at next beat. (level 5)*



*Figure 6-14. Start-variation. Initial pitch different, but equal succeeding pitches at concurrent positions and identical initial pitch of next beat. (level 5)*

The more general pitch set similarity, concerns similarity in terms of identical highest or lowest pitch over pairs of beats - identical pitch set boundaries. This more general pitch set similarity is classified as level 3 similarity in the analysis.





### R-sequences – Similarity at end of sequence

R-sequences (reversed similarity) are designated by successive similarity regarding pitch contour or pitch set at the end of sequence, while the beginning of the sequence is different. Thus they mirror the NIL sequences and can be regarded as end-oriented interpretations of the same similarity structure as determines the correspondent NIL sequence.

Figure 6-19. Relationship between a R-sequence and the corresponding NIL-sequence. Sequence with 50 % similarity.

R-sequences are generally not considered as structurally significant as NIL-sequences, due to signal quality of sequences identified by start similarity within a metrical context.

Figure 6-20. Hierarchical sequence structure identified by end similarity on level 5, R-sequence of eight beats with four beats of equality, (code (0 8 R 5 4) <start> <length> <type> <level> <equal length>) and R-sequence of four beats with two beats of equality. (Each pair of sequences is marked by a letter, which designates identified similarity within the hierarchical level and a number which designates variation within the group. Shaded areas represent identified similarity)

Yet another example will be provided below, where the similarity becomes more evident at the end the sequence. The very general similarity of the beginning is not sufficient to identify the sequence.

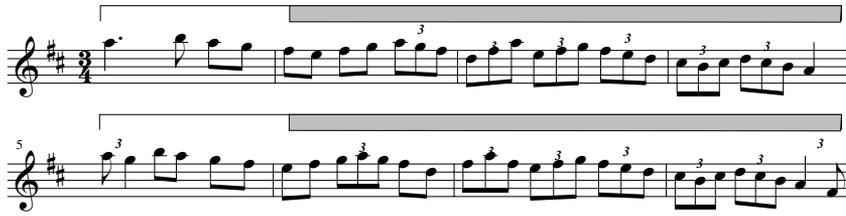


Figure 6-21. Excerpt from *Polska* played by Per Danielsson, Swedish tradition music (From “*Svenska Låtar*” V, no 13). The sequence is identified as a R-sequence, with nine out of twelve beats similar at level 3 (pitch set similarity).

### O-sequences – Start variation by start – end overlap

A specific structural feature of many melodies in different traditions at least within European music<sup>81</sup> is variation of the first beat of the structural repetition. This is sometimes connected with what can be interpreted as structural ambiguity (/confusion) concerning the first beat of the sequence which can be conceived as the ending of the previous segment at the same time as the beginning of the new segment. Such structural ambiguity can be interpreted as creating a linkage between the segments, or a melodic motion or *drive*, through gestalt grouping overlapping the beat-grouping. When start-end-overlapping segments and segments with no overlap alternate within the same melody, it is crucial for the identification of the segment structure to identify possible start-end-overlaps. Hence, O-sequences (start-overlapped) are created which mirror the NIL-sequences, but starting one beat earlier.

Consider the example below, taken from “*Jesu Meine Freude Bleibet*” (J.S. Bach, Cantata no 147, Obligato melody). The first beat of the sequence, marked by the brackets, can be conceived as the last beat of the previous sequence. The first interpretation is supported by the periodical structure of the melody with starts at corresponding positions, while the gestalt qualities favor the second interpretation due to the distance which enhances a conception of group start after the first note of the beat. Some Western listeners prefer a segment boundary at precisely this position, as a pickup to the ‘real’ start at the second beat. (see section 6.3.6.2, Experiment 3)<sup>82</sup>

<sup>81</sup> Western classical, Scandinavian folk, Balkan folk and others

<sup>82</sup> Still, most of the listeners in my test seem to perceive the start of the first beat as being the point of maximum structural gravity in this piece.



Figure 6-22. Sequence identified as equal from the second beat-group in the sequence. O-type similarity – sequence. Level 3 pitch contour similarity of five beats out of a total of nine beats. The portions rated similar are shadowed. Different possible segmentation regarding start – end overlap are marked by overlapping brackets.

But start-variation is not always connected to start-end-overlap. Below are two examples from different repertoires, with no evident structural start-end-overlap ambiguity.



Figure 6-23. Waltz played by Per Danielsson, Mörsil, Jämtland, Sweden (From Svenska Låtar, V, no 18), first and second start within sequence, perfect similarity from second beat



Figure 6-24. Cetvorno (Traditional Balkan tune, Bulgaria), first and second start within sequence, perfect similarity from second beat.<sup>83</sup>

<sup>83</sup> This would, however, be identified as a NIL sequence as well.

### C-sequences – Non-successive melodic pitch contour similarity on a general level

Melodic pitch contour in metrical melodies can, according to the general assumptions (see section 6.1.5.1, *Assumptions* no. 22-23) be conceived as change of pitch between beats. The general assumption behind this view is that the pitches of groups of tones can be summarized and thus pitch can be conceived to change between groups of tones.(see section 3.3)

This leads to the question of how pitch change, and specifically MPC change, between groups of tones is determined?

In sequences determined by similarity regarding melodic pitch contour (NIL-sequences, level 4/3), described above, the order between MPCs within beat were generally significant, with some limited exceptions. For sequences determined by pitch set similarity (NIL-sequences, level 5/3), the order of pitches within the beat-groups were assumed to be less significant, on the other hand demanding the repetition of the actual pitches.

A more general view of pitch change between beats is to regard the pitch change as basically a change of MPC set between beats, a change of register. Then each beat-group is defined by the total MPC set of that beat-group (regardless of order) and the change is a matter of intersection between the MPCs at one beat and the following.

This is the core of the more general analysis of similarity of melodic pitch contour, C-sequence, implying a categorization of pitch change from beat to beat into categories up, down and same.

The most important structural qualities of register are the limits, the pitch register of the beat-group hence determined by the lowest and highest MPC within the beat MPC set. The description of pitch qualities of beat-groups must, however, be somewhat more elaborate. The structural significance of the transition between the initial pitches of two beats nor, the relative frequency of different MPCs within a beat cannot be ruled out entirely.

Based on the assumption of change of register between beats, when does a change of pitch takes place? **It is herein assumed in this model that a change takes place when the total MPC register that differs between the two beats is larger than the MPC register in common.** This means that the register overlap between the two beats should be less than 2/3. (Cf. the *categorical proportion rule* section.3.1)

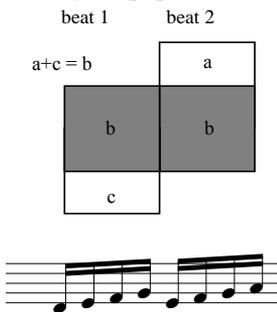


Figure 6-25. Illustration of the register overlap limit for pitch shift between beat-groups. If pitch register overlap is  $\leq 2/3$  it will be regarded as a change of pitch, otherwise it will be regarded as a repetition.

The other factors, frequency of occurrence and initial transition, are implemented through a relative MPC mean index, with the initial MPC of the beat counted twice.

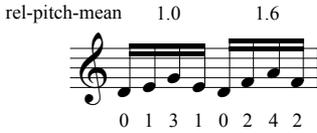


Figure 6-26. Quantification of pitch change based on mean pitch. The mean of the relative MPC values with the first MPC of each beat counted twice. For the first beat:  $(0 + 0 + 1 + 3 + 1) / 5$ . Approved change ( $> 1/3$  difference).

The general view behind the evaluation of the most general similarity of melodic pitch contour is that the search for similarity does not have to be limited to perfect contiguous similarity on the same level. We may very well be able to determine similarity based on non-contiguous similar elements represented at different levels of structural significance.

Thus, the similarity evaluation of C-sequences is based on a similarity rating of the corresponding beats of the sequence, where the beats exhibit similarity at different levels or no similarity at all. The similarity rating extends from very general similarity of registral pitch change between super-beats consisting of 2-4 beats at tactus level, giving low similarity value, to specific similarity of melodic pitch contour (as in NIL sequences, level 4), giving high similarity rating. The final evaluation is based on the total value, which allows sequences with generally low general similarity as well as sequences of shorter, but more specific similarity.

A basic assumption (see *Assumption no 10*) is that more general melodic similarity is more likely to be recognized when the segment boundaries are more evident from discontinuity. Thus, a basic requirement for C-sequences is that either the mid-sequence boundary is marked by global discontinuity (see Chapter 5, section 5.2.4.6, *Global pitch changes and global rhythmical changes*) or when both starts exhibit strong local discontinuity on a comparable level. Another constraint is that global discontinuities must conform also within a C-sequence.

Let us look at a few typical examples of C-sequence similarity, general similarity of melodic pitch contour.

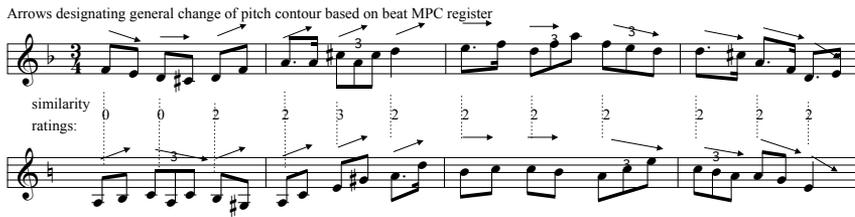


Figure 6-27. Excerpt from *Polska* after Bengt Bixo and Per Danielsson. Same melody, different versions (the second transposed)

In this example the general melodic contour similarity rating is based mainly on inter-beat register change (described above). Only one of the ratings is considered to be on a more specific level, giving rating 3. Note that the beginning is not regarded as similar.

Chapter 6

Yet another example shows the similarity rating based on similarity regarding pitch change between groups of beats.



Figure 6-28. First section of Bridal March played by Per Danielsson (SvL 5 Jämtland/Härjedalen no 21)

This melody has very strong similarity between the first 8 bars after the pickup and the next eight bars, which is generally considered a repetition, i.e. a typical sequence. But there is also a more general melodic similarity considering the first two bars after the pickup and the subsequent two bars, which could be said to have a melodic turn downwards in common beginning at the second bar of each sequence. They also have a raise of general pitch to the subsequent measure in common. Yet if we are looking at the beat-to-beat level, there is little that tells that these four measures have much in common.

Here is an example of when the very general similarity between super-beats can be structurally significant. If we look at groups of two beats at a time and consider the change between them, the similarity become apparent.

SIMILARITY OF SUPER-BEAT PITCH CONTOUR

Similarity of pitch change between groups of 2 beats indicated by arrows at the middle of bars above upper system  
Similarity rating of sequence of first four beats of Bridal March (SvL 5 no 21b)

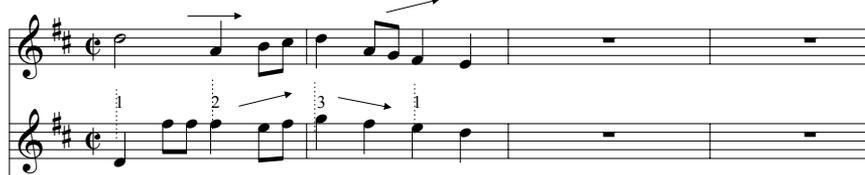


Figure 6-29. Similarity rating between first two measures and the subsequent two measures approved as C-sequence of four beats.

The value 1 represents the most general level of similarity. As can be viewed above the similarity level changes within the sequence, being most obvious at the significant turn of the second measure.

To identify the substructure at this level this similarity has to be considered.

The more striking similarity between first and second repetition can be viewed below:

Similarity rating of first and second repetition of Bridal March no 21b

Figure 6-30. Similarity rating of first and second repetition of Bridal March no. 21b. The repetitions are displayed on parallel systems.

The start difference will be identified as start variation of similarity at the NIL –sequence level.

This melody has been noted by the folk music collector Olof Andersson (see Andersson 1926:16) to be a variation of another melody, “*Venus, Minerva*”, by Swedish poet and troubadour Carl Michael Bellman. To identify this appointed general similarity we need to consider change of MPC register between pairs of beats.

Similarity of pitch change between groups of 2 beats indicated by arrows at the middle of bars above upper system

Figure 6-31. Similarity rating of Bridal March no. 21b and appointed parallel melody by C.M. Bellman.

## I- and X-sequences – Segmentation by general similarity and global discontinuities

The most general type of sequences are the I- and X-sequences. These types are, like C-sequences, determined partly by discontinuities at sequence boundaries. However, in these cases, the start and mid-position of the sequence have to be marked by strong discontinuities at a global level.

I-sequences demand that both start- and mid-boundaries are marked by both rhythmic and register changes at a global level. I-sequences also require that global changes within the sequence halves are less marked, which means the structural change at sequence boundaries must be discrete in relation to the sequence content. In addition they require compatible global register change equal beat durations as well as at least six beats of average similarity value higher than 1 (see above), which means more than super-beat similarity. The general concept is that general motif similarity will be recognized when the segment boundaries are discrete in relation to the content of the segments.

Below is an example of identification of sections/global phrases by I-sequence similarity. It is the melody line of *Boléro* by M. Ravel. The similarity between each contiguous pair of segments is very general, and is more obvious only at the beginning of the sequence. The first

Chapter 6

of these pairs are not of the same length in terms of beats. The I –sequences listed below allows the identification of the motif structure of the melody. Major segments can be regarded as variations of the same melodic idea.

COMMON SHAPE OF MELODIC CONTOUR:

The figure shows five staves of music from Bolero by Maurice Ravel. Each staff has a vertical dashed line indicating a start position. Below each staff is a similarity rating sequence. The ratings are as follows:

- Staff 1: Similarity rating for sequence (0 24 1 3 7), start-position beat 0, sequence-length 24 beats, approved similarity 7 beats. Rating: 1 1 1 2 2 0 1 0 0 0 0
- Staff 2: Similarity rating for sequence (24 30 1 3 7), start-position beat 24, sequence-length 30 beats, appr. sim. 7 beats. Rating: 1 1 2 2 3 0 0 0 0 0
- Staff 3: Similarity rating for sequence (54 12 1 3 10), start-position beat 54, sequence-length 12 beats, appr. sim. 10 beats. Rating: 2 1 2 2 2 3 3 1 0
- Staff 4: Similarity rating for sequence (66 12 1 3 10), start-position beat 66, sequence-length 12 beats, appr. sim. 10 beats. Rating: 2 3 2 1 1 2 2 3 3 1

Figure 6-32. Excerpts of Boléro by M Ravel (1928)), the identified similar segments are superposed. The similarity rating values are displayed between each pair of segments for each corresponding beat.

As can be read from this example, the similarity between the segment pairs is rather general. The similarity is made up from different structural features, a combination of general similarity regarding register change and more specific similarity regarding global rhythmic changes and beat-to-beat register change. Such general similarity would not have been structurally significant under different circumstances and the perceptual reality of this similarity can be questioned. However, these sections were identified as similar by a majority of the subjects in *experiment 3*, see section 6.3.6, which is important for the syntactic analysis of the melody.

X-sequences are determined only by similarity regarding global changes. An X-sequence does not require any similarity at a local level. Thus, the discontinuity at the sequence boundaries must be greater than within the sequence. In addition, an X-sequence must have corresponding global change of melodic direction within the sequence.

Consider the following example, identified as an X-sequence. It consists of a successive global melodic direction upwards of about nine beats, followed by a global melodic direction downwards of the same length. The similarity is determined by continuous change of register

for nine beats even though the change is in the opposite direction in the corresponding segments. Possible points of change of melodic direction are coded by numbers above nine (10, 20, 30) depending on the ambiguity and strength of indication. Rising and falling melodic lines are coded as 1 and -1 respectively. Repetition is coded as 0.

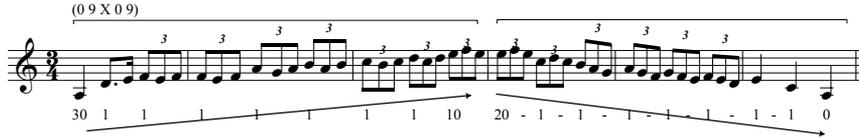


Figure 6-33. Segmentation by corresponding global melodic direction: Code of global melodic direction displayed below the system.

The similarity can be regarded as similarity in terms of compatible changes. It implies that e.g. the composition technique of melodic *inversion* will be recognized as structurally significant, while a melodic *retrograde* will not be recognized if it does not imply concurrent successive changes. Since global melodic changes seldom depict precise segment boundaries, the interpretation of X-sequences is highly contextual.

### T-sequences – Rhythmic similarity

Sequences defined by rhythmic similarity are labeled T-sequences in the current model. In analogy to the categorization of segmentation based on similarity regarding melodic pitch contour, categorization of rhythmic similarity is based on the relationship between interonset structure and metrical structure at tactus/central beat level.

The different aspects of this relationship have been discussed in (section 6.1.5.2 *Rhythmic Structure*). It is manifested in a categorization of rhythmic similarity at different levels based on the categorization of the rhythmic qualities of beats. These range from equal successive duration changes within beats (level 3 similarity) to the most general rhythmic similarity, which concerns the categorization of beats into divided, undivided and tied beats (level 1 similarity).

T3 similarity, the most specific level of rhythmic similarity, is based on the successive duration relationships between pairs of events within and between beats categorized as longer, shorter or equal. This implies that the successive duration change between each event is discriminative, not the actual duration. In other words, two events both of which can be categorized as the first event being either longer than, equal to or shorter than the next event will be considered equal.

T3 similarity - values designate successive rhythmic relationships

A1

A2

Figure 6-34. T3-sequence similarity. Similarity is determined by successive rhythmic relationships between events within and between beats, based on a categorization of each beat. The two systems represent two segments categorized as equal.

The first beat of the two segments above is considered equal according to this categorization because the first event is longer than the second event. The second beat is analogously categorized as two events of equal duration followed by a longer event, etc. Thus, the difference in regarding actual duration is disregarded. The reason is that more specific similarity is not assumed to be structurally significant.

T2 similarity represents a somewhat more general type of rhythmic similarity. It is based on the categorization of beats into divided, undivided and tied. In addition, the level of division, i.e. the number of events dividing a beat, is regarded as structurally significant. This categorization differs from T3 similarity in that it disregards the successive duration relationships between events. Hence, it represents a more general level of similarity.

T2 -similarity: numbers designate division per beat

A1

A2

Figure 6-35. T2 similarity. Similarity is determined by the division of beats.

As can be seen in the present example, the division of the corresponding beats in the upper and lower systems is equal, while the successive relationships between events within beats are different.

The indifference regarding specific successive rhythmic relationships within beats and makes it possible to trace similarity by division which is structurally significant but not obvious. The following example is an excerpt from the solo part of *Bolero* (M. Ravel), in which the segmentation is revealed by T2-similarity:

T2 similarity not identified on T1 or T3 level  
sequence of undivided, div. in 2 and div. in 3 events per beat

1 2 3 1 2 3

Figure 6-36. Excerpt from *Bolero* (M. Ravel). Segmentation by rhythmic similarity on T2 level. Sequence marked by brackets.

The most general level of rhythmic similarity is labelled T1 similarity. This is like T2 similarity based on a categorization of beats into divided, undivided and tied, but disregarding the level of division of the beat. In fact, T1 categorization is inherent in the higher level T similarity categories. This categorization reflects the primacy of metrical structure, implying that segments consider the grouping of beats.

T1 similarity is more specifically based on a categorization of beats into four different *duration-classes*: beats without onset at beat start (*at-tie*), divided beats (*short*), undivided beats (*long*) and beats with a duration which exceeds the length of the beat (*>long*). Note that this classification involves the exclusion of articulative rests and ornaments, i.e. events not considered rhythmically significant as well as the exclusion of short ties – anticipative/pre-starts – without assumed structural significance.

T1-similarity: 1 is divided beat, 2 is undivided beat

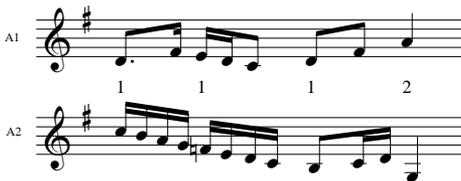


Figure 6-37. T1-similarity. Similarity based on the classification of beats into divided, undivided and tied.

As is demonstrated by the above example, the similarity determined by T1-sequences is very general. Thus, the salience of T1 similarity is related to the categorical discreteness of the sequence, i.e. how discrete a sequence is in terms of duration-classes, and inversely related to the length of the sequence in terms of beats.

T1 similarity can even be more general and allows differences between sequences regarding *duration-classes* if the majority of beats within the sequence are identical regarding *duration-classes*.

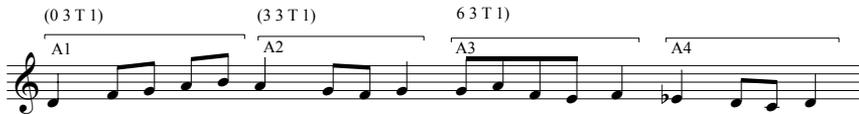


Figure 6-38. Extended T1 similarity. Allows imperfect similarity regarding duration-class to be identified. Identified sequences designated by <start-position (beat no) > <sequence-length (in beats)> <sim. type> <sim. level> below.

T1 similarity has proved to be the most frequently used type of rhythmic similarity for determining segmentation throughout the test material, which is not overruled by or replaced by pitch similarity. (see section 6.3.2.2 *results*)

T6 similarity concerns similarity regarding beat length. This type of similarity is evidently only in question when there is change of beat length within the melody. But in such melodies beat length (T6) sequences are most structurally significant, implying that sequences that are not concurrent regarding beat length will have extremely weak structural significance. It is also

## Chapter 6

a quite common structural signifier in such melodies judging from the test material (see section 6.3.5).

The following example is an excerpt from a south Indian raga “Abhogi” (K. Shivakumar 1992). It is a typical example of a melody in which the major structure is determined basically by T6-similarity. The excerpt is also an example of T3-similarity as structural signifier.

The image shows two systems of musical notation, A1 and A2, for an excerpt from Raag Abhogi. System A1 is on a treble clef staff with a 2/4 time signature. It contains a sequence of notes with several brackets above it. The longest bracket is labeled '(4 5 T 6) 3+3+3+3+4'. Below it, a shorter bracket is labeled '(4 2 T 3) (512 512 512) (1024 512)'. Below the staff, five groups of notes are delineated by hooks, with MIDI duration numbers 1536, 1536, 1536, 1536, and 2048 written below them. System A2 is on a treble clef staff with a 3/4 time signature. It contains a sequence of notes with a bracket above it labeled '3'. Below the staff, three groups of notes are delineated by hooks, with MIDI duration numbers 1536, 1536, and 2048 written below them.

Figure 6-39. Excerpt from Raag Abhogi (after K Shivakumar): T6-similarity (long brackets) and T3-similarity (short brackets). Hooks delineate beat-groups. Numbers between systems represent beat length in MIDI duration.

Rhythmic similarity (T-sequences) is generally more forceful than pitch similarity on local level as a consequence of the primacy of interonset structure locally (see section 5.1.3 *Metrical analysis*). In contrast pitch similarity is more forceful on a global level (see also section 5.3.8.3 *Evaluation of metrical phrase and section analysis*).

## 6.2.4 Evaluation of sequence structure

### 6.2.4.1 Outline – general design

#### General principles

The sequence structure obtained by the sequence analysis described above is evaluated successively through the melody in order to model the successive conception of melodic structure predicted by the general model of melody cognition (see Chapter 3).

The fundamental principles behind the evaluation are:

- (1) Specific similarity is superordinate to more general similarity
- (2) Longer/complete similarity is superordinate to shorter/incomplete similarity.
- (3) Discrete sequences are superordinate to less discrete.

Moreover, there is a general difference regarding valuation of rhythmic sequences (T sequences) in relation to pitch sequences (NIL, R, O, C and I sequences). Rhythmic similarity is superordinate to pitch similarity regarding short/local sequences (shorter than the categorical present  $7+/-2$  beats), while pitch contour similarity is superordinate to rhythmic similarity regarding longer/global sequences. This reflects the basic assumption that onset/offset information is superordinate locally – at the level of the categorical/perceptual present (See section 5.1.3 *Metrical analysis, Assumption no. 18*)

The discreteness of a sequence is evaluated based on the discontinuity value of the sequence boundaries.

The evaluation also involves the secondary and third level of general grouping principles (see section 6.1.7). The principle of constancy/good continuation is reflected in the preference for sequences similar to previously identified sequences, preference for established structure. Once a sequence has been identified and is considered valid/salient locally, a new search for starts similar to the current sequence which can result in new sequences which are added to the body of sequences to be evaluated.

The principle of constancy/good continuation is also reflected in the preference for continuity regarding segment length (and hence sequence length) on corresponding hierarchical levels.

The evaluation further involves the principle of symmetry, which is reflected not only in a general preference for symmetrical segmentation, but also in that symmetrical segmentation is assumed to be hierarchically consistent, creating expectation of consistent and hierarchically coherent symmetrical segmentation. The expectation of symmetrical segmentation, can according to the assumption of secondary grouping give rise to implied/inferred segmentation, which is implemented in the model through *implied* sequences (Y-sequences) derived from previous or hierarchical symmetry. This is only implied segmentation, not based on any similarity between the segments.

## **Method - basics**

Technically, the evaluation is based on a structural prominence index value. This is in turn based on a segment boundary value, which is the sum of the discontinuity values for each segment boundary. This value is then multiplied by different factors based on secondary and third grouping principles - the *gestalt* salience/*prägnanz* and consistency features mentioned above, forming a structural prominent index. One reason for using factors of weighted importance rather than fixed values for each quality is that these factors either apply to the entirety of the sequence or depend on the distribution of the individual segment boundary values.

The structural prominence index value (described in detail below, see section 6.2.4.2) is one important piece in the evaluation of sequence structure. However, most aspects of secondary and third grouping principles, such as good continuation /consistency of structure, symmetry and salience / *prägnanz*, cannot be evaluated solely with regard to the features of the individual sequence, since they depend on context. A sequence of low structural prominence value can be salient / relatively prominent in a context with no competing sequences while a strongly marked sequence of great integrity can be inaudible in a context where it is non-consistent with an overall structure. Therefore, a successive sequence evaluation is performed in which the structural prominence index plays an important role, but with the addition of several other factors. This successive evaluation is, as mentioned, performed in two steps: first, evaluation of sequences at section level and then evaluation of sequences at phrase level (see section 6.1.3). The design reflects the general model of melody cognition (Chapter 3), which assumes that new information can imply a revised conception of previous structures and that global structure – sections – rules local structure – phrases. The result of the section analysis must then be input to the phrase analysis.

There is yet another important reason for the separation of section and phrase analysis, namely that the influence of different structural factors are quite different. On a large scale level, periodicity, constancy and symmetry have considerably less significance and precision than within smaller time-spans. This design allows the differences between these two scopes to be formulated and valued.

The two evaluations are, however, identical regarding the general concept of the evaluation.

### **Method – involving contextual factors**

The successive evaluation for section and phrase level respectively is briefly performed in six steps:

1. Evaluation of sequences beginning at the same position;
2. Evaluation of first half of sequence – a part;
3. Evaluation of sequence hierarchy (may include recurrence of 2.);
4. Analysis of symmetrical implication;
5. Possible addition of new sequences based on similarity with the identified sequence;
6. Evaluation of second half of sequence and sequence overlap – b part.

The first step is based on the structural prominence index of the sequence. The sequence with the highest structural prominence index value is, in the second step, tested against the index values of subsequent sequences starting within the first half of the sequence in question.

If the sequence is regarded as salient, then, in the third step, a second evaluation of the sequences beginning at the same position take place, with regards to sequence hierarchy. This hierarchical analysis is, besides the structural prominence index, also based on factors of symmetry and integrity of each hierarchical level regarding sequence length (section 6.2.4.3). If a hierarchy is identified, the largest sequence at the position is tested for highest structural prominence index value against the sequences beginning at the first half of the sequence. The sequences at the same hierarchical position receive a value boost. If the sequence in question passes in this test, it is regarded as significant.

In the fourth step, the symmetrical implications of the identified sequence are analyzed, based on the structural features of the sequence and other sequences at concurrent positions. If implicated symmetry is recognized new sequences – not based on actual similarity but implicated symmetry, Y-sequences – are added to the sequence list (section 6.2.4.4). This reflects the assumption that secondary grouping principles can implicate segmentation. The fifth step involves the search for new sequences based on similarity with the identified sequence and is reflects the assumption that starts which are similar to already identified starts will be easier to recognize. In certain conditions, identified sequence starts are designated as defined-global-motifs indicating the signal function. If such sequences are found, they are added to the sequence list (section 6.2.4.5).

The last step considers the second half of the identified sequence. Since the identity of a sequence is based on the integrity of the sequence starts, the start of the first and the second half of the sequence are essential for its identification, a repetition or a similarity being heard. In this model, it is assumed that the integrity of the sequence is not as dependent on the second half of the sequence since the similarity which constitutes the sequence does not have to be complete with respect to the basic rule of more similar than dissimilar. The model thus allows end sequence overlaps, i.e. new sequences beginning before the end of the current

sequence and mid sequence overlaps, i.e. new sequences beginning at the mid position of the sequence (section 6.2.4.5)

The evaluation of mid- and end sequence overlaps against implied sequence end involves secondary and third grouping principles as well as structural prominence index valuation. In the following, the crucial factors and procedures in the successive evaluation of sequence structure will be described more thoroughly.

### 6.2.4.2 Quantifying structural prominence of sequences

#### Outline

The fundamental assumption regarding the relationship between grouping by similarity and discontinuity states that sequences are superordinate to discontinuities globally as determinant of melodic structure, but that discontinuities are superordinate to sequences locally, determining the phase of the sequence. (See *Assumption no 9*) In other words, the prominence of a sequence is related to its integrity, which is dependent upon how salient and discrete its boundaries are.

This relationship constitutes the fundament of the sequence evaluation, implying that the basic value of prominence of sequences is made up from the local and global discontinuity values of the segment boundaries, the start-, the mid and the end of the sequence. (See Fig 40 below.)

To make it easier to follow the presentation below sequences are described by the internal encoding of sequences within the model (see section 6.2.3.3., e.g. fig. 20). This entails the designating at a sequence by start, length, type of similarity, level of similarity and length of concurrent parts, (<start> <length> <type> <level> <equal part>), where position and length are measured in beats at central pulse/tactus level. The code (0 6 NIL 4 3) hence designates a sequence that starts at the first beat (0), with the length of six beats of each similar segment (6), with pitch contour similarity at level 4 (NIL 4) and three similar beats per sequence half (3).

Sequence: (0 6 NIL 4 3)

Discont. val.: 12 START-POS      4 MID-POS      5 END-POS

Figure 6-40. The three sequence boundary positions with discontinuity values for each position for the sequence (0 6 NIL 4 3). The concurrent parts of the sequence halves are shadowed.

The sequence in this example would have the basic prominence values of 12, 4 and 5, resulting in a total basic prominence value of 21. The basic features of the analysis of discontinuity values will be described below under *Sequence boundary evaluation*.

Apart from this, the structural prominence of the sequence is evaluated according to a number of different similarity, contextual and *gestalt* salience factors described below under *Structural prominence index*, through the multiplication of these factors with the basic prominence value.

#### Sequence boundary evaluation

The basic prominence values are based on analysis of local discontinuities, analysis of global discontinuities regarding interonset/rhythm, change of pitch register and change of

melodic direction (or successive change of pitch register). The analysis of local discontinuities almost exactly replicates the analysis of local discontinuities described in Chapter 5, Section 5.2.4.6 *Table 1. Conditions for segment boundary valuation of melody events*, the most important features overviewed in the preceding pages.

The main differences are due to the object of the valuation. In the present case, the valuation concerns sequences, as a consequence of which value enhancement due to concurrence with sequence boundaries is less useful. Instead, similarity regarding rhythmic properties and global changes at the start of the two sequence halves are used as input to the local discontinuity evaluation.

The major differences in relation to the local discontinuity valuation (or start indication value) as displayed in Chapter 5, section 5.2.4.6, *table 1*.

The interpretation of *xlong* duration-class after *tied*, *divided* or *undivided duration-class*<sup>84</sup>, paragraphs D2 and D3, is dependent on rhythmic duration-class similarity between the starts of the two sequence halves.

The exclusion of discontinuity/start indications based on return from *mp*, label V, and sequence boundary concurrence, label W.

Apart from the local discontinuity (or start indication) value global discontinuities contribute to the discontinuity value. The analysis and valuation of global discontinuities are described in Chapter 5, section 5.2.4.6, *Analysis of global discontinuities*. Below is an example of how discontinuity values can contribute to the total discontinuity value for the start-, mid- and end-position of a sequence.

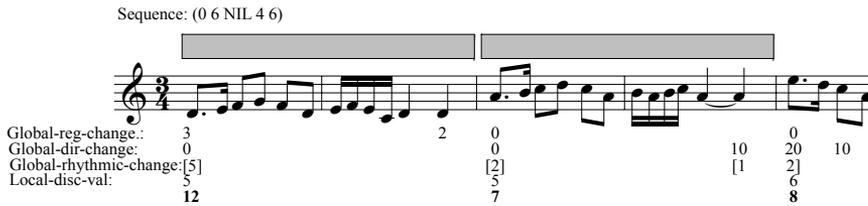


Figure 6-41. The different components of the total discontinuity value for the sequence (0 6 NIL 4 6), displayed for the start-, mid- and end-position respectively

## Structural prominence index

While the sequence boundary evaluation provides the basic valuation of the integrity of the sequence, its structural prominence is assumed to be influenced also by other structural factors considering the *prägnanz* and *salienc*e of the sequence. This is dependent on the specificity/level and type of similarity, the consistency in the delineation/markings of the sequence, the uniqueness/integrity of the sequence in relation to other sequences and, in relation to the substructure of the sequence, the position of the sequence in relation to established structure or implied structure by symmetry or periodicity.

<sup>84</sup> see section 6.2.3.3, *T1-sequences for explanation or duration-class*

In the model the influence of these different factors are measured as ratios by which the basic prominence value is multiplied, forming a structural prominence index labeled `tot-eval-value`. Thus the totality of the influence of these factors on the integrity/salience and prominence of the sequence is reflected in the measurement. The actual factor values are chosen to reflect the structural hierarchy of prominence as simply as possible.

In the following the different factors that are used in the calculation of the structural prominence index will be explained further.

1. Type and specificity of similarity – *seq-type/c4-factor*

The valuation of structural prominence of sequence types is related to the specificity of similarity, but also to how structurally determinative different structural features are: The most specific similarity is categorized as NIL 5 – sequences (see section 6.2.3.3), which is a true repetition of an array of notes and durations. Two of the most general types of similarity are the T1 – similarity regarding duration-classes of beats (see section 6.2.3.3) and T6 – similarity regarding beat length. Still the T1 and the T6 similarities are among the most structurally prominent – at the same level as NIL 5- sequences, since they represent the most basic types of rhythmic and metric compatibility between segments, which are most determinative on a local level. Thus, it is assumed here to be highly unlikely that we will consider two segments with e.g. different beat lengths to be structurally equivalent.

However, the structural significance of T-sequences (rhythmic similarity) is assumed to be strongest regarding short sequences within the temporal/categorical present, to decline regarding longer sequences where interonset similarity will be harder to recognize.

Sequence type	Max. factor	Restriction
NIL 5	4.0	marked, with comp. beat duration changes
NIL 5 / O 5 T1	3.0	without beat dur. changes 3-4 beats
NIL 4 / O 4 T6 T3 I3	2.0	not unmarked general (2-12 beats) 5-12 beats
NIL 3 C 3	1.5	
R T1-3 X	1.0	general
X	0.5	> 16 beats
T6	0.125	> 12 beats

Table 27. Sequence type enhancement factor (*TYPE/C4 VALUE*). Table shows maximum enhancement factors, also dependent on minimum strength in sequence boundary indication. (R, longer T and X sequences does not receive any type/level factor enhancement)

As can be read from the table, more specific similarity regarding pitch is ranked before less specific similarity, while more general rhythmic specificity is ranked before less specific in some cases. Reversed similarity (R sequences) and similarity regarding global changes (X-sequences) are regarded less structurally significant.

The order of structural significance as shown in the table above reflects the fundamental assumption of the relationship between rhythmic similarity and pitch similarity as structural determinants (*Assumption no 35*). This implies that pitch contour sequences are structurally superordinate to rhythmic sequences regarding longer sequences, while rhythmic similarity is superordinate to pitch similarity regarding short sequences.

The salience and hence structural prominence of a sequence can be regarded as proportional to the level of specificity regarding pitch contour similarity of the sequence, since it affects the uniqueness/integrity of the sequence (*Assumption no 25*). This is reflected in the higher level of type factor enhancement for pitch change sequences of higher specificity.

The prominence of start-oriented grouping over end-oriented grouping is reflected in the lack of enhancement for R sequences regardless of level of specificity.

2. Range of concurrent similarity – *seq-len-factor* and *complete-seq-factor*

The integrity and salience of a sequence is also dependent upon how large portion of the sequence that is actually similar. This is reflected in a factor the value of which is related to the proportion of similarity. This does not concern rhythmic sequences (T-) since these are only considered when the entire sequence conforms to the similarity condition.

But especially longer pitch contour sequences can be structurally significant even though only part of the sequence is recognized as similar. The basic conditions for recognized similarity are either that more than half the sequence conforms to the similarity condition or at least 5 beats (the lower limit of the categorical present 7-2).

The reduction factor values are shown in the table below. Besides the basic >50% category and identical sequence category (> 90%, less than 3-5 beats not similar), the most important limit is the 2/3 -limit, which implies that two segments are similar when the content which is similar is greater than the content which is dissimilar (see also section 6.2.3.3 *C-sequences*). The 1/3 limit is the reversal of the 2/3 limit, giving the possibility for a identification of content between other segments.

Apart from the ratio, the actual number of beats involved is of crucial importance. The limits regarding the absolute number of beats involved are based on the category limitations of the perceptual present (see Chapter 5, section 5.2.4.7, *Beat scopes*). This is the basis of the assumption that 5 beats of similarity between segments is enough to identify distant melodic segments as similar (the minimum categorical limit). This gestalt constraint also implies that differences below 3 beats can be disregarded for sequences that encompasses more than the range of a basic beat scope (12 beats).

Ratio/no. of beats of concurrent similarity	Min. reduction/ Max. enhancement factor	Restriction
< 1/3	0.50	(< 1/4 $\Rightarrow$ 0.125 for non section-search)
< 1/2	0.75	
> 2/3	1.25	
> 9/10	1.50	C-sequences (< 20 beats) O/NIL-4/5
> 2/3 and > 5 beats conc.		
> 8/10 and < 3 beats diff. > 9/10 and < 2 beats diff.	2.00	O/NIL-4/5 C-sequences (> 20 beats)
> 1/2 and > 15 beats conc.	3.00	NIL-4/5

Table 28. Reduction factor values for different levels of similarity concurrence. NB. Does not apply to T-sequences or X-sequences.

3. Consistency and strength in the delineation of the sequence – *r-cue-factor*, *non-overlap-factor* and *seq-type/c4-factor*

The discreteness of the sequence boundaries cannot be regarded only as the mere sum of the individual sequence boundary discontinuity values. The distribution of discontinuity values also matters.

Obviously the start- and mid-positions of the sequence are more important than the end point of the sequence for the salience of the sequence, following the assumption of primacy of start orientation (*Assumption no 6-8*). Thus, it does not matter how strongly the end position of a sequence is marked if the mid position is unmarked, in which case the sequence probably would not be recognized. The most important quality of the distribution is, however, that it concerns two successive sequence boundaries. Discrete changes at the mid- and end-position can function as delineating a sequence by its endings as well as discrete structural changes at the start- and mid-position by its start.

Therefore, corresponding equal global level rhythmic changes (regarding duration-class, beat-duration change etc.) at two successive sequence boundaries – *rhythmic cue factor* - can give a value enhancement from 1.5 to 3.0. T1-sequences without corresponding global level rhythmic change at start- and mid-position are practically ignored.

Also, evident start positions, such as the first beat of a melody or the first beat with note onset after a rest of several beats, which are unlikely to be part of a previous sequence are interpreted as to provide a general enhancement to the integrity of the sequence. This is reflected in the *non-overlap factor*, which in combination with successive marked start- and mid-position can give a value enhancement of 1.5 to 3.0.

Strong discontinuity values at start- and mid-position can by itself give value enhancement to a sequence, but the *marker distribution factor* is generally defined in negative terms, giving a reduction factor between 0.75 and 0.5 for X-sequences, T-sequences and longer contour sequences with start- and/or mid-boundary value below 2. A start- or mid-boundary value  $\leq 0$  gives a reduction factor between 0.1 and 0.5 depending on the type, length and concurrent similarity of the sequence.

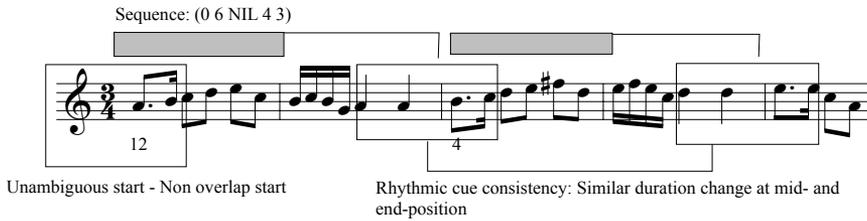


Figure 6-42. Example sequence with non-overlap factor because of evident start position and rhythmic cue factor due to similar rhythmic changes at mid- and end-positions and duration-class similarity regarding entire sequence

- The uniqueness and integrity of the sequence – *chained-seq-factor* and *latticity-factor*  
 The integrity of a sequence is evidently related to its uniqueness – if it cannot be mixed up by other sequences of the same length but different phase it will be considered discrete in relation to the surrounding structure. Thus, the first in a line of similar sequences will be more likely to be recognized than succeeding sequences.

This is reflected in the *chained sequence factor*, which gives a reduction with 0.5 times the prominence value for sequences which can be derived from a sequence starting on the preceding start position.



Figure 6-43. A root sequence with a subsequent chain of sequences

In the above example all sequences of the same length, type and level will receive the chained sequence reduction factor in the calculation of its prominence value.

Another factor that affects the integrity of a sequence is the *latticity* factor level<sup>85</sup>, i.e. the level to which a sequence can be conceived as a composition of equivalent substructures. The more salient the sub-structural level, the less salient the sequence as a whole.

Consider, for instance, a boarded wall or a typical wallpaper pattern. In such repeated structures it is usually possible to imagine a number of different superstructures of low structural integrity, since the content of the superstructure will not be unique and discrete in relation to other superstructures.

One aspect of this is the number of equal/similar sub-sequences composing a sequence: The greater the number of equal/similar sub-sequences, as defined by type, level and range of similarity, the less salient will the super-structure will become.

<sup>85</sup> My own term, here suggested, to be read as the level to which as structured can be conceived as latticed.

Seq.: (0 12 NIL 5 12)  
Seq.: (0 6 NIL 5 6)

Seq.: (12 6 NIL 5 6)

Figure 6-44. Example of high latticity factor. The sequence (0 12 NIL 5 12) is composed of two equal sub-sequences of NIL 5 similarity, (0 6 NIL 5 6) and (12 6 NIL 5 6), resulting in a reduction factor of 0.25.

5. The position and identity of the sequence in relation to established structure – *previous-sequence-equality-factor* and *inherited-sequence-start-factor*

The grouping principle of good continuation/constancy (see section 6.1.7 *General grouping principles*) is reflected in the structural prominence index in various ways. One aspect of this is the relationship to established structure, which can be viewed from two different perspectives: (1) The similarity with already identified segments and (2) The concurrence of the sequence with already identified segment boundary positions.

This is reflected in the calculation of the structural prominence index through the previous-sequence-equality-factor and the inherited-sequence-start-factor. The first of these refers primarily to the structural start similarity between the sequence in question and previously identified sequences, while the latter primarily refers to the concurrence between the start position of the sequence and previously identified start positions.<sup>86</sup>

Seq.: (12 6 NIL 4 3)

Figure 6-45. Example of previous-sequence-equality factor. The sequence (12 6 NIL 4 3) receives an increase of the value by a previous-sequence-equality factor of 2, through the start similarity with the preceding sequence (0 6 NIL 4 3) even though the sequences are not regarded as wholly similar at level 4.

<sup>86</sup> Since the evaluation of sequence structure is performed in two steps, beginning with the section analysis concerning the major structural segments, this factor involves position concurrence with section segments in the subsequent phrase analysis (lower level segments).

Note that the successive consistency of start positions and sequence length is also valued in the successive and contextual evaluation of sequences and not primarily through the index value.

6. The position of the sequence in relation to implied structure by symmetry or periodicity – *metrical position* and *symmetry factor*

Yet two other aspects of secondary grouping principles are implemented in the structural prominence index: good continuation regarding consistency of metrical position and concurrence with implied symmetry.

When a sequence structure implies consistent repetition of sequence length within the limits of periodicity perception (see Chapter 3, *general model*), this will create the cognition of a regulative metrical structure, a metrical grid, which implies that segments that concur with the metrical grid will be structurally prominent. This does not concern beat structure, which is input to the metrical phrase and section analysis, but rather higher level metricity, such as measure and *period* level. This is reflected in the structural prominence index through a reduction factor that is employed when a sequence length does not conform to the established metrical grid at measure level.

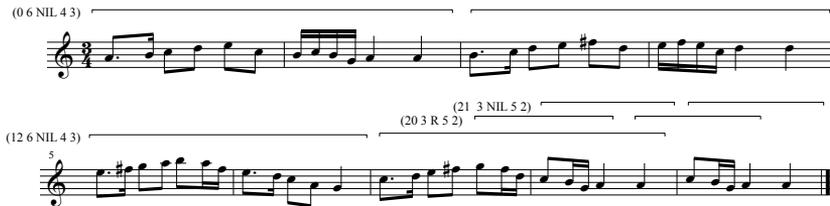


Figure 6-46. Example of metrical position factor. The sequence (20 3 R 5 2) receives a reduction by metrical position factor 0.5, while the sequence (21 3 NIL 5 3) is congruent with the metrical structure established by the sequences (0 6 NIL 4 3) and (12 6 NIL 4 3) and does not receive any index value reduction.

Analysis of implicated symmetry is performed in another part of the evaluation of sequence structure and will be explained separately (See section 6.2.4.4). When symmetrical structure is recognized by the model, sequences that do not concur with the implicated symmetry receives a reduction factor similar to the metrical reduction factor.

### 6.2.4.3 Quantifying structural hierarchy

As described in the section 6.2.4.1 sequences beginning at the same position are evaluated regarding their hierarchical relationship. This involves the influence of secondary and third grouping principles, such as symmetry, good continuation and prägnanz/ gestalt integrity besides the evaluation based on the structural prominence index value described above.

In the sequence analysis, hierarchical relationships are identified to select the hierarchically discrete sequences beginning at the same position, excising the sequences which interfere with the hierarchy.

It is assumed here that an ideal hierarchical relationship between a higher level and a lower level implies the division of the higher level in two equal parts or, inversely, the forming

of a higher level by combination of two equal parts. More generally, a hierarchic relationship should ideally imply the division of a higher level in as few parts as possible. It is generally assumed in this study that a hierarchical relationship on the level of melodic segments is limited to a maximum of four segments a lower level forming a higher level segment - a 4:1 relationship between levels – but a 2:1 relationship being strongly preferred. This is exactly the limitation of implicative symmetry (see section 6.2.4.4), but in the case of hierarchical relationship the 2:1 relationship is extremely dominant since a symmetrical relationship can be retrieved from a hierarchical structure, while the conceived hierarchy is formed continuously by the grouping of contiguous segments into higher levels. In practice, only division in two and three has to be considered in the hierarchical analysis, since these relationships require the excision of interfering possible segments to be formed.

Categorized as the division of a higher level into small whole numbers, the assumption of hierarchical relationships gives the possible hierarchical relationships between two segments being either 1:1, 2:1, 3:1, i.e. either equal length, divided in two or divided in three.

This leads to a definition of hierarchical levels as discrete categories (*the categorical proportion rule, cf. chapter 3*). Thus, a 1:1 relationship (equal length) is delineated by the relationship to a 1:2 ratio (duple division), stating that it is 1:1 ratio when it is closer to 1:1 than to 1:2, which implies a ratio above the geometrical mean, i.e.  $1:\sqrt{2}$  ( $\approx 2:3$ ). Consequently 1:2 ratio (duple division) is delineated by its relationship to 1:3 ratio (triple division), stating that it is 1:2 division when it is closer to 1:2 than 1:3, i.e. from  $1:\sqrt{6}$  ( $\approx 2:5$ ) and above. The inclusion of the boundary ratio in the duple division interpretation is due to the assumed preference for duple division. These two categorical definitions of equality of length and duple division is here called *the categorical proportion rule* or *2/3 – 2+3 rule*.

*Figure 6-47. Illustration of discrete hierarchical level categorization, categorical proportional rule or 2/3 – 2+3 – rule. The limit for the categories equal length and duple division, based on approximation of the mean between the proportions to the arithmetic mean.*

Below follows two musical examples that illustrates the maximum allowed deviation from perfect duple division, 1:2 ratio in the model:

Lower limit of duple hierarchical relationship - 2:3 between dividing segments

Seq.: (0 15 NIL 5 15)  
Seq.: (0 9 NIL 4 5)

Seq (15 9 NIL 4 5)

Upper limit of duple hierarchical relationship - 3:2 between dividing segments

Seq.: (0 15 NIL 5 15)  
Seq.: (0 9 NIL 4 6)

Seq (15 9 NIL 4 6)

*Figure 6-48. Maximum deviation from perfect duple division (1:2), which is limited to between 2:3 and 3:2 relationship between the divisive segments (2:5-3:5 division)*

Apart from the ratio definition of the interference of sequence length (< 1/3 difference) mentioned above, there is also a measure of interference which states that two sequences that differ less than three beats regarding length, will be considered equal if it is not a 1:2 length ratio. This is based on the assumption that a discrete gestalt/group requires three beats – two changes (see Chapter 5, section 5.2.3.2)

Interfering sequences are evaluated based on gestalt salience and integrity, the structural prominence index as well as other structural indications such as similarity type, similarity level and *start-* and *midmarker* values. Furthermore, interfering sequences are evaluated on grounds of even division: A sequence which divides a superordinate sequence into equal parts is regarded structurally prominent, while a uneven division is regarded being less prominent.

#### 6.2.4.4 Segmentation by symmetrical implication

##### Implicated symmetry sequences

As mentioned above, symmetry is an influential factor in the evaluation of structural hierarchy. But the influence of symmetry goes beyond that. It is assumed here that symmetry, as a secondary grouping principle (see section 6.1.7), can create segmentation by symmetrical implication. This refers to segmentation not based entirely on features of the melodic structure, such as similarity or discontinuity but rather on expectation of and preference for symmetrical grouping. It means that, if melodic phenomenal structure complies with a

symmetrical segmentation and is indicated by symmetrical segmentation through sequences or discontinuities, it will be applied to all substructures within the superordinate structure.

Thus, the sequence structure cannot be evaluated without taking symmetrical implication into account. The influence of symmetry regards both the evaluation of sequences and segmentation based on symmetry, referred to as implicated sequences.

The method of involving implicated sequences in the evaluation of sequence structure in the current model is accomplished by creating ‘empty’/‘dummy’ sequences based on symmetrical implication, designated type Y sequences. Like other types of sequences, these are involved in the subsequent evaluation of sequence structure and are evaluated with regards to their structural prominence index value.

## General principles

The analysis of implicated symmetry relies on two types of structural indications:

- (1) Sub-sequences concurring with symmetrical division of a principal sequence.
- (2) Global discontinuities concurring with symmetrical division of a principal sequence

The first of these indications is assumed to be more significant than the second, in compliance with the general assumption of primacy of segmentation by similarity before segmentation by discontinuity regarding large-scale melodic structure.

Further, it is contingent upon the quality of the symmetrical indication. This concerns:

- a) The degree of symmetrical division, i.e. the number of dividing segments.  
It is assumed here that the significance of the symmetrical implication is related to the number of dividing segments, a simple division being more significant – giving stronger implication – than a more complex division. (see also section 6.2.4.3, *Quantifying Structural hierarchy*). The reason for this assumption is that the temporal dimension is regarded a limiting factor for the possibility of conceiving a division as symmetrical, while the maximum division of symmetrical implication is set to 1:4
- b) The perfection of the symmetrical division.  
It is assumed that perfect symmetry, i.e. strictly symmetrical division of a principal sequence into equal parts is more prominent than a slightly asymmetrical division. Note, however, that deviations from perfect symmetry are accepted since implicated symmetry is regarded to be preferred / searched for in the conception of a melody.
- c) The structural integrity of the division  
This relates to the afore mentioned *latticity-factor* (see section 6.2.4.2, *Quantifying structural prominence*). This reflects the assumption that the more evident/salient the sub-structures are, the less salient the principal structure becomes, the structural integrity of the principal structure being diminished. This is influenced by the number of substructures (level of division) but also by the degree of similarity between the substructures and the degree of structural discontinuities within a principal structure.
- d) The structural prominence of the dividing segments  
It is assumed that a structural prominence of the dividing segments will influence the degree of symmetrical implication, the more prominent the stronger the implication.
- e) The general evidence of symmetry within a structure  
It is assumed that symmetry on one level implies symmetry on all other levels, symmetrical implication being hierarchical (see 6.2.4.2 *Quantifying hierarchical structure*).

## Method

The method of examining the symmetrical properties of a given sequence reflects the assumed prominence of simple division in relation to complex division: First, duple division is examined, then triple division and finally quadruple division.

The primacy of segmentation-by-similarity is reflected through the involvement of structural indications of sub-sequences before structural indications of global discontinuities. Analogously, sequences that are symmetrical divisions of a principal sequence are evaluated before other sequences, reflecting the assumption of hierarchical consistency of symmetry.

The most problematic part of the analysis of symmetrical implication concerns the categorization of a division into the four possible categories: no division, duple division, triple division and quadruple division. One central question is how much a division can deviate from perfect division and still be conceived as a symmetrical division with symmetrical implication. Following the principles presented above in section 6.2.4.3, *Quantifying structural hierarchy*, segments which differ with less than 1/3 of the longer segments are regarded as equal in length.

The limits of the categories of symmetrical division are set to the mean ratio between the adjacent ratio categories, displayed in the following table

Table 29. *Symmetrical Categorical Implication (limits are approximated to the arithmetic mean)*

Symmetrical div. category	Limit
no division 1:1	> 2:3 (1:1.5)
duple div. 1:2 (diptych)	<= 3:5 (can extend >= 2:5 (1: 2.5))
triple div. 1:3 (triptych)	< 2:5 >= 2:7 (1:3.5)
quadruple div. 1:4	1:4

As can be seen in the table, the duple division (1:2) is favored when the mean limit value falls between duple and triple division. Likewise triple division is preferred before quadruple division, the latter requiring perfect division and perfect structural integrity of the units to be recognized. This reflects the assumption that a more simple division is easier to recognize and is structurally prominent.

This implies that one segment of a duple division and two segments of a triple division can overlap, which is why the categorization has to be supplemented by an examination of the division of the dividing segment has to be employed. Thus, in addition, two segments of a triple division must employ more than 4:7 of the principal segment to be regarded significant.

Consider the examples of deviations from perfect symmetry below. In the first case (duple division), the largest dividing segment is 3/5 of the principal segment. In the second case (triple division), two contiguous segments employ more than 4/7 of the sequence half.

Chapter 6

Example of duple symmetrical division (diptych), with max. deviation from perfect symmetry

Seq.: (0 15 NIL 5 15)  
 Seq.: (0 6 NIL 4 5)

Seq (15 6 NIL 4 5)

Figure 6-49. Duple symmetrical division

Example of triple symmetrical division (triptych), with max. deviation from perfect symmetry

Seq.: (0 17 NIL 5 14)  
 Seq.: (0 6 NIL 4 5)

implied segment

implied segment

Figure 6-50. Triple symmetrical division

In the above cases, the division of the principal sequence originates basically from sub-sequences. However, in the second example of triple division, the third segment is implicated and its boundary is defined by a major discontinuity. Since symmetry is assumed to be a strong structural factor, a symmetrical division can actually be based entirely on global structural changes.

Example of duple symmetrical division based on structural discontinuity

Seq.: (0 12 NIL 4 6)  
 Seq.: (0 6 Y 0 6)

implied sequence

implied sequence

Figure 6-51. Duple symmetrical division based on structural discontinuity

The general assumption of the hierarchical consistency of symmetrical division leads to the assumption that symmetrical implication will be carried out on lower levels if it is confirmed

on higher levels and there is no contradicting structural indication at the level of division. In the example below, the lowest level (the three beat phrase level), which is in fact identical with most probable measure level, is applied because of assumed hierarchical symmetry. There is actually no structural change – discontinuity – strong enough to either confirm or refute this level of grouping.

Figure 6-52. Implicated symmetry based on hierarcbic symmetry, resulting in three beat sub-phrases, (0 3 Y 0 3) etc. Output from computer model. (The exclusion of the last beat from the last phrase is due to graphical limitations in the application)

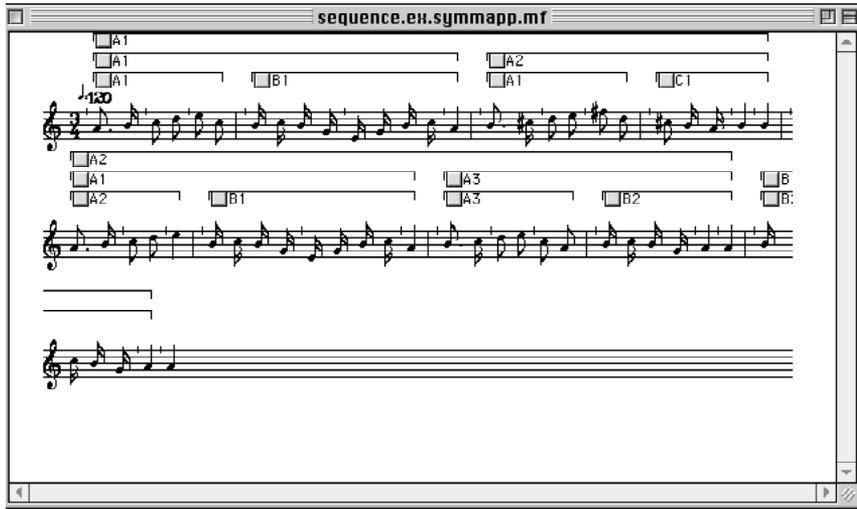


Figure 6-53. Interpretation of metrical structure at measure level originating from the phrase analysis above. Output from computer model. (The exclusion of the last beat from the last phrase is due to graphical limitations in the application)

The quadruple symmetrical implication is employed when there is a chain of sequences with mid-sequence overlap with perfect symmetry, which makes a twofold division non-discrete.

Example of quadruple symmetrical division



Figure 6-54. Quadruple symmetrical division based on mid sequence overlap (0 6 NIL 4 3) - (6 6 NIL 4 3)

### 6.2.4.5 Successive evaluation of sequence structure - overview

#### General design

The successive evaluation of sequence structure is performed in two steps, the section analysis and the sequence analysis. The first of these refers to the segmentation of a melody into the longest sections determined primarily by melodic similarity. The result then provides input to the second step of sequence analysis, which considers more local structure. This

design reflects the basic assumption that global structure is prominent in relation to local structure and that the structural significance of similarity is related to length, longer sections of similarity being prominent in relation to shorter.

The reason for separating these two steps is that the adjacency of segments is a determinative factor in the evaluation of local segment structure while it is not as significant regarding large-scale structure above the level of metrical periodicity which is in the current model set to a maximum of 48 beats. (*Assumption no 15*). Above this level it is assumed that periodicity will be hard to perceive and will thus not be an influential factor in the structural conception at this level. The underlying assumption is that longer periods of time are harder to estimate. This implies that section structure may be determined by non-adjacent sections of melodic similarity which are not symmetrically balanced regarding their extension in time. This further implies that there may be conflicts between more local grouping influenced by aspects of periodicity and more global grouping based on similarity alone. We will return to that in the more detailed exploration of the section analysis below.

In the current model, it is assumed that revision of a local melodic structure due to the conception of large-scale structure is performed continuously. For the benefit of evaluation of the different parts of the analysis, they are separated in the current implementation.

Besides the different weight of periodicity factors, section and sequence analysis is performed exactly the same way. The input to the analysis is a list of sequences. They are evaluated successively based basically on the structural prominence index and the structural hierarchy quantification and when a sequence is considered valid, interfering sequences are removed. As a consequence of the general assumption that longer similarity, rules shorter similarity longer sequences are evaluated before shorter sequences.

Good continuation /constancy is involved in the evaluation through preference for similar and symmetrical sequence length (within the limits given above) and concurring sequence boundaries. A sequence which starts on a position determined by a previous sequence and which is identical or compatible in length with the previous sequence, is considered prominent.

Besides the determination of the sections of a melody by melodic parallelism, extreme global discontinuities are, in fact, regarded in the analysis of sections and sequences. Lack of onset during a categorical/perceptual present (max. 9 beats / 5.625 “) or a contextually extreme period is treated as a Grande Pause/interception. This implies a section boundary not to be overlapped in the sequence analysis.

## **Example – Section analysis**

The following example may serve as a general demonstration of the method of sequence evaluation. The melody in fig.55 below is a folk tune from Sweden adapted from a published transcription with the notated ornamentation removed<sup>87</sup>. For the purpose of the demonstration the time signature has been arbitrarily altered from 3/4 time to 16/8 time. Note that the notated time signature is not used in the analysis.

---

<sup>87</sup> One rhythmic figure has also been slightly changed due to limitations regarding possible rhythms in the implementation,

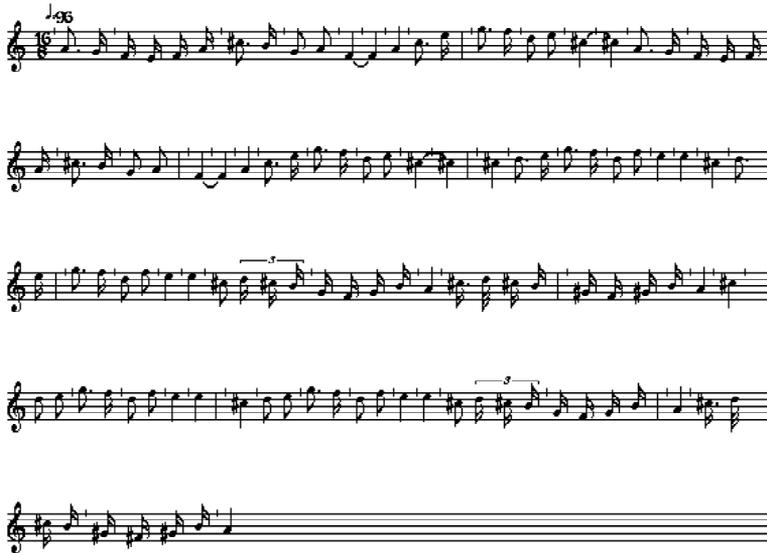


Figure 6-55. Melody adapted from *Svenska Låtar* (Andersson 1922) Polska no 83, after Bleckå Anders Olsson, Orsa. Ornamentation and chords omitted and time signature changed from 3/4 to 16/8. Output from metrical analysis, quarter note designated tactus level.

The metrical analysis provides - in the case that the beat atom analysis finds a consistent meter at central pulse/tactus level - only the central pulse level, here at quarter note level marked by small hooks between the beats in the figure above.

This is the input to the metrical phrase and section analysis. Initially an analysis of possible sequences is performed, which means that all possible sequences are identified and collected. Regarding the section analysis, sequences of the most general type of similarity, X-sequences, are not included, since these are significant only on local level and section analysis regards mainly large scale similarity.

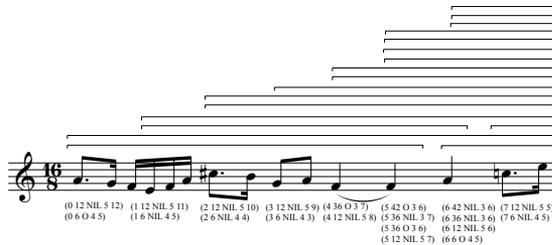


Figure 6-56. Beginning of sequence list for section analysis. Polska no. 83 (see above). Sequences are designated also by sequence code, e.g. (0 12 NIL 5 12), start-position sequence-length sequence-type similarity-level similarity-length. Sequences are displayed under the beat representing the start position.

The sequences obtained by the analysis of possible sequences are then evaluated successively. In this particular case, the first prominent sequence with the highest index-value is (0 12 NIL 5 12) , a perfect repetition of segment. The resulting section-structure would in this case have been identified also without section analysis: The largest major sections of this melody is determined by the sequences (0 12 NIL 5 12) and (24 18 NIL 5 17.)



Figure 6-57. Output from section analysis. Polska no 83.

In the figure 57, the two halves of a sequence determining a section, are designated by brackets with the same letter and number at the left end of each bracket, e.g. A1. Here the two sections are considered different by the subsequent analysis, and are therefore designated by different letters (called *root*). The number 1 means that they are the first members of a similarity family at a particular level.

### Example – Sequence analysis

The next step, the sequence analysis, gets a slightly different input sequence list since extremely long sequences with non-significant similarity are omitted and shorter sequences of more general similarity are added. Sequences interfering with section boundaries are also omitted.

(0 19 X 0 12) (1 6 NIL 4 5) (2 6 NIL 4 4) (4 3 X 0 3) (6 6 O 4 5) (7 5 X 0 5)  
 (0 18 X 0 12)  
 (0 12 NIL 5 12)  
 (0 7 X 0 5)  
 (0 6 O 4 5)  
 (0 6 C 3 4)

Figure 6-58. *Polska no 83*. Sequence list, input to sequence analysis.

Once again the sequence (0 12 NIL 5 12) will receive the highest structural prominence value, this time also supported by already being identified as a section. This implies analysis of symmetrical implication as described in section 6.2.4.4 above. A test of the structural hierarchy at the position of the current sequence results in the deletion of all sequences beginning at position 0, except (0 12 NIL 5 12), (0 6 O 4 5) and (0 6 C 3 4). Implicative duple symmetrical division is recognized by the analysis resulting in Y-sequences at all corresponding positions within the (0 12 NIL 5 12) sequence: (0 6 Y 0 6) and (12 6 Y 0 6). These will in fact not play any role since there are already sequences of these lengths at these positions.

The identification of each sequence results in yet another search for sequences with similarity with that particular sequence allowing identification by similarity to gradually evolve.

The evaluation is continued in this manner resulting in the following graphical analysis after the actual sequence analysis:

Figure 6-59. *Polska no 83*. Output from initial sequence analysis.

This graphical output is based on the list of sequences, which are considered valid by the analysis:

```
(0 12 NIL 5 12)
(0 6 0 4 5)
(12 6 C 3 4)
(24 18 NIL 5 17)
(24 6 NIL 5 6)
(24 3 Y 0 3)
(30 3 Y 0 3)
(36 3 NIL 5 3)
(42 6 NIL 5 6)
(42 3 Y 0 3)
(48 3 Y 0 3)
```

Note that the sequence (24 3 Y 0 3) and corresponding sequences are determined only by symmetrical implication (Y-sequence). It is a result of a sublevel duple symmetry to the general triple symmetry of the second major section.

The graphical output displays the sequence structure at three parallel hierarchical levels. This is quite sufficient with regard to the current melody, but it is very common in longer melodies where the result incorporates up to six simultaneous hierarchical levels.

At each hierarchical level the designation of the sequences (left box) by *root letter* and *identity number* is based on structural similarity *within* the same hierarchical level. A segment which is an A segment on one level may, therefore, very well be a B or C segment at another

hierarchical level.<sup>88</sup>

Lets look at what this designation of identity by similarity implies:

Figure 6-60. Segment similarity at middle hierarchical level (segments of six beats)

Here, the three first different segments are considered similar, but not identical. The first sequence (0 6 0 4 5), exhibits level 4 melodic contour similarity except for the first beat. The sequence A3, which belongs to the B-part of the melody, displays level 4 melodic contour similarity with the A2 segment for the first four beats. Thus, the similarity can be recognized on this level despite the A3 segment belongs to a different section. It can be interpreted as if the B section of the melody is built on the ending of the A section.

The second step in the sequence evaluation involves the evaluation of the implications of the sequence structure obtained. This is implemented as a part of the evaluation of segmentation by discontinuity to which we will return in section 6.2.5.

Here symmetrical implication and implication by periodicity and good continuation are evaluated and the sequence structure is completed. The graphical output (see fig. 61) shows that the three beat segment level is completed by symmetrical implication, and that the six beat segment level is completed by periodic implication. From this output it is possible to make a metrical interpretation at measure level, which would correspond with the original metrical interpretation, resulting in a 3/4 meter at measure level.

The categorization of segment similarity does have its shortcomings and can be developed further. The similarity between the B and C phrases on the lowest level is, for example, not reflected in the naming of the segments. However, the syntactic structure is the main object of the analysis and the similarity is, in fact, recognized in the naming process. It would thus be possible to provide also a tree of similarity for the similarity relationships between segments – general melodic parallelism.

<sup>88</sup> The actual root letter that is chosen depends on how many roots are designated at the same level. However, due to programming problems, this is not always sequential and can be confusing when reading the output.

Figure 6-61. Polska no 83. Final sequence analysis. Output from computer model.

### Global motifs – specific aspects of section analysis

The main object of the section analysis is to reveal the major segments of the melody that rule subordinate sequence analysis. These sections are determined either by major similar segments or by major discontinuity – major interceptions.

The need for a top-down process where sections rule sequences is shown below in fig. 62.

A local sequence with near perfect melodic contour similarity (66 6 NIL 4 5) marked by brackets in fig. 62, would have resulted in a global segment incorporating the two segments if analyzed in isolation. But the two sequence halves does really belong to two different sections since the major repetition starts at the second half of the sequence, which cancels the sequence. We might believe that the whole thing will start over again at the beginning of the sequence, but it turns out that the major repetition starts at the second half of the sequence.

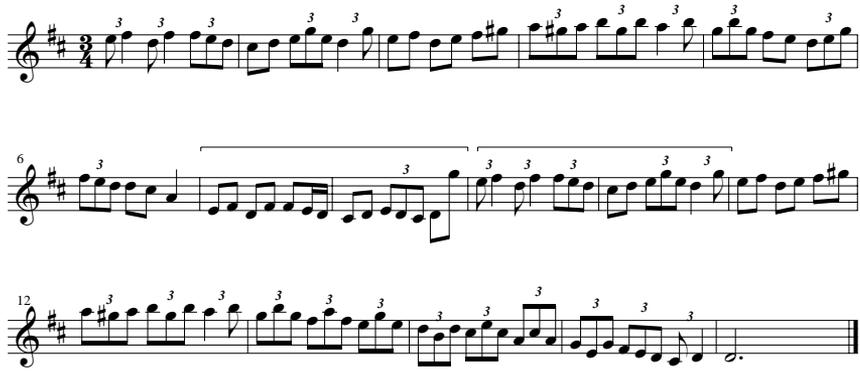


Figure 6-62. Excerpt from *Polska no 4*, Per Danielsson, *Svenska Låtar V*. Local sequence similarity obstructing section boundary marked by brackets.

Figure 6-63. Output from computer analysis (excerpt) of the *polska no 4*

What determines the section is the longer more specific similarity of the section (48 24 NIL 5 13), which incorporates and suspends the implications of the shorter sequence. As can be read from the computer analysis, all three hierarchical levels concur with this division, forming a B1 and a B2 section at the highest displayed level, while the subordinate E1, E2, E3 similarity is displayed at the lowest 2-measure level.

In the above example, the sections were determined by longer sequences – contiguous and adjacent similar segments – incorporating the shorter contiguous similarity, and thus invalidate the shorter sequence as a structural determinant.

But in longer melodic structures one can frequently find similar structures, which are not adjacent, but still function as structural determinants – recurrent melodic structures which signals the beginning of a melodic segment. This kind of structurally significant, but not necessarily adjacent, segments determined by similarity is here designated *global motifs*. I assume that once a segment is determined by e.g. adjacent sequence or discontinuity it becomes a *global motif*. We can use structural similarity between this motif to determine future segments by structural similarity alone, without the conditions regarding structural integrity that applies for segmentation by sequences.

In the example below (Fig. 64, Obligatto melody from ‘*Jesu bleibet Meine Freude*’, BWV 147 by J.S. Bach) the similar sub- and major segments of the melody are displayed in a column.<sup>89</sup>

---

<sup>89</sup> The identification of similar segments is my own. However, this was recognized also by the test participants in experiment 3 (see section 6.3.6.3).

Since this melody is really a part of a polyphonic composition, it is obviously not composed to be experienced as a standalone melody. Thus, the structure of the piece is evidently not only determined by the structure of this *obligatto* part alone. However, it is, in fact, sometimes presented as a melody with accompaniment today.

The actual start positions of the segments are conceived differently by different listeners. This is, however, not important in this context. This problem will be addressed in the evaluation of metrical phrase and section analysis (see section 6.3.6.3)

## Jesu bleibet meine Freude

Obligatto melody JS Bach BWV 147

The musical score for 'Jesu bleibet meine Freude' by J.S. Bach, BWV 147, is presented in a single system with multiple staves. The score is in G major and 3/8 time. The sections are marked with capital letters (A1, A2, A3, A4, A9, A10, A11, A13, A14, A1) and lowercase letters (a1, a2, a3, a4, a5, a6, a7, a8, a10, a11, a12, a13, a14, a15). The sections are numbered with measure numbers. The score is written in a single system with multiple staves. The first staff (A1) contains measures 0-15. The second staff (a3) contains measures 16-32. The third staff (A1) contains measures 33-63. The fourth staff (a3) contains measures 64-83. The fifth staff (A4) contains measures 84-92. The sixth staff (A1) contains measures 93-108. The seventh staff (A2) contains measures 109-125. The eighth staff (A9) contains measures 126-136. The ninth staff (A3) contains measures 137-149. The tenth staff (A11) contains measures 150-152. The eleventh staff (A4) contains measures 153-161. The twelfth staff (A13) contains measures 162-167. The thirteenth staff (A14) contains measures 168-184. The fourteenth staff (A1) contains measures 185-205.

Figure 6-64. Segments with similar beginning in obligatto melody from 'Jesu bleibet Meine Freude' by J.S. Bach (BWV 147). Possible major sections determined by adjacent sequences are marked by capital letters and numbered with regards to general similarity (NIL 5, >50%), single segments with similar beginning are designated by lower-case letters and numbered with regards to perfect similarity. Beats are numbered. (The continuation of the A1 sections are clipped.)

This is an example of a melody where the recurrent similar melodic structure becomes a motif, a signal for a new segment to begin. The recurrence is so frequent, test participants had difficulties in determining a major structure when trying to segment the piece (see section 6.3.6.3 *Experiment 3, Results*).

Still, some major sections can be identified, by adjacent symmetrical sequences, which can be interpreted as sections as opposed to phrases. There are also recurrences, such as the single measure which can be experienced as a false start that is inhibited by the succeeding start of longer similarity.

To determine this structure, starts determined by similarity with the recurrent global motif must not be overridden by global or local adjacent sequences. Therefore, the section analysis determines similar structures, which are considered stable, independent of local structure. Stability is considered to relate to level of similarity and the length of the similar structure, involving the number of beats, interonsets and the duration of the segment. If a structure has high level of similarity (melodic contour, type NIL-level 4 and above) and a length of a 24-beat-scope (see section 5.2.4.7), at least twice the number of interonsets and a duration exceeding two standard perceptual presents ( $> 11.25''$ ) (see Chapter 5, section 5.2.3.2 *Beat-atom analysis*) and involves a complex substructure of at least four sub-segments, it is regarded valid as a stable similarity at section level. This means that it will remain a section start regardless of the adjacent sequence structure except for combinations with segments of the same structural dignity.

In short, an indisputable section must have the complexity of a melody. In the above example there are four such recurrences, marked A1. (The other sections are determined basically by major discontinuity)

The *global motif* similarities below this level are also identified by the model to be of start quality, however not categorically, which is why they can be incorporated in sequences. This level is represented by lower-case letters in the example above.

It is assumed in the model that the identification of similarity is a dynamic process in which comparison is not only performed with the original segment but also with the latest identified similar structure. This makes it possible for the conception of similarity to evolve, creating a motif family in which the most distant members may not be discretely similar.

Consider the similar segments noted above, almost every recurrent segment is slightly dissimilar, either in the beginning or at the end, regarding interval size, melodic direction or length. The model finds this through applying more liberal similarity conditions when a global motif is determined.

The similarity with a global motif is then involved in the sequence evaluation through the structural prominence value.

In the figure below the computer implementation output from the sequence analysis of “*Jesu Bleibet Meine Freude*” shows that the inclusion of lower-level similar phrases in higher-level phrases/sections can be found.

Figure 6-65. “Jesu bleibet meine Freude” (J.S. Bach). Output of computer analysis. Note that major section level analysis is omitted since only three levels can be displayed simultaneously. Some phrases are merged or omitted due to limitations in the graphical interface, which explains why e.g. grand pauses are included in preceding or succeeding phrases.

## 6.2.5 Local phrase boundary analysis – Segmentation by discontinuity, periodicity and symmetry

### 6.2.5.1 Outline

After the segmentation based on melodic parallelism - *the sequence analysis* - the local *phrase boundary analysis*, follows. Local phrase boundary analysis is based mainly on structural discontinuity and implicative segmentation based on secondary grouping principles such as good continuation and symmetry. This part of the segmentation process is fundamentally different from the sequence analysis since segments are determined by structural change or implication of structural articulation. It does not primarily concern grouping by similarity through the clustering of similar events – global segmentation – but rather on structural differences which can be regarded as changes in a continuum – local segmentation.

This part of the analysis can result in reevaluation of segment boundaries, even though the implication of structural change is involved in the sequence analysis. But since discontinuity basically is regarded as subordinate to similarity as a structural determinant, the local phrase boundary analysis primarily involves the segmentation of melodies and melodic sections that are not yet thoroughly segmented.

The local phrase boundary analysis is based on local and global structural changes on different level and yields numerical marker values that correspond to different levels of structural prominence. As in the evaluation of the sequence structure this valuation is based on a set of conditions applied to each beat in the melody. This involves a hierarchy of structural indications where onset is superordinate to offset, changes of duration is superordinate to pitch change, local sequencing is superordinate to local pitch changes, changes of register is superordinate to changes of melodic direction and composite changes are superordinate to simple changes. In fact, this valuation of structural change is performed by exactly the same set of rules that is involved in the complex metrical analysis (see Chapter 5, section 5.4.2.6 *Analysis of Global Discontinuities*).

Based on this valuation of structural change at every beat and implication of segmentation based on periodicity and symmetry, the different possible segment boundaries are evaluated successively with regard to congruence of structural hierarchy, symmetry, good continuation, and *prägnanz*/salience and are valued in relation to the segmentation by sequence structure.

Finally, the segmentation is evaluated with respect to consistency, hierarchical congruence and the similarity between segments is evaluated once again based on all segments, whether or not they are obtained by sequence analysis or local phrase boundary analysis. This valuation of similarity includes both successive and absolute similarity between segments.

The evaluation of hierarchical congruence results in a grouping of segments into different levels based on the general assumptions of structural hierarchy (see section 6.2.4.3 *Quantifying structural hierarchy*).

Phrase- and sections are named after basic similarity (*root/family*) and perfect similarity (number/variant), e.g. A1, A2 etc. In the computer output phrases are indicated by brackets above the system with the phrase ID to the left end of the bracket. The different levels are displayed by vertical placement above the system (see e.g. figure 65 above).

### 6.2.5.2 Method of local phrase boundary analysis

#### Design

The actual local phrase boundary analysis is basically divided into two steps, finding indications of segmentation and the evaluation of these indications.

There are three types of indications of segmentation that are acknowledged in the analysis. These are:

1. Implication of segmentation by periodicity of segmentation from sequence analysis
2. Implication of segmentation by hierarchical symmetry
3. Indication by local and global discontinuities

As have been mentioned above the result of the sequence analysis is input to the local phrase boundary analysis. On the basis of this analysis, the implications of periodicity are implemented, the fundamental assumption given that periodicity is a secondary grouping principle which is significant within the limits of periodical conception when the basic metricity exists. This limit is generally within the method set to a maximum of 48 beats at central pulse/tactus level. This implies that when a certain segment length is subsequently repeated more than once it is assumed to create an expectation of segmentation at subsequent positions as long as it is not contradicted by interfering segmentation by the existing sequence structure.

The second implication of hierarchical symmetry is based on the assumption that when symmetrical hierarchical segmentation exists at a superordinate level the listener will expect symmetrical segmentation on subordinate levels. Thus, if the sequence analysis provides such evidence at one level, symmetrical segmentation is implicated on subordinate levels as long as the existing sequence structure does not contradict this.

The indication of segmentation is, as mentioned above, based on analysis of local and global discontinuities at beat positions. The conditions which govern valuation of indication of local segmentation are demonstrated in *Metrical Analysis*, Section 5.2.4.6. *Analysis of local discontinuities*, and the conditions of valuation global segmentation are described in section 5.2.4.6 *Analysis of global discontinuities*. However, since metrical analysis mostly concerns the level of primary grouping resulting in beat-grouping and phrase and section analysis relies on primary beat-grouping and concerns complex grouping, they typically employ a different set of conditions.

A segment or section is recognized as being possible to subdivide into sub-segments when it can be subdivided into a minimum of two groups of at least two beats per group. However, when the sequence structure does not indicate a grouping level below four beats per segment, a subdivision will not be performed.

When a segment is recognized as possible to subdivide into sub-segments, the different possible segmentation division points are given a structural prominence value; *division-index*, based on the above indications, 1) the symmetrical qualities of the division, 2) the consistency of divisions (the relationship to previous divisions), 3) the congruency between the discontinuity values of the segments and the integrity of the segmentation indication with regards to discontinuity values of the segments boundaries in relation to the discontinuity values within the segment. Of these, the last quality is of crucial importance, since the segmentation must be discrete.

**Example**

Consider the following example (fig. 66). This shows parts of the result of the sequence analysis for “*Raag Suddha Danyasi*” (Shivakumar 1992). As can be read from the example, the sequence analysis does not provide a segmentation of the entire piece of music. Quite a large portion of the melody is not segmented and mostly larger segments are identified by the analysis.

The image shows a musical score for Raag Suddha Danyasi in 140 bpm. The score consists of eight staves of music. Above the staves, there are several annotations indicating segmented sections:
 

- A box labeled 'A1' spans the first staff.
- Boxes labeled 'A2' and 'A3' are positioned below the first and second staves, respectively.
- A box labeled 'A1' is positioned below the third staff.
- A box labeled 'A1' is positioned below the fourth staff.
- Boxes labeled 'C1' and 'C2' are positioned above the fifth staff.
- Boxes labeled 'A1' and 'A2' are positioned below the sixth staff.
- A box labeled 'A1' is positioned below the seventh staff.

 The music is written in a single melodic line on a treble clef staff with a key signature of one flat (B-flat). The notation includes various rhythmic values and accidentals.

Figure 6-66. Output from sequence analysis of Raag Suddha Danyasi (Shivakumar, K. 1992).

Among the missing segments is the very beginning of the piece. However, the start of the next segment is defined by the sequence analysis.

The local phrase boundary analysis therefore considers the unsegmented part to the first sequence start to be the subject of segmentation, either as a part of a segment including the subsequent sequence or as a segment of its own with possible sub-segments.

In the figure above, the length of the unsegmented start is equal to the length of the following sequence half. Thus the unsegmented part can be considered a segment at this level, balancing the following sequence half. Since there are no prior segments, there is no implicative segmentation, and potential sub-segmentation must rely on the valuation of discontinuities.

In the figure below (fig. 67) the discontinuity values are displayed below the system. The interonset distance between beats two and three indicates a possible sub-segmentation into two segments of equal duration.

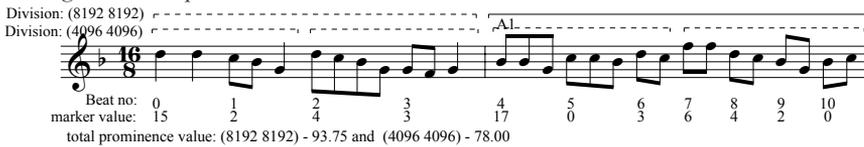


Figure 6-67. Possible segmentation of the initial part indicated by dotted brackets. Existing sequence indicated by line brackets. The beats at central pulse/tactus level are numbered and displayed below the system. The discontinuity values are displayed below the beat numbers.

This subdivision gives two discrete sub-segments with discontinuity values (15 2) and (4 3) delineated also by the subsequent event with discontinuity value 17. The entire segment is also discrete, delineated by greater discontinuity values at the start, mid and end position than within the segment: (15 2 4 3) (17 0 3 6 4 2 0) (15...).

Based on the boundary values – the discontinuity value integrity, hierarchy and symmetry factors – total prominence values for the possible segments are calculated. These values are used in the evaluation of possible segments for discriminating significant segments from non-significant. In the present example, both segmentations are regarded significant.

The segments obtained by the analysis of the initial segment are used in the analysis of subsequent segments. Thus, the two-fold division of the initial segment is also applied to the second major segment.

The fundamental assumption behind the metrical implication is that when a sequence period is repeated more than once it gives rise to an expectation of periodical segmentation. In the figure below (fig. 68) the implicated segments by metrical implication or implication by continuation of previous segmentation are indicated by dotted brackets.

## Raag Suddha Danyasi

after K Shivakumar

Division: (8192 0)  
 Division: (4096 4096)

Beat no: 0 1 2 3 4 5 6 7 8 9 10  
 marker value: 15 2 4 3 17 0 3 6 4 2 0

3 A2 A3  
 11 12 13 14 15 16 17 18 19 20 21 22 23 24  
 15 4 7 2 0 0 2 19 4 3 2 4 6 3

5 A1  
 25 26 27 28 29 30 31 32  
 25 2 4 1 13 0 1 9

7 A1  
 33 34 35 36 37 38 39 40  
 21 2 4 1 13 0 1 3

9 C1 C2  
 41 42 43 44 45 46 47 48 49 50 51 52  
 21 5 15 0 13 5 14 7 7 1 1 1

11  
 53 54 55 56 57 58 59 60  
 10 2 8 3 10 2 3 1

13 A1 A2  
 61 62 63 64 65 66 67 68 69 70  
 25 1 0 17 0 0 19 3 9 1

Figure 6-68. Implicated or discontinuity based segments marked by dotted brackets. Sequences marked by line brackets. Note that implicative periods are introduced successively based on prior sequences.

Note that when a different segment period is determined by the sequence structure (e.g. measure no 9) or by discontinuity factors (e.g. measure no 10), the metrically implicated segment period is inhibited. But once a metrically implicated period is determined and established, it is assumed to be inherent in the structure and easy to re-establish (see e.g. measure no. 11).

danyasi.tot.mf

The image shows a musical score for Raag Suddha Danyasi in 16/8 time, marked with a tempo of quarter note = 140. The score consists of eight staves of music. Above the staves, there are numerous annotations consisting of small boxes containing letters and numbers (e.g., A1, B1, C1, D1, B2, D2, D3, B3, F1, G1, H1, I1, J1, K1, C1, H1, I1, J1, N1, N2, D1, K1, G1, H2, L1, P1, Q1, E1, M1, R1, L2, P2, S1). These annotations are connected to the musical notes by horizontal lines, indicating the results of a computer model analysis, specifically local phrase boundary analysis.

Figure 6-69. Output from computer model analysis including local phrase boundary analysis performed on Raag Suddha Danyasi

In the example above, there is no symmetrical implication since hierarchic symmetry is not established by the sequence analysis. Symmetrical implication does occur only when hierarchical symmetry is established by sequence structure as in the example below (Figure 70).

Figure 6-70. Excerpt from *Polska no 83, Svenska Låtar I* (N. Andersson 1921). Sequences are marked by line brackets, while implicated segments by symmetry are indicated by dotted brackets.

Here, a symmetrical hierarchic structure is established, while the principal sequence is divided into two subordinate sequences of equal length. This symmetrical hierarchy implies a general twofold sub-segmentation in equal parts.

Figure 6-71. Output from computer analysis including local phrase boundary analysis. *Polska no 83, Svenska Låtar I* (N. Andersson 1921).

## 6.2.6 Final analysis – start- and end-oriented interpretation

The final analysis of metrical phrase and section analysis regards the identification of segments by similarity, completion of phrase structure and finally the creation of an end-oriented interpretation of segment structure.

The identification of segments by similarity involving the categorization of similarity into root/family and variant designated by root-letter and variant-number has been described in section 6.2.3, *Sequence analysis*. The identification of similarity between segments in general is performed by a similarity rating involving all segments in the melody at the same hierarchic level. The root/family letter is defined by the position of the first appearance of the root in its context. The similarity evaluation involves general non-successive melodic contour similarity on C-sequence similarity level (see section 6.2.3.3 *C-sequences*). This implies that the total similarity value is the basis of the similarity rating. However, factors determining the level of similarity, i.e. specificity of similarity, initial and successive similarity, are also involved in the similarity rating, allowing the similar segments to be ranked from the most similar to the least similar segment.

Pure rhythmic similarity is thus not recognized as a factor that determines identification of segments on an inter-contextual level. This is because of the assumed lack of possibility to address inter-contextual identity due to the one-dimensionality of rhythmic similarity. (see section 6.2.3.3 *T-sequences*).

Neither is general global change similarity (X-sequence similarity) recognized in the identification process. Since X-similarity lacks local and specific similarity, it is assumed to be impossible to trace inter-contextually.

Completion of phrase structure involves making each hierarchical segment level complete, basically by symmetrical implication and implied periodicity performed in essentially the same way as the local phrase boundary analysis (see section 6.2.5, *Local phrase boundary analysis*).

The last step in the final analysis of metrical phrase and section analysis is the mirroring of the default start-oriented interpretation to an end-oriented interpretation. The start-oriented interpretation is default because the inclusion of a pickup or an anacrusis in segments can be subject to variation without having any significant impact on the structural and syntactic relationships between segments. In turn this is a consequence of the metrical structure, whenever present, is the primary level of structural organization of time (see section 6.1.2 *Start- and end-oriented grouping*). Start-oriented interpretation hence considers the syntactically significant level of segmentation.

End-orientation interpretation is, however, often perceptually more significant. Lerdahl & Jackendoff argues in “A General Theory of Tonal Music” (1983) for the distinction between meter and grouping by an example in which an end-oriented grouping does not concur with the metric structure (see also Chapter 5, *Metrical analysis*, 5.2.4.2, *Analysis of low-level grouping*.) In their principal example, a parallel metrical and grouping analysis of the initial theme from *Symphony no 40* (K550), by W.A. Mozart, the grouping interpretation is consistently end-oriented.<sup>90</sup>

The image shows a musical staff in G minor, 2/4 time, measures 1-14. Above the staff, horizontal lines indicate the 'Grouping structure', showing phrases that are end-oriented. Below the staff, dots indicate the 'Metrical structure', showing a regular 2-beat pattern. The notation includes a treble clef, a key signature of two flats, and a common time signature.

Figure 6-72. Analysis of meter and grouping structure in the opening theme of Symphony no40 in G minor, K550, 1<sup>st</sup> movement by W.A. Mozart, adapt. from Lerdahl & Jackendoff (1983:27).

In an experiment the subject of which was to determine the listeners conception of central pulse level<sup>91</sup> the participants was to notate the rhythmic structure of the opening melodic line (measures 1-14, tempo: ♩ = 120) of Mozart’s G-minor Symphony. The subjects were given a score displaying only the note heads and were instructed to complete the score with the rhythmic and metrical notation. The stimulus was dead-pan performance of the score produced by a notation software. The test results of this test is summarized in the figure below:

<sup>90</sup> Note that I do not object to the separation between grouping and meter as concepts.

<sup>91</sup> This test included in Experiment 2 (see Chapter 5, Metrical Analysis, section 5.3.4.3). The test participants and the apparatus were identical to the one used in Experiment 2.

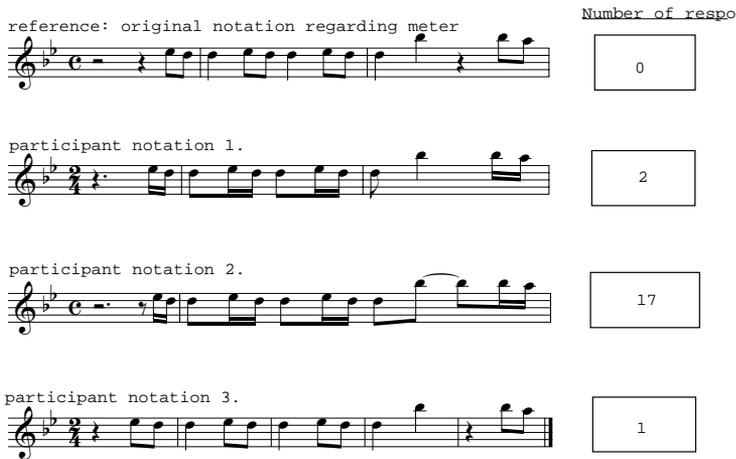


Figure 6-73. Results from test of notation of musical meter and rhythm in the opening. Provided notations are shortened.

As the summarized result shows, most of the test participants did not notate this melody in accordance with original score (even though one of the test participants had actually performed the piece). Given the relatively high tempo which is typical of performances of this piece the majority of the test participants seem to have regarded the half note-level of the original notation to be the central pulse level.

Given the subsequent task to apply grouping indications none of the participants notated the lowest group level suggested by Ler Dahl & Jackendoff, but favored the third and fourth levels. Even if the use of notation as a means of indicating conceived grouping is problematic, this can be interpreted as if the conception of grouping primarily regards levels above the tactus for these test participants. This supports the assumption of the metrical grid at central pulse level provides the fundamental temporal mapping of the music.

As was suggested in section 6.1.2, the start-oriented grouping thus can be regarded the most structurally determinant grouping level, which means that the inclusion of an upbeat or not will not fundamentally alter the structural conception. Therefore, the model makes use of start-oriented grouping in the analysis of phrase and section structure. End-oriented grouping output is within the model derived from the result of the start-oriented grouping.

A start-oriented grouping can be viewed by the default output of the computer model of the current method.

Figure 6-74. Start-oriented interpretation of the initial melodic theme from Mozart G minor symphony. Computer output.

Note that the lowest phrase level (2/4 level) in this example corresponds to the tactus beat level of the test result notation exemplified above. As suggested this level is possibly not perceived as a typical phrase level because of the small number of events per group and low structural integrity in the latter part of the melody.

Even if such start-oriented interpretation is possible on all levels, the connection between the two eighth notes preceding the first quarter note in the first melodic phrases is obvious through interonset proximity, preceding interonset distance and repetition of the pattern. The end-oriented interpretation suggested by Lerdahl & Jackendoff would thus probably be more perceptually evident to Western listeners. The model is designed to provide both the start- and end-oriented versions, but in this case the default interpretation is the end-oriented due to the repeated upbeat. The end-oriented interpretation of the same segment analysis can be viewed below.

Figure 6-75. End-oriented grouping interpretation of the initial melodic theme from Mozart's G minor symphony. Computer model output

This interpretation, which largely resembles the interpretation provided by Lerdahl & Jackendoff, is obtained by checking the neighboring events to each phrase boundary for largest interonset distance, largest pitch distance and onset-offset indication. The position of the highest structural distance and most consistent relationship to the start-oriented grouping is then chosen as the segment boundary. Thus, the end-oriented grouping is the complementary inversion of the start-oriented grouping.

The results of the experiment 3 referred to in the evaluation of metrical phrase and section analysis (see section 6.3.6.3) indicates that a valid prediction of the perceptual salience of start- versus end-oriented grouping can be impossible to make even for a group of listeners with no significant differences in cultural background.

Lerdahl & Jackendoff, who are actually trying to present the culturally informed listener's most plausible interpretation of grouping, still accept that grouping conception can vary among listeners. (see Lerdahl & Jackendoff 1983:63) In the case of the following melody from *Mozart A major Sonata K. 331* they present two alternative groupings

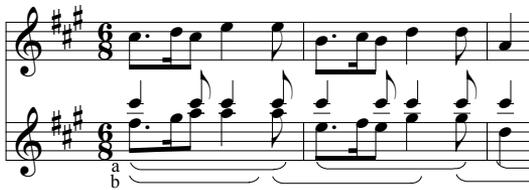


Figure 6-76. "... it is supplied with two possible groupings. (We favor grouping a, but grouping b has not been without its advocates; see Meyer 1973)." (Lerdahl & Jackendoff 1983:63)

These two interpretations corresponds to the computer models start- and end-oriented grouping interpretation, displayed below.



Figure 6-77. Start- and end-oriented grouping interpretation performed on the initial melodic theme from *Mozart's Piano Sonata in A major K. 331*

Chapter 6

The figure shows a musical score in 3/8 time with a key signature of one sharp (F#). The melody is written on a single staff. Above the staff, several horizontal lines represent chord groupings. The first line has a box labeled 'A1' spanning the first six measures. The second line has boxes for 'A1' (measures 1-3), 'B1' (measures 4-5), and 'A1' (measures 6-7). The third line has boxes for 'A1' (measures 1-2), 'A1' (measures 3-4), 'B1' (measures 5-6), 'C1' (measures 7-8), 'A1' (measures 9-10), and 'D1' (measures 11-12). Below the staff, there are two empty staves, with a box labeled 'C1' positioned between them, spanning measures 7-8.

Figure 6-78. End-oriented interpretation

However, in many cases start- and end-oriented grouping concur. This is, for example, true for the *Polska no. 83* (Svenska Låtar, Andersson 1922), which was used as example of the successive evaluation of sequence analysis (see section 6.2.4.5)

The figure shows a musical score in 3/8 time with a key signature of one sharp (F#). The melody is written on a single staff. Above the staff, several horizontal lines represent chord groupings. The first line has boxes for 'A1' (measures 1-6), 'A1' (measures 7-10), 'A2' (measures 11-14), and 'A1' (measures 15-18). The second line has boxes for 'A1' (measures 1-2), 'B1' (measures 3-4), 'A2' (measures 5-6), 'B2' (measures 7-8), and 'A1' (measures 9-10). Below the staff, there are two empty staves, with boxes for 'B1' (measures 3-4), 'A2' (measures 5-6), 'B3' (measures 7-8), 'A3' (measures 9-10), 'C1' (measures 11-12), and 'A2' (measures 13-14). The third line has boxes for 'D1' (measures 15-16), 'C1' (measures 17-18), and 'D1' (measures 19-20). The fourth line has boxes for 'B1' (measures 1-2), 'A3' (measures 3-4), 'A3' (measures 5-6), 'D1' (measures 7-8), 'A3' (measures 9-10), 'C1' (measures 11-12), and 'D1' (measures 13-14). The fifth line has boxes for 'D1' (measures 15-16), 'C1' (measures 17-18), and 'D1' (measures 19-20). The sixth line has boxes for 'B1' (measures 1-2), 'A3' (measures 3-4), 'A3' (measures 5-6), 'D1' (measures 7-8), 'A3' (measures 9-10), 'C1' (measures 11-12), and 'D1' (measures 13-14). The seventh line has boxes for 'D1' (measures 15-16), 'C1' (measures 17-18), and 'D1' (measures 19-20). Below the staff, there are two empty staves.

Figure 6-79. Identical start- and end-oriented grouping interpretation in *Polska no. 83* (Svenska Låtar, Andersson 1922)

## 6.3 Evaluation of metrical phrase and section analysis

### 6.3.1 The means of evaluation

#### 6.3.1.1 General problems

The evaluation of the method of phrase and section analysis is, for several reasons, even more difficult than evaluation of metrical analysis; While meter is usually notated in some way or another, the segment structure is seldom notated in scores or transcriptions in any elaborate or consistent way. While physical response to metric structure as well as metrical accompaniment is common to people from different musical cultures and thus can be used in experiments, there are seldom any evident social or cultural control systems, natural physical responses or unambiguous performance codes connected with the conception of segment structure in melodies. (see also section 6.2.6, *Start- and end-oriented grouping*).

This common lack of control systems connected with segment structure may lead to greater variability in peoples' conceptions of segment structure than that of e.g. metrical structure. The lively discussions about the performance interpretation of musical works in almost all musical tradition that I have come across, suggests that this is true. Musical structure in general can be conceived differently by people belonging to the same cultural sphere.

This issue is not always addressed in the older literature about melodic segmentation, reflecting the view of the analyst as the ideal listener, but has become increasingly more noticed by scholars:

*“One difficulty with studying grouping lies in determining the “correct” analysis – the one listeners would perceive – for a given input. With meter, the correct analysis for a piece is usually clear, at least up to the level of the measure. With grouping, the correct analysis is often much less certain...The problem is compounded by the fact that our intuitions about grouping are often ambiguous and vague, much more so than meter.” (Temperley 2001:60)*

*“Although we are unaware of any past research that directly supports this claim, melodic segmentation is most certainly an ambiguous affair. Rather, the process is influenced by a rich and varied set of contexts, motives, harmony, melodic parallelism), where local structure (gestalt principles), higher-level structure (e.g. recurring motives, harmony, melodic parallelism), style-dependent norms and the breaking of these norms all have an impact on what is perceived as “salient”...” (Hötker et al 2002a: 168)*

The goal of the evaluation of the metrical phrase and section analysis - metrical segmentation - is to establish, if this method is able to provide plausible results of melodic structure, which may reflect people's conceptions of melodic structure better than chance. This goal is quite problematic because of the above noted general lack of control systems, musical as well as extra-musical.

#### 6.3.1.2 Comparisons with expert analyses and score information

The most elaborate existing sources of segment conception in melodies come from expert analyses, analyses made by music researchers *'by hand'* often according to some method

of structural analysis. The nature of these sources makes it hard to draw conclusions about people's conception of segment structure from listening experience in general. Still I am going to use such sources in the evaluation of the generality of the method. I am using some particular examples of segmentation that originate from the literature of ethnomusicology, systematic musicology and studies in music cognition and perception (see sections 6.3.3 - 6.3.6). But I am also making a comparison of the results with a larger corpus of analyses made by a single analyst, Einar Övergaard, a Swedish folk music researcher, who made analyses of about 3500 Swedish folk melodies, (Övergaard 1935, MS).

Ordinary transcriptions and notations of music can also provide some information about e.g. major structural elements such as recapitulations and sections. I have used this kind of information, for example, in the analyses of the South Indian and Balkan melodies as well as some melodies from the area of Western art music.

Obviously I have also used my own judgment as to how musically plausible a certain phrase structure seemed to me during the development of the method. However, I have learned from experience that introspection can be a rather problematic method when it comes to predictions about the conceptions of other listeners.

### 6.3.2 Listener tests

The most valuable sources are perhaps listener experiments, since if they are well designed might reflect different listeners various conceptions of melodic structure. A traditional musical analysis based on score analysis, may be influenced by the application of a theoretical methodology or visual experience of melodic form.

However, the problem with listener experiments regarding melodic form is how to design them? My first approach was to ask the listeners to tap on an electronic keyboard when there was a segment boundary in the melody. This methodology was used by e.g. Deliège (1987) with good results.

But in my pre-test this turned out to be a problematic method for several reasons: Since the conception that a segment start had occurred was realized at the time of the segment start or after, the response was generally too late to reflect the conception of the listener. To remember where the taps should be performed throughout a second listening to correct the mistakes of the first turned out to be difficult. Further, an experience of a hierarchic segment structure was not easy to show through a tapping. Last, but not the least, tapping was felt by the listeners to be an appropriate behavior/response regarding metrical experience but not for the experience of melodic form. It was a likewise difficult not to tap on beats or measures.

Krumhansl (e.g. 1996:401-432) has developed a method by which subjects respond via a computer interface, which have proved to be very successful. Unfortunately I was not aware of the method used by Krumhansl at the time of the experiments. The results reported Krumhansl suggests that it would be possible to further develop experimental equipment for the purpose of investigating people's conceptions of musical surface structure - perhaps resembling an advanced control station or joystick - with which people could give physical multidimensional response to form conception in a measurable way.

Because of the reasons discussed above, I tried two other experimental designs, *response by notation in scores* and *response by valuation of different performances* constructed as to reflect different structural conceptions of the melodic structure. The first method limits the group of possible subjects to people skilled in reading music, and adds an uncertainty factor in the unknown

influence of the individuals' skill of reading music. It also introduces the problem of involving a visual representation of music, which may (and probably will) influence the results. I tried to limit the influences of the visual representation by presenting the notes of the melody evenly distributed with chance distribution in relation to anticipated grouping together with the sound of a deadpan performance of the melody.

The great benefit of the visual response is that conception of hierarchical melodic structure is easy to represent visually. Also it makes it possible to represent ambiguity in melodic structural conception.

However, the limitations of using score representation of the music are severe, basically because it excludes people unskilled in reading music from the experiments. But, even for educated persons participating in the tests, the task was obviously difficult and exhausting.

The other method was to present listeners with different interpretations of the same melody designed to correspond to different segmentations and make the listeners rank the performances with regards to how well the performances matched their conception of melodic structure. The general methodological problems concerning this design are discussed in section 5.3.4.1, *Evaluation of metrical analysis*, but, in the case of evaluation of segmentation, the difficulty of representation of structural hierarchy may be the most important problem.<sup>92</sup>

However, the combination of these two experimental methods together with both comparisons of the results of the analyses with structural analyses made by experts and structural conceptions reflected in notations of music may give a general measure of the generality of the method.

### **6.3.2.1 How to evaluate the success of the model in relation to listener segmentations?**

A general problematic issue common to both comparisons with expert analyses and results from listener tests is how to evaluate the success of the model in relation to the results from manual analyses.

This issue has recently been discussed by Thom, Spevak and Höthker in a series of articles/papers. (Höthker et al 2001, Thom, Spevak and Höthker 2002a, 2002b, 2002c). They have performed a systematic evaluation of the results of two computer-based models of melodic segmentation in relation to segmentations made by different musicians on a certain corpus of melodies. Their interesting approach involves addressing the question of ambiguity in segmentation.

*“... the data clearly corroborates the hypothesis that ambiguity is an inherent property of the segmentation task and should not be ignored. Even for the “simple” folk song displayed in Fig. 1, nineteen musicians produced nine different segmentations on the sub-phrase level. In a more complex excerpt (Fig. 3), the number of segmentations rose to eighteen (only two excerpts was unanimous agreement of the nineteen musicians gave the same answer). In none of the ten melodic excerpts was unanimous agreement obtained.” (Höthker et al 2002a:171ff)”*

---

<sup>92</sup> Also regarding these experiments it would be possible to develop a design that would give more evident results, e.g. by letting the subjects interactively make a virtual segmentation of the sounding music. This would allow for hierarchical interpretation.

As will be apparent from the results of experiment no. 3, (see section 6.3.7.3), variability in the segmentations is typical also for the results of my experiments.

Thom, Spevak and Höthker point out some causes for these differences in the subjects responses:

*“One cause for this ambiguity concerns granularity – one musician’s notion of a phrase might more closely coincide with another’s notion of a sub-phrase – yet in terms of identifying “locally salient chunks,” both are musically reasonable. In other cases, musicians might agree on a phrase’s length, but disagree on location, and in this situation, ambiguous boundaries are often adjacent to one another. (Höthker et al 2002a:171ff)”*

In their proposed model of calculation of a fit value for a certain segment boundary they are incorporating the test groups preferences for positions of segment boundaries in connection with segment length between adjacent pairs of boundaries and the total number of segments preferred.

Since this rather sophisticated model is based on somewhat different presumptions and different input<sup>93</sup>, it cannot be directly applied to my experiment. But it points out some important aspects of an evaluation of segmentation: (1) The positions of segment boundaries cannot be treated independently of one another, but (2) segment length, (3) phase of segment and (4) hierarchical level have to be incorporated in the evaluation.

I have not used any comprehensive measure of fit of the models segmentation in relation to manual, but rather focused on the model’s *ability to account for the different manual segmentations* by a few, rather simple statistical measures that involve the above aspects.

### 6.3.2.2 Outline

In the following, I will present the results of the models behaviour with regards to different melody material focusing on different problems of segmentation. The comparisons with manual expert analyses of melodic structure will be presented in connection with the repertoire it concerns.

## 6.3.3 Phrase structure in Swedish instrumental folk tunes

### 6.3.3.1 General problems

Folk music is often used in studies of musical structure as an example of simple and straightforward musical structure. But even if many folk music melodies, such as e.g. many folk songs, have the analytical benefit of being relatively short in relation to e.g. symphonic works, the structure is not always so unambiguous or simple, even in the shortest melodies.

The probably most common basic form type in Swedish instrumental folk music can be described as consisting of two major repeated sections – reprises/turns – of eight measures. Each of these is divided into two major (generally repeated) super-phrases of four measures which are, in turn divided into two phrases of two measures at what is often considered the

---

<sup>93</sup> It is e.g. based on a note-list instead of beat list as input; it does not explicitly allow simultaneous parallel hierarchical levels but assumes an ideally one-dimensional segmentation: *“to focus on breaking a melody into a segment stream that contains a series of non-overlapping, contiguous fragments.”* (Höthker et al 2002a:168)

central phrase level. A sub-phrase level of one measure is sometimes also recognized, depending on the metrical structure of the tune. Generally, phrase structure is conceived to concur with the metrical structure, i.e. phrases begin at beats and on measures and can be thought of as combinations of measures as the basic unit. (see e.g. Övergaard (MS) 1935)

But even if there are plenty of examples of this type of consistently symmetrical, hierarchical structure in the corpus of Swedish instrumental folk music, examples of more complex or quasi-symmetrical are probably just as common. Moreover while there are no overall statistics of the aforementioned form analyses of *Polska* tunes by Övergaard, the great number of various form schemata he describes, suggests that other forms in fact dominate the repertoire. To account for this variability actually demands a comprehensive model.

There are some typical problems regarding segmentation of Swedish folk tunes that I have encountered:

1. Low level of local structural contrast. The continuous flow of events with relatively few long rests or durations make discontinuity indications and boundary positions weak. This particular repertoire shares this structural feature with many other instrumental folk music repertoires; It is, for example a typical characteristic of many dance music styles in Europe and Asia.<sup>94</sup> The generally rich ornamentation and floating boundary between ornamentation and melody makes this problem even more manifest with audio-to-midi transcriptions.
2. Disguised segment boundaries, not coinciding with tone onsets. This is a typical feature of performance style found in many sub-styles within e.g. Swedish folk music, which can make it difficult for listeners not familiar with the style to conceive the melodic and metric structure when first listening to it. This is evident in brief score notations by frequent “ties over bar lines”. This stylistic feature can also be found in many different styles, which sometimes can make it impossible to identify the culturally dominant metrical interpretation and phrase structure from the melody alone. (See e.g. the discussion of meter in Norwegian *Gangar* and *Halling* melodies in section 5.3.4.3 *Discussion*). Such ambiguity regarding meter and phrase structure may be a valuable feature in dance music styles, since it can stimulate the interaction between music and dance movement.
3. Sequences determined by brief/vague similarity. The typical structural features of a musical style may sometimes provide such strong expectations of a certain structure, that vague inherent structural melodic cues can still be recognized. The two- to three-measure phrase/sub-phrase level is such a typical feature of Swedish instrumental folk tunes. There may be a level of similarity at which a sequence-determinant similarity will not be recognized by a person not familiar with the style. For the model to adapt to such style-dependent idiosyncrasies would require the model to be style-perceptive and this is precisely what it is not intended to be (see discussion in Chapter 2). The learning of the model is restricted to a melody.
4. Asymmetrical segment lengths within a generally symmetrical framework. Quasi-symmetrical segmentation seems to be very frequent within certain individual and local repertoires. There might either be a general level of symmetry with

---

<sup>94</sup> In e.g. the instrumental folk music styles of the Balkans and Middle East this is even more apparent.

asymmetrical deviations at a lower level or perfect symmetry at a lower level within an asymmetrical structure at section level. The lack of a culturally determined formal method of composition or interpretation makes the variability in this matter considerable.

5. Variability concerning the degree of structural hierarchy. Quasi-hierarchical structure makes the implications of hierarchy ambiguous. This sometimes relates to the adaptation of a strictly hierarchical (and symmetrical) melody of the European popular repertoire, which have been interpreted non-hierarchically, hence giving ambiguous melodic structural cues.

In general, phrase structure is surely sometimes ambiguous up to a level where it becomes almost impossible to make predictions about people's conceptions of phrase structure. But I will return to this question of predictability in the discussion of the results of the listener tests (see section 6.3.7.3).

In the following sections, I will address some of these problems by actual examples in order to demonstrate the performance of the model. I will basically use examples from the repertoire of one fiddler, Per Danielsson (1823-1922) from Mörsil, Jämtland and, in one case, from his pupil Bengt Bixo (1879-1962), both published in *Svenska Låtar* (Andersson 1926) in order to illustrate the great variety of structural problems within even a limited repertoire and well-defined local style.

### 6.3.3.2 Structural ambiguity by asymmetry within a symmetrical general structure

#### Example. Polska after Per Danielsson

Indication of asymmetrical sub-structures within a symmetrical general melodic structure is, as mentioned above, a cause of structural ambiguity in many Swedish folk tunes. The following example may serve to illustrate this phenomenon.

10. P O L S K A

*A-bas.*

Figure 6-80. Polska no. 10. *Svenska Låtar*, Jämtland (Andersson 1926:11). Played by Per Danielsson, Mörsil, Jämtland.

At the section this tune conforms level to the most common structure of *Polska* melodies, which can be characterized as two major repeated sections, two turns/repeats. These are determined by quite evident structural indications (according to this notation), perfect repetition and strong interonset contrast at the endings, i.e. tones of long duration as opposed to the generally short tones within the sections. This is thus a typical example of the generally low contrast regarding interonset durations, which characterizes most of this melodic style.

According to the assumption of hierarchical consistency of symmetrical implication (see section 6.2.4.4, *Segmentation by symmetrical implication*), symmetry on one level implies symmetry on subordinate and superordinate levels. This means that, once we recognize a symmetrical relationship regarding length, we assume this to be a consistent structural feature of the entire structure. In combination with the principle of hierarchical structure, hierarchically subordinate levels to the symmetrical level receive stronger symmetrical implication than superordinate levels.

If this assumption holds, the repetition of a major segment would make us expect symmetry on lower levels. The duple division implied by the repetition of the first major section would hence show on the lower levels and so forth.

### Section I. Inherent ambiguity on sub-phrase level

If we consider the first major section of this melody, we will find such a hierarchical perfectly symmetrical division determined by similarity:

beat pos. 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17  
 disc. values (0 18 C 3 18) [A]  
 (0 9 Y 0 9) [X]  
 (0 6 C 3 3) [A] [A]  
 (0 6 X 0 6) [X] [X]  
 (9 3 X 0 6) [X] [X]  
 (12 3 NIL 4 2) [B] [B]

7  
 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33  
 [A]  
 [X]  
 [A] [A] [X] [X]

Figure 6-81. The first section of *Polska* no. 10. Global discontinuity indications marked by shapes above the systems and sequences displayed below the system. (see explanation below)

The sequence (0 18 C 3 18) demonstrates a general duple division of the section, which means that, on a non-specific level of melodic contour and rhythmic similarity, the section consists of two entirely equal parts. In this case, there are only minor local differences between the two parts.

Below this level we cannot find any entire duple division of the super-phrases determined by similarity, and now the structural indications starts to be more ambiguous. The first melodic repetition to be found below the super-phrase level is represented by the (0 6 C 3 3) sequence, which indicates general similarity on C-level regarding of three beats out of six. This similarity concerns the two first changes between the initial tones of the beats, which makes a up-down return to a close tone:



This sequence similarity is extremely brief and would not be recognized by the model if it were not supported by general global structural changes concurring at the boundaries of the sequence. This is manifested also by a sequence determined by compatible global changes (0 6 x 0 6) at this instance. The arrow above the system in the figure 82 shows that the mid-position of the sequence, beat position 6, is considered to be the ending of a gradually falling melodic line at global level. At the end position of the sequence, beat position 12, there is also a turn of global melodic direction which is clear and significant. In addition, there is a sequence determined by high level similarity, (12 3 NLL 4 2), beginning at the end-position of the sequence.

Thus, segmentation is supported by vague similarity regarding melodic contour, and relatively strong delineation of the sequence boundaries, determined by significant change of global melodic direction and the start of a sequence of high level similarity at the end position.

But the symmetrical implication of the super-phrase and section level – 72 beats divided into two sections of 36 beats, which in turn is divided into two super-phrases of 18 beats - implicates a duple division of the super-phrase, a division in 9 + 9 beats. This symmetrical implication is supported by a global stepwise change of register to beat position nine, determined by the tones at the six preceding beats being significantly lower than the tones of the next six beats. Additionally, a change of global melodic direction takes place at this point.

The total discontinuity value supports this interpretation, giving the highest total value of structural prominence to the (0 9 Y 0 9) sequence. At the same time, the clear change of melodic direction at beat 12 together with the relatively salient sequence starting at this beat position, makes this beat position a salient segment boundary and starting point. The strong structural indications at both of these points also make it possible to conceive the sequence (9 3 x 0 3), i.e. a sequence determined by global change compatibility.

Symmetrical implication, which is the basis of the *sequence* concept, suggest that, when we recognize a similarity between two contiguous streams of events, we expect the second stream to be of equal length as the first stream; the next segment boundary is anticipated at double the length of the first segment.

If we view the current example according to this concept, it can be interpreted as a repeated inhibition of symmetrical expectations:

- The sequence (0 6 x 0 6)/(0 6 C 3 3) implies a segment boundary at beat 12...

- but it is interrupted by a global structural break at beat 9 thus one expects that the first nine are being a segment (0 9 Y 0 9) which raises expectation of a segment boundary at beat 18...
- but it is interrupted by a global structural break at beat 12 which in turn creates the expectation of the sequence (9 3 X 0 3), which would imply a new segment start at beat 15...
- but this is interrupted by the sequence (12 3 N 4 2), which means that beat 12 turned out to be the start of a new sequence pair rather than the middle; this creates expectations of a new end/segment start at pos 18...
- and this is confirmed by the repetition of the whole segment from beat 0 to 18, (0 18 C 3 18). The global repetition also confirms the (0 9 Y 0 9) sequence and clarifies the other segment implication as sub-segments.

This results in the following output solution from the computer implementation of the model:

The image shows a musical score for a piece titled 'sul50010.mf'. The score is written on a single staff in treble clef with a 3/4 time signature. It contains several triplet markings. Above the staff, there are multiple levels of segment brackets, each labeled with a letter and a number (e.g., A1, C1, E1, D1, A2, C2, E1, D2). The labels are not consistent across different levels of the hierarchy, illustrating the complexity of the analysis.

Figure 6-82. Output solution for the first turn of the Polska no.10. Note that due to the graphical design of the implementation, which allows only for three simultaneous segment levels to be displayed, not all identified segments at sub-phrase level are displayed in the output view. The naming of the phrases at sub-phrase level are also not consistent, since the method is not fully implemented.

To conclude, this interpretation of the sub-phrase levels of implies that the symmetrical division of the super-phrase (0 18 C 3 18) into two segments of nine beats (0 9 Y 0 9) includes an asymmetrical division of this phrase level into two sub-segments of 6 plus 3 beats in the first half and 3 plus 6 beats in the second half of the sequence.

However, it is quite possible and even likely that different individuals would conceive the melodic structure below the super-phrase level differently, because of the inherent ambiguity of the structural cues.

## Section II. Overlapping phrases

If we turn to the second turn/major section of this melody, we discover another variant of asymmetrical substructure within an overall symmetrical structure. Here the total length of the repeated section encompasses 42 beats.

Here, (see fig. 83) one cannot find any indication of a symmetrical duple division of the section. Instead, the section begins with a sequence at super-phrase level with recurrent high level similarity after 12 beats, (72 12 NIL 5 6), i.e. exactly the same notes recur after 12 beats. This is a relatively strong structural indication of a segment, giving strong implication of symmetrical structure with a new segment to begin after the repetition, at beat 96.

The figure displays three systems of musical notation for the second turn of Polska no. 10. Each system consists of a staff with notes and rests, followed by beat positions and structural analysis symbols.

**System 1 (Beats 72-83):**

- beat pos: 72 73 74 75 76 77 78 79 80 81 82 83
- prom. value: 12 4 5 2
- Structural analysis: (72 12 NIL 5 6) [A], (72 6 Y 0 6) [X], (72 3 R 3 2) [A], (78 3 NIL 4 3) [A]

**System 2 (Beats 84-98):**

- beat pos: 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98
- prom. value: 5 2 3 5
- Structural analysis: (84 6 X 0 6) [X], (84 3 C 3 3) [A], (90 12 C 3 9) [A], (90 6 X 0 6) [X], (96 3 NIL 4 2) [A]

**System 3 (Beats 99-111):**

- beat pos: 99 100 101 102 103 104 105 106 107 108 109 110 111
- prom. value: 2 10 2 1
- Structural analysis: (102 3 NIL 4 2) [A]

Figure 6-83. Second turn/ section of Polska no. 10. Global structural changes are illustrated above the system. Sequences below the system. Structural prominence values for each sequence segment boundary displayed below systems. (For explanation of shapes see fig. 81).

The first half of this sequence contains implications of two sub-structural hierarchical levels of perfect symmetry. The duple division of six beats (72 6 Y 0 6) indicated by a global change of melodic direction after six beats and the end point of a shorter sequence (72 3 R 3 2) is determined by general similarity regarding melodic contour and the start of the sequence (78 3 NIL 4 3) at the mid position. These shorter sequences indicate yet another sub-structural level, a continuous segmentation in segments of three beats.

The 12 beat phrase level, the 6 beat phrase level and the 3 beat phrase level exhibit perfect hierarchical symmetry and imply this symmetry can be found at the subsequent half of the 12-beat phrase. But only half of the 12-beat sequence is repeated. Instead, a new structure is formed from beat position 90 on, also containing 12 beats, (90 12 C 3 9); It is more complete, but of more general similarity than the previous sequence and with very strong discontinuity indication of the mid position by e.g. stepwise global register shift and concurring sequence boundaries.

Throughout this sequence the symmetrical implication of the previous sequence is inhibited and the new sequence becomes the master segment of the latter part of the section. The significance of this change is reinforced by the use of the first ending of the second 12-beat phrase as the first 12-beat phrase. The central concept of the model is to attribute significance to phrases *when at least half the initial half of the sequence is confirmed by the subsequent sequence half*, which makes both the initial 12-beat phrase and the subsequent 12-beat phrase significant. Below this level, all sequence and discontinuity indications imply hierarchic duple division of perfect symmetry. Thus, an asymmetrical phrase structure at higher phrase level is obtained using a perfectly symmetrical phrase structure at a lower phrase level.



Figure 6-84. Output from computer model. First part of second turn of Polska no. 10.

The total analysis of the segment structure of the melody is displayed in the figure below.





**C motif family**

C1, C2, E1

ORIGINAL

C3, C4

C3/pitch set-similarity

**D motif family**

D1

ORIGINAL

D2

NIL 5 - similarity (50%)

D1

NIL 5 - similarity

D3

NIL 5 - similarity

D4

NIL 5 - similarity (50%)

Figure 6-86. Motif families as identified by the computer model. Arrows and lines above the systems refer to similar global melodic direction. Connected notes marked by circles refer to pitch equality.

We are now approaching the question of the cognitive reality of the analysis obtained by the model. It is most important to realize that there will be listeners who will not hear any coherent sub-structural segmentation in this melody, not to mention the general similarity between distant sub-segments, regardless of his or her familiarity with the style. The degree to which a listener chunks a melodic line into a coherent structure, the granularity of the segmentation, differs radically between listeners according to experiments made both by myself and others (see section 6.3.8). Moreover, there are probably experienced listeners who will conceive the basic phrase levels (level 1-2), but who will not conceive the higher structural levels. The melody will then be conceived as an on-going chain of segments with no higher order, but perhaps with suggestive power. The model can but give a prediction of what can be conceived and try to segment the melody down to the smallest possible units.

### 6.3.3.3 Structural ambiguity by disguised phrase starts

Another problem, which was mentioned among the general problems regarding the structural analysis of the repertoire of Swedish folk music, is the common stylistic feature of disguised phrase starts by absence of tone onset on the implied phrase start.

Consider the following melody example:



Figure 6-87. Example melody (constructed) consisting of a repeated melodic line with the first note of the repetition omitted. Start-oriented interpretation marked by solid lines, while end-oriented interpretation is marked by dashed line.

The problem regarding this example is that while the first sequence half has a clear and evident starting point by an upbeat to the virtual start, the second half has a rest at the corresponding beat position.

One of the strongest determinants of discontinuity of melodic structure in this model, as well as other models that are known to me (see e.g. Cambouropoulos 1998:72, Temperley 2001:68 ff), are relatively long interonset intervals and pauses. If we would use only interonset interval as the cue for grouping in this case we would end up with two segments of different total duration consisting of 14 beats and 10 beats respectively. These segments would not be recognized as implying perfect symmetrical duple division and would hence not constitute the four measure level as a structural unit, implying subsequent segmentation based on the same period.

To find the structurally significant division, we have to assume that sequences can start within a rest, within an offset-onset interval. This is obtained in the current model by assuming that the last beat position before an onset can be a possible starting point of a local discontinuity boundary, considering this position as accentuated because it is the position *from which a change takes place*. (see also Chapter 5, section 5.2.4.7, *Initial successive evaluation*). Global rhythmic breaks (major interonset distances) are assumed to allow for starting points at also more beat positions within the break, if the rhythmic structure is sparse indicating a complex beat level above the designated tactus/central pulse level (see section 6.2.2).

The C and I similarity levels allow for incomplete similarity of varying levels of specificity. Search for brief similarity locates the current segmentation, resulting in the sequence (3 12 I 3 9), an I-sequence (segment boundaries determined by global breaks) with nine beats of equality.

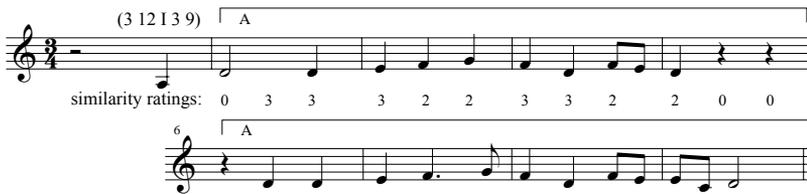


Figure 6-88. The similarity rating of the beat events of the sequence (3 12 I 3 9)

The output of the phrase and section analysis model includes a sub-segmental level of duple symmetrical division, based on compatible change of melodic direction at the mid of the sequence halves.

**filtertest1.mf**

The image shows a musical score for a file named 'filtertest1.mf'. It consists of two staves of music in 3/4 time. The tempo is marked as quarter note = 150. Above the first staff, there are several horizontal lines representing chord annotations: A1, B1, A2, and B2. The melody starts with a rest for one measure, followed by a sequence of notes. The second staff continues the melody with more notes and rests.

Figure 6-89. Output solution from computer model for test melody (Ablbäck 1992)

Below, there is yet another example from a traditional melody which exhibits rest-note variation between repetitions. In certain traditions of northwestern Sweden, such empty starts are a common stylistic feature.

**svl050046.263.mf**

The image shows a musical score for a file named 'svl050046.263.mf'. It consists of four staves of music in 3/4 time. The tempo is marked as quarter note = 110. The melody is highly rhythmic, featuring many eighth and sixteenth notes, as well as triplet markings. Above the staves, there are horizontal lines representing chord annotations: A1, B1, A2, and B2. The score shows a complex structure with multiple repetitions and variations of the melody.

Figure 6-90. Polska no. 46, second section, variation, played by Bengt Bixo, Mörsil. From *Svenska Låtar, Jämtland* (Andersson 1926:28) Output from computer model. The lowest phrase level is missing in the last super-phrase half – the last five measures.

The section level here consists of a perfect repetition of ten measures / 30 beats, each of which begin with a rest. Each section repetition is divided into a super-phrase level of 15 beats (0 15 C 3 13); below that level, there is a segmentation based on discontinuity which divides the super-phrase into nine plus three beats.

In this case, the repetition of the super-phrase level of 15 beats includes not only sequence start on a rest, but also a disguised start by a tie to the beginning of the second sequence half. This is displayed in the computer output by the phrase start beginning on the onset before the tie, while the end of the previous phrase is displayed by a bracket sign right after the tie. The variation between rest and tie start in the repetition indicates interchangeability between no onset due to rest and no onset due to tie.

We will return to disguised starts by tied tones over phrase starts with regard to 20<sup>th</sup> century popular music (see section 6.3.5)

#### 6.3.3.4 Structural ambiguity by melodic variation within sequences

##### Example 1. Evaluation of general structural similarity involving general pitch set similarity

Another general problem regarding the segmentation of melodies within the Swedish instrumental folk music repertoire is the melodic variation within repetition which makes similarity between correspondent segments vague and ambiguous.

Consider the following example, the second section of a *Polska* tune notated in Svenska Låtar (Andersson 1926:11).



Figure 6-91. Second section of *Polska* no. 11 played by Per Danielsson, Mörsil, Jämtland, Sweden. *Svenska Låtar* (Andersson 1926:11)

I conceive this section as consisting of two super-phrases of 12 beats determined by similarity, and I think many experienced listeners within the tradition also would recognize such a segmentation of the section. What are the structural cues for this segmentation of the section?

From the point of view of Western harmony, the segmentation would easily be explained as defined by a repeated harmonic pattern<sup>95</sup>. But the model does not use implicated harmony as a structural indication, since it is designed for analysis of melodies without regard to the existence of a harmony conception within the style. Even in this style, where many melodies are composed with regards to harmonic structure, use of harmony as an input to segmentation would give very weird results. Moreover, the same implied harmonic pattern would fit into many other melodic lines, which would not be regarded as having a similar melody.<sup>96</sup>

Can the melodic similarity between the two sections be determined without assuming culturally determined preconception of functional harmony?

<sup>95</sup> The basic harmonic progression assigned beat-wise would e.g. be I: A D A, D D A, A D A, A E A, II: A A A, D D A, A D A, E A A, which is almost identical although with different cadences.

<sup>96</sup> The harmonic structure of this section is indeed very similar to the one of the first section of the melody

Chapter 6

Locally, regarding beat-to-beat and note-to-note transitions, there are many obvious differences between the two implied sequence halves. But if we look at the global changes we see that each implied segment consists of a rising and a falling melodic line. In addition, there is a stepwise change of register during the first measure of the segments. The general structural feature, which relates to harmonic structure, is pitch set similarity. In this respect – regarding global pitch set similarity – the two sequence halves are also significantly similar.

- Arrows above systems indicates global melodic direction (continuous change of pitch register)  
Change of global melodic direction gives discontinuity value
- step shape indicates global register change (step-wise change of pitch register)
- this shape indicates global rhythmic break (major interonset distance)
- correspondent boxes at both systems with connecting lines indicates global pitch set similarity in relation to preceding pitch set

Figure 6-92. Global changes and global similarity displayed for the second section of Polska no. 11

This is the basis for identification of the sequence, the section is divided into two contiguous segments of 12 beats.

disc.prominence val. 6  
(60 12 C 3 9)

sim. values 2 2 2 3 0 0 1 1 2 1 3 2

Figure 6-93. Sequence (60 12 C 3 9) of section 2 of Polska no. 11 identified by similarity on different levels.

The similarity values between the systems refer to different levels of similarity. The value where 0 represents non-significant melodic contour similarity<sup>97</sup>; 1 represents global melodic contour similarity regarding groups of beats (with pitch set similarity); 2 represents general beat-wise similarity; and 3 represents a perfect (level 4 and 5) melodic contour similarity.

When such brief similarity is recognized it allows for sequences defined by global similarity or compatibility to be found on many correspondent positions, allowing a great number of sequences to appear.

In the fig. 94 below, only the sequences which do not interfere with the initially identified section structure are displayed for this section. Still these can result in a number of different segmentations in the section.

The figure displays two systems of musical notation for 'Polska no. 11'. The first system shows a melodic line in 3/4 time with a key signature of two sharps (F# and C#). It features a triplet of eighth notes followed by a quarter note, and another triplet of eighth notes followed by a quarter note. Below the notation are several horizontal lines representing similarity sequences, each labeled with a sequence of numbers in parentheses, such as (60 24 NIL 5 23), (60 12 C 3 9), (60 11 X 0 11), (60 6 X 0 6), (60 5 X 0 5), (61 23 X 0 23), (61 12 C 3 11), (61 5 X 0 5), (63 3 R 4 3), (64 3 NIL 4 2), (65 7 X 0 7), (65 6 X 0 6), (66 18 X 0 18), (66 8 C 3 5), (66 7 X 0 7), (66 6 X 0 6), and (66 5 X 0 5). The second system shows a continuation of the melody starting with a quarter note, followed by a triplet of eighth notes, another triplet of eighth notes, and a final triplet of eighth notes followed by a quarter note. Below this notation are more similarity sequence labels: (72 12 C 3 9), (74 10 C 3 6), (75 3 R 4 2), and (76 3 NIL 4 2).

Figure 6-94. Sequences identified by the initial sequence analysis in *Polska no. 11*

The evaluation process of these sequences is (as have been explained in section 6.2.4) based on the calculation of a total prominence value for each sequence, which is in turn based on the local discontinuity valuation for the three boundaries of the sequence.

<sup>97</sup> General pitch set similarity is not specific enough in this case to contribute to the similarity value for all beats.

Table 31. Initial total prominence values for the sequences identified for the second section of Polska no. 11

Sequence	Total boundary value	Boundary marker values	Total prom. value
(60 24 NIL 5 23)	28	(6 8 14)	630.00
(60 12 C 3 9)	20	(6 6 8)	135.00
(60 11 X 0 11)	16	(6 4 6)	24.00
(60 6 X 0 6)	16	(6 4 6)	24.00
(60 5 X 0 5)	9	(6 1 2)	1.69
(61 23 X 0 23)	9	(1 8 0)	0.56
(61 12 C 3 11)	2	(1 0 1)	0.09
(61 5 X 0 5)	9	(1 4 4)	1.13
(63 3 R 4 3)	9	(3 4 2)	11.25
(64 3 NIL 4 2)	6	(1 2 3)	12.00
(65 7 X 0 7)	8	(1 6 1)	2.00
(65 6 X 0 6)	7	(1 4 2)	1.75
(66 18 X 0 18)	14	(4 8 2)	7.00
(66 8 C 3 5)	14	(4 4 6)	17.50
(66 7 X 0 7)	6	(4 0 2)	0.19
(66 6 X 0 6)	12	(4 6 2)	6.00
(66 5 X 0 5)	10	(4 4 2)	10.00
(72 12 C 3 9)	20	(6 8 6)	90.00
(74 10 C 3 6)	14	(4 8 2)	17.50
(75 3 R 4 2)	5	(1 2 2)	7.50
(76 3 NIL 4 2)	9	(2 1 6)	11.25

As can be read in this table, the (60 12 C 3 9) sequence receives the highest total prominence value, partly because it has high discontinuity values at both the first and second segment start and partly because the starting point is already determined by the section analysis.

Through the process of successive evaluation of sequences, which involves also aspects of structural hierarchy and symmetry, the remaining sequences are shown in fig. 97 below. The sequence (60 6 X 0 6), which originally has a lower total-prominence-value than some other competing sequences, remains because of implicated symmetry and hierarchical structure.

disc. prominence val. 6  
 (60 24 NIL 5 23) |  
 (60 12 C 3 9) |  
 (60 6 X 0 6)

2 2 2 3 1 1 1 1 2 1 3 2

5 6 2

(72 6 Y 0 6)

Figure 6-95. Remaining sequences of second section of Polska no. 11.

This sequence identification and evaluation forms the basis of the output of the model displayed below:

B1 C1 E2  
 D1 G1 E1 FT

B2 C2 A5 B2

B1 C1 B2 F2 H1 G1 E1

Figure 6-96. Structural analysis of second section of Polska no. 11. Output from computer model.

## Example 2. Melodic variation including variation on global rhythmic level

I will provide yet another example of the melodic variation from the same repertoire, another second section of a *Polska* played by Per Danielsson.



Figure 6-97. Polska no. 13 played by Per Danielsson. *Svenska Låtar* (Andersson 1926:12).

As in the example of the previous section, I recognize a duple division of this section into two super-phrases, the first of which consists of 12 beats and the other of 15. This is within the limit of the duple symmetrical division category as defined in section 6.2.4.4.

In this case the segmentation is supported basically by a similar global melodic direction which includes similar global pitch set between the implied sequence halves as well as certain parts that exhibit perfect beat-to-beat similarity. It is also supported by distinct step-wise change of register at the boundary between the two implied segments.

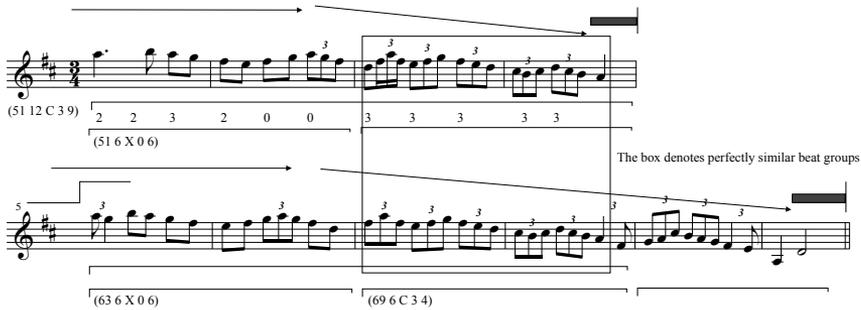


Figure 6-98. Global structural similarity and similarity rating for duple division of second section of Polska no. 13. Similarity ratings are displayed between the systems. Sub-phrases indicated by sequences are also displayed.

However, there are also obvious dissimilarities, the most incriminating being the difference of duration in the beginning of the sequence. The long duration would imply a segment boundary after the two first beats, because of disguised two first beat onsets. There is also significantly different melodic direction at beat-to-beat level in the second measure.

That similarity prevails and, in spite of these differences, can be conceived to emphasize the importance of regarding similarity between groups of events and not only to use duration of *interonsets* as the decisive property of melodic structure.

In this section, there is also, in fact, a “hidden” perfect sequence on note-to-note level, which is assumed to be insignificant on a structural level since it interferes with the basic metrical structure. It is only indirectly recognized by the analysis within the global pitch set and global melodic direction analysis.

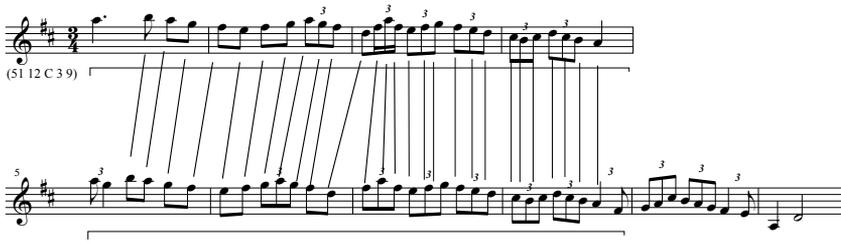


Figure 6-99. Note-to-note perfect similarity between sequence halves at super-phrase level

If this sequence were structurally significant, it would have implied a sequence start at the second note of the section, which would have been inconsistent with the section structure. According to my experiments (See section 6.3.7.3), sequences that are inconsistent with regards to primary metric structure are not likely to be recognized or to be conceived as structurally significant.

Is this note-to-note similarity between the segments the musician's way to use the same fingering and tones put in another position in relation to the metric structure to achieve melodic variation?

### 6.3.3.5 A quantitative test of the model in relation to melodic segmentations by E. Övergaard

Einar Övergaard was a folk music researcher and collector of predominantly instrumental folk music from the Western parts of Sweden. Among his interests as a folk music researcher was the phrase structure of the melodies. (For a survey of his work, see Ramsten 1982:9-33). He analyzed the form of about 3.500 melodies using his own method of describing the phrase structure. His study was performed entirely on *polska* melodies, which was the form which seemed to fascinate him most. The *polska* melodies in his study are all notated in 3/4 time. This study was never published but exists only in manuscript. (Övergaard 1935)

The major repeated sections, the turns, form the basis of his analysis. Since these are notated in the transcriptions, he does not include section analysis in his method. The aim of his method is to describe the melodic structural units below the section level in a fundamentally non-hierarchic manner, appointing one phrase level to be the most pertinent within the sections of a melody. However, he recognizes phrases on all levels, from phrases consisting of entire sections (which can include more than 30 beats) to sub-phrase levels coinciding with the measure level, phrases including only three beats. But most frequently he uses the two/three- and four/six-measure level of phrase length. It is noteworthy that he assumes phrases always to coincide with measure starts.

The manuscript shows that Övergaard was frequently running into problems in his attempt to assign the most salient phrase level, and sometimes he added a second phrase level in his description. What determines his assignment of central phrase level is not clear from the material and seems to be a general weakness in his method (see Ramsten, 1982:31).

This fundamentally one-dimensional design of Övergaard's study makes the results of the current model difficult to compare with the results of Övergaard by a simple measure, such as the F score measure. Since the current model is hierarchical and the melody material exhibits

features that is interpreted as hierarchical, the current model generally assigns considerably more phrase levels, hence segment boundaries, than what is noted by Övergaard.

Because of this basic difference between the materials, I have used two different measures for the evaluation of the current model with regards to Övergaard's results:

1. The level to which the current model succeeds in finding the same phrases, including position of segment boundaries and phrase length.
2. The level to which the current model succeeds in finding the segment boundaries assigned by Övergaard

These two measures indicate how well the current model can account for the segmentation performed by Övergaard.

The second of the two measures is added because of the lack of consistency in Övergaard's choice of pertinent phrase levels. This implies that the current model and Övergaard's model can end up with concurrent phrase boundaries originating from phrases at different hierarchical levels.

Sixty melodies are included in the comparison, originating from three different parts of the Svenska Låtar collection, part 1 Dalarna (melodies no. 1-22), part 5 Jämtland (melodies no. 1-35), part 13 Västergötland (melodies no. 3-34). In some cases, melodies have been excluded from the comparison, since Övergaard had not performed an analysis of the whole melody or because the manuscript was difficult to interpret.

The choice of repertoires was directed by the wish to test the current model on melodies of different rhythmic and melodic structure within the *polska* category, to obtain a result which would have relevance for the entire material in Övergaard's study. The result of the test is summarized in the table below.

Table 32. Results from comparison between the results of the current model applied to melodies in a study of phrase structure in polska melodies by Einar Övergaard (Övergaard 1935)

Category	No. of tune in Svenska Låtar collection	No. of phrase bound. Överg.	Ratio of correctly identified phrases	Median of identified phrases	Ratio of correctly identified phrase boundaries	Median of identified phrase boundaries
Svenska Låtar Västergötland no. 13 (22 melodies)	no. 2-3,6,8,10-11,13-14,16,18-19,21-24,26-27,29-32,34	323	81.0 %	0.83	94.5 %	1.00
Svenska Låtar Dalarna no. 1 (19 melodies)	no.1-19	312	78.5 %	0.88	88.7. %	1.00
Svenska Låtar	no.1-17,25,35	384	92.3 %	1.00	92.3 %	1.00

Jämtland no. 5 (19 melodies)	17,25,35					
TOTAL		1019	82.1 %	0.92	93.7 %	1.00

As can be seen in this table, the model was reasonably successful in locating the phrases identified by Övergaard. The model managed to find an average of 82.1 % of the 1019 phrase segments. Moreover, 93.7 % of the 1019 phrase boundaries were identified.

The median is added since it reveals something important about the characteristics of the result. In most cases, the median is considerably higher than the mean. This reflects the typical behaviour of the model, which, when it fails to find e.g. the correct section structure by appointing the upbeat the section start consistently misses all the appointed phrases and phrase boundaries, which cause the result to be considerably below chance level. This is true for two of the melodies of the Västergötland repertoire and two of the melodies in the Dalarna repertoire.

Another general failure, which consistently affects the entire resulting phrase structure below section level, considers misinterpretation of symmetrical hierarchy. This is the case in two of the tunes of the Jämtland repertoire, two of the Dalarna repertoire. This causes the result to drop to chance level or below.

This can be viewed in the table below, which displays the number of correctly identified phrases, the number of not identified phrases for each tune and the resulting ratio.

Chapter 6

Table 33. The detailed result of the comparative test regarding phrases (boundary connected to length) from output result of the current model in relation to analyses by Övergaard. Results on or below chance level are shaded.

Tune number	Correctly identified phrases	Not identified phrases	Ratio of correctly identified phrases
svl130002	8	0	1,00
svl130003	8	4	0,67
svl130006	12	4	0,75
svl130008	7	3	0,70
svl130010	16	4	0,80
svl130011	6	4	0,60
svl130013	7	13	0,35
svl130014	18	6	0,75
svl130016	8	17	0,32
svl130018	8	0	1,00
svl130019	4	0	1,00
svl130021	16	0	1,00
svl130022	14	0	1,00
svl130023	24	1	0,96
svl130024	8	4	0,67
svl130026	12	4	0,75
svl130027	19	3	0,86
svl130029	12	0	1,00
svl130030	16	9	0,64
svl130031	6	0	1,00
svl130032	8	0	1,00
svl130034	10	0	1,00
svl010001	14	4	0,78
svl010002	14	2	0,88
svl010003	14	0	1,00
svl010004	10	0	1,00
svl010005	0	16	0,00
svl010006	12	0	1,00
svl010007	12	0	1,00
svl010008	11	11	0,50
svl010009	20	2	0,91

svl010010	14	4	0,78
svl010011	29	23	0,56
svl010012	3	11	0,21
svl010013	16	0	1,00
svl010014	15	1	0,94
svl010015	9	3	0,75
svl010016	10	2	0,83
svl010017	12	0	1,00
svl010018	4	0	1,00
svl010019	11	3	0,79
svl050001	12	0	1,00
svl050002	16	2	0,89
svl050003	12	0	1,00
svl050004	21	1	0,95
svl050005	36	0	1,00
svl050006	12	2	0,86
svl050007	10	2	0,83
svl050008	20	4	0,83
svl050009	4	16	0,20
svl050010	21	1	0,95
svl050011	8	0	1,00
svl050012	40	0	1,00
svl050013	27	3	0,90
svl050014	3	25	0,11
svl050015	29	1	0,97
svl050016	20	0	1,00
svl050017	8	0	1,00
svl050025	12	0	1,00
svl050035	16	0	1,00

Another way to interpret the result is to consider it to be very good for 52 out of 60 melodies, 86.67 % of the material, for these melodies giving a result of which accounts for 88.9 % of the phrases identified by Övergaard. The remaining 8 melodies give an erroneous result, in most of the cases due to deficiencies in the implementation of the current method. For two of the melodies, this is also due to lack of sufficient structural information or to high level of inherent structural ambiguity.

## 6.3.4 Phrase structure in Norwegian gangar melodies

### 6.3.4.1 General problems

Some of the *gangar* and *balling* melodies of the Norwegian instrumental folk music tradition offer extreme challenges for any method of analysis of melodic structure. In contrast to the Swedish *polska* melodies, this melody style is characterized by asymmetrical, ambiguous structure at section level. Symmetry is more common on lower phrase levels, though often not consistently. There is no *timed* general structure in this music; correspondent sections and phrases can differ in length between repetitions within the melody, which makes implications of hierarchy and symmetry ambiguous and vague. Moreover, motifs can change syntactic roles within the structure; the ending motif of one phrase come the starting motif of another. This makes the melodic structure nested, and require it almost to be deciphered by a listener, as well as an analyst. I have yet not met anyone, has not been brought up with the music, who has found this music easy to grasp or described it as having a simple or clear structure. On the contrary, these tunes can be conceived as musical riddles, possibly without any answers.

A Danish musicologist, Morten Levy, has made a comprehensive study of a small group of melodies within this repertoire in his thesis, “The World of the Gorrlaus slats” (Levy 1989). In this extensive work, he develops a general theory of musical structure adapted to the particular repertoire of his study, influenced by concepts from semiology and linguistics. A great deal of his work is devoted to melodic segmentation, and he argues for the necessity of involving relationships between segments in the analysis of segmental structure. He also argues for the need to allow ambiguity, both on the level of segment length and regarding the motif identity of segments.<sup>98</sup>

His theory of melodic structure is based on the identification of structural units called *circuits* (cf. the *sequence* concept of the current model), delineated by melodic parallelism/repetition and characterized by certain melodic, tonal and rhythmic-metric features (called *fields*). These are combined into higher order hierarchical levels, sections, which form the major structural units of the melody.

As mentioned above, the problems regarding the analysis of the melodic structure of this type of melody are extensive. To cite Levy:

*‘I shall mention:*

- *a first segmentation of a slått playing on the basis of simple, preselected (i.e. independent of the material) criteria does not immediately offer itself. I am thinking particularly of a mode of division suggested by Nicholas Ruwet (Ruwet 1966), according to which the pauses in the music are taken as the first basis of division. It falls to the ground of itself here, for the simple reason that there are practically no pauses.*
- *that the slåtts in the selected material differ widely in extent, especially in duration, but also in range.*

---

<sup>98</sup> The work by Levy is unique within the realm of analytical works of Scandinavian folk music, but to my knowledge no further studies have been performed using his interesting and challenging concept. He presents a comprehensive theory of musical structure, which due to its general approach can be compared with far more influential works, as e.g. the works by Lerdahl and Jackendoff, Nattiez, Narmour and others.

- *that a slätt consists of a free number of voltas* [times of repetitions of the entire chain of sections which constitutes the melody].
- *that a volta consists of a free number of sections.*
- *that a section consists of a free number of circuits* [phrases, defined primarily by sequencing].
- *that a circuit consists of a free number of fields* [characteristic elements of the circuit].
- *that the metre is free and non-periodical.* “(Levy 1989:11)

In other words, there is asymmetry and structural inconsistency on all hierarchical levels, which would indicate that symmetrical and hierarchical implication play no role in the conception of melodic structure. Yet simultaneously Levy points to the sequences, which he calls *circuits* or *repeat structures*, as being the main determinant of primary melodic units on phrase level. Thus, symmetry in metrical terms exists, however variable, on a lower structural level, and which makes symmetrical implication possible. Likewise, hierarchy exists, despite the asymmetry and variability regarding the relationship between higher and lower levels.

The inherent ambiguity regarding the structural cues of this melodic material makes it interesting to study the behaviour of the current formalized model. Levy's analyses are not based on a formalized method of segmentation or motif identification. They do not contain any description of a consistent method of defining which repetition of all possible should be recognized – the phase of the repetition - or the degree of melodic similarity for a repeat-structure/circuit to be recognized, even if these questions are discussed. Further, the identification of the pertinent substructures of a circuit is not formalized in a way, which can be tested by others. It is rather a qualitative theory, which contributes by formulating important aspects about the nature of the musical structure.

We will view the behaviour of the current model in two examples, the first of which exhibits a lower level of structural complexity than the second. The model has been tested on several other melodies within this tradition, but, for the benefit of comparison with an existing structural analysis only one in-depth analysis, will be given here.

#### **6.3.4.2 Example 1. Sordölen. A chained sequence structure**

The first example is a *gangar* melody, called *Sordölen*, adapted from a transcription of performance by Eivind O. Hamre (1853-1934), Hamre, Setesdal, Norway. (Lande 1983:116)

### B. Sordölen.

Etter Eivind O. Hamre.

The musical score for 'Sordölen' is presented in six staves. The first five staves are polyphonic, with larger note heads indicating the melody. The sixth staff is a single-line melody labeled 'Conda. (2. gang.)'. The score includes various musical notations such as slurs, accents, and dynamic markings.

Figure 6-100. Transcription of Sordölen, gangar after Eivind O. Hamre, *Setesdal* (Lande 1983:116)

A typical problem regarding this repertoire, which is generally performed on the *bardingfela* or the *violin*, is the polyphonic setting, which my model cannot handle in the present state. But it is a common practice in the notations of these tunes to designate one of the simultaneous notes to be the melody note, while the other notes are regarded as accompanying notes. This is marked in the notation by assigning larger note heads to the melody notes than to the accompanying notes.<sup>99</sup> This makes it possible for me to reduce it to a one-line melody. That this interpretation of hierarchy of simultaneous notes in this musical style is not only a matter of notation practice is indicated by the fact that these tunes are also played on instruments which cannot produce chords<sup>100</sup> and with different level of polyphony in the fiddle playing, but are still being recognized within the culture as identifiable versions of the tune. The concept of melody thus seems to be applicable to at least a large portion of this repertoire.

This melody has a relatively low level of complexity in comparison with the more complex tunes within this repertoire. This version of the tune exhibits, however, both some of

<sup>99</sup> To the extent that this represents a listener's conception, it will be possible to include such a separation in the model, which is outlined in the final discussion (see section 6.3.9).

<sup>100</sup> e.g. jaws harp and voice

the typical structural features of the style, and some important problems common to many melodies.

Basically, the structure has the same low contrast regarding duration of events, as was characteristic of the Swedish *polska* music, as was pointed out by Levy in his characterization of the musical style. The pitch structure does not provide enough sufficient indications of major and consistent changes, for a more thorough and consistent segmentation.

But if we turn to melodic parallelism, segmentation based on similarity, the structure gives considerably more clues.

Below (Fig. 101) are the initially identified sequences in this version of *Sordölen*.

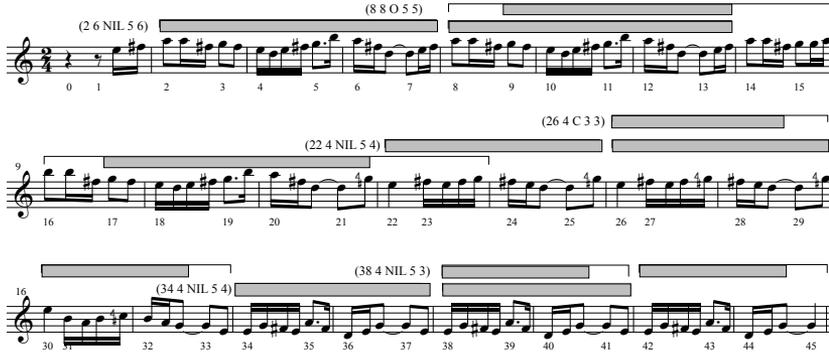


Figure 6-101. Sequence structure identified by the current model, within the sequence analysis. Sequences are here displayed over the systems, designated by brackets between the boundaries of each sequence half. Similar parts of the sequences are shaded.

From this figure (fig. 101) we can see that the tune begins with a sequence of six beats (2 6 NIL 5 6)<sup>101</sup> with perfect repetition, implying a segment start at position 14. The second half of this sequence is, however, included in an overlapping sequence of eight beats, with a different first beat (8 8 0 5 5), implying a segment start at beat 24. This is, in turn, overlapped at the end by a sequence starting at beat 22, defined by four beats of perfect similarity (22 4 NIL 5 4) implying a segment start at beat 30. But this sequence is chained by a sequence starting at the mid-position with more general similarity but of equal length (26 4 C 3 3), implying a segment boundary at position 34. This implication is confirmed by a new sequence starting at this position (34 4 NIL 5 4), exhibiting perfect similarity. But this sequence is in turn chained by a similar sequence starting at the mid-position.

The key words for this structure might be chain and overlap. Sequence implications are consistently inhibited by new sequences starting at the mid position of earlier sequences, determined by varying degrees of similarity.

This solution does not, however, fully comply with my intuitive conception of the sequence structure in one particular part of the melody. Consider the second sequence, the overlapping (8 8 0 5 5). This interferes at first with the implied sequence ending at beat

<sup>101</sup> NB Start-oriented grouping, an end-oriented grouping solution would include the pickup notes.

position 14, which is determined by a sequence of perfect similarity. Further, it interferes also with the start of the sequence (22 4 NIL 5 4), which also is very salient due to perfect repetition and salient local discontinuity indications (no onset on beat preceding its start). I do not spontaneously conceive the (8 8 0 5 5) sequence but rather I conceive a sequence starting at the position implied by the previous sequence (beat pos. 14). This lasts to the beginning of the next sequence at beat 22 as an expanded variation of the first sequence, determined by one beat similarity in the beginning and a similar ending.

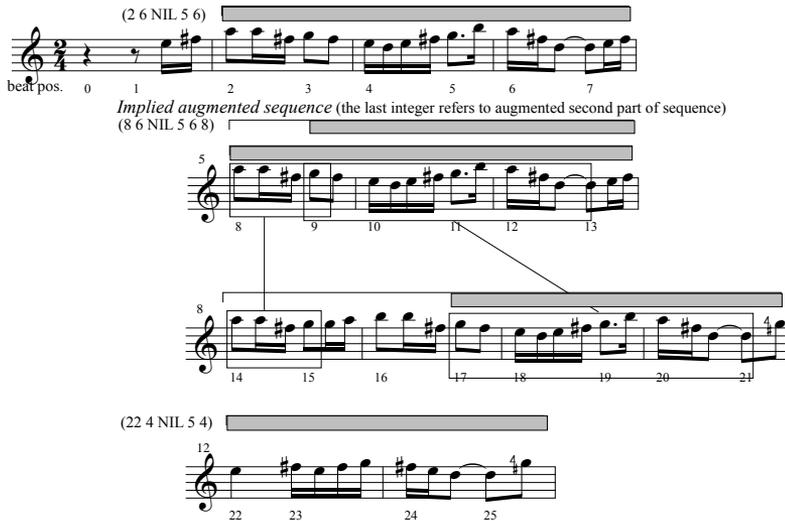


Figure 6-102. Sequence structure of Sordölen, including sequence defined by augmentation of sequence repeat. Connected boxes designate corresponding melodic content.

In the above figure, the similarity between the expanded segment and the original segment, which it is a variation of is indicated by boxes around the equal events with connecting lines. The perfect similarity, regarding the first beat, strengthens the start implication originating from the first sequence and the perfect similarity between the end of the varied repeat. The rest of the original segment strengthens the categorization of the two segments as a repetition with variation.

If my conception is relevant, there is a need for the handling of phrases defined basically by end- or reversed similarity (see section 6.2.3.3, *R-sequences*), but which are *expanded* by variation in middle of the sequence. Since the sequence analysis more easily handle the opposite – early phrase-overlaps through end variation – this kind of sequences are constructed within the model from the initially identified sequences. They are internally designated by a sixth integer in the sequence-ID, which designates the length of the second part of the sequence.

Regardless of the interpretation, this kind of variation, expansion and contraction of corresponding segments is indeed very common in this repertoire, as also pointed out by Levy (1989:90), which raises the level of ambiguity in the structural analysis.

The current model gives an output, which complies with the concept of the augmentation with reversed similarity, resulting in a chain of similar sequences (see fig. 103 below). The categorization of similarity, groups them however into three motif families, the A, B and C motives. These categories are, however, not discrete, and the categorization algorithm uses the level of difference in similarity to group them. From this analysis, one might construct a higher structure that would correspond to Levy's *section* level (1989:88), which would indicate a structure of three sections forming one *volta* of the tune, to use Levy's terminology.

A. Start-oriented version

The image displays a musical score for a start-oriented version, consisting of five staves of music. The first staff begins with a treble clef, a 2/4 time signature, and a key signature of one sharp (F#). A tempo marking of ♩=96 is present. The score is divided into six distinct motifs, each indicated by a horizontal line with a small square icon and a label above it: A1 (spanning the first two measures), A2 (spanning the next two measures), B1 (spanning the next two measures), B2 (spanning the next two measures), C1 (spanning the next two measures), and C2 (spanning the final two measures). The motifs A1 and A2 are characterized by a rhythmic pattern of eighth notes and quarter notes with a sharp sign. Motifs B1 and B2 show a similar pattern but with a different rhythmic arrangement. Motifs C1 and C2 feature a more complex rhythmic structure with eighth and sixteenth notes.

The end-oriented version is shown below; It differs from the start-oriented interpretation by the inclusion of up-beats. Since this is a regular feature it is likely to be conceived as end-oriented (Cf. discussion in section 6.2.6).

Except for the lack of higher level analysis, which would require clustering of similar contiguous segments into larger groups, the analysis does not perform any sub-phrase segmentation in this case though it would be possible to conceive this. This is due to limitations of the current implementation of the method.

B. End-oriented version

Figure 6-103. Output from analysis by computer model: The phrase structure of Sordölen. A start-oriented and B end-oriented version.

### 6.3.4.3 Example 2. Norafjells. Asymmetrical hierarchy.

#### Introduction

The cardinal example of Morten Levy's work is the tune *Norafjells*, a highly complex tune which exists in many structurally very different versions among a relatively small number of performers in a small area in the southern parts of Norway, Setesdal.

I have used this tune in a version that I have transcribed from the playing of Andres Rysstad (1893-1984), a renowned fiddler within the Setesdal tradition.<sup>102</sup>

When using a transcription of a tune as material for analysis, it is worth considering that transcription of this music into common notation is not a trivial matter. Accordingly I have not yet found two transcriptions of tunes with this level of complexity that are entirely similar<sup>103</sup>. A fundamental difficulty in the absence of a video recording of the performance, is to determine the bow stroke onsets, which are highly significant for the conception of the

<sup>102</sup> A very similar version is also transcribed by Morten Levy in his thesis (Levy 1989, III :111)

<sup>103</sup> This is also pointed out by Levy (see e.g. 1989 III:3)

accent structure of the music. Another obvious difficulty is to determine the boundaries between ornaments and core notes of the melody. Sometimes there is no such clear boundary, but rather a diffuse transition between the two categories.<sup>104</sup> Moreover, the distinction between melody tone and accompanying tone is not always possible to make in a consistent way, since a tune can include sections with pronounced chord structure. Yet another typical difficulty, however not of vital importance in the current context, is the categorization of intonation differences.

## Norafjells

Transcription from recording with Andres Rysstad, Hylestad, Setesdal

transcr. Sven Ahlbäck 1992

The image displays a musical score for the piece 'Norafjells'. It consists of five staves of music, each with a treble clef. The time signatures vary across the staves: 6/8, 6/8, 3/8, 3/8, and 6/8. The melody is written on a single line, with accompanying notes indicated by smaller note heads. The score includes various musical notations such as slurs, accents, and dynamic markings.

Figure 6-104. Original transcription of Norafjells (Initial part), played by Andres K. Rysstad (1893-1984), Hylestad, Setesdal, Norway rec. 1973. Assigned melody notes indicated by larger note heads than accompanying notes.

This original transcription was then simplified, omitting ornamentation, articulation and accompanying notes, to be used as the material for the analysis. This process resulted in the

<sup>104</sup> Levy's and my transcription differs consistently in this respect (cf. Levy 1989, III:111)





Figure 6-106. Norafjells. Section starts initially identified by the computer model (end of melody analysis clipped)

The section analysis identifies three major repeated sections within the tune. These are internally designated by the section sequences (0 98 NIL 3 14), (24 92 NIL 5 11) and (36 97 NIL 5 21). The order of the section-starts between the corresponding section halves indicates a global repetition – a second round or *volta* (Levy 1989 I:13) of the tune - at beat 98 (at the mid of the first system of the right column), indicated by the root letter A. Each global section then consists of a chain of sections A, B and C.

This indication is used by the sequence analysis and local phrase boundary analysis as an *imperative* start, a start position which is determined by the global structure and must not be overridden by local sequences or local structural boundaries. In this case, this procedure is essential since the start of the second *volta* of the melody plays the role of the ending of the previous phrase.

This major segmentation of the melody conforms to the first major division in *voltas* made by Levy in his analysis of this tune. (Levy 1989 II:698).<sup>105</sup>

## The phrase analysis

As have pointed out above (see section 6.3.4.2, *Sordölen*), the analysis of major repeated segments is not sufficient for determining the entirety of the macro structure. It must be based on the analysis of the more local structural segments, on phrase level. The result of the

<sup>105</sup> Levy is, however, making a re-interpretation of the variation of the B section in the second turn, by which he induces an A phrase/circuit between the B and C sections of the second volta, which seems inconsistent with his interpretation of the first B section.

phrase analysis is displayed in the fig. 109 below, including beat categorization. Note that the sections obtained by the initial section analysis are not included in the recognized phrase structure.

Figure 6-107. Norafjells after Andres Rysstad. Output from computer model analysis of phrase structure. Start-oriented interpretation. (the entirety of the analyzed part of Norafjells)

Compared to the analyses of the *polska* melodies by the current model presented in the previous sections (section 6.3.3), this output is strikingly different with its almost total lack of hierarchical levels of phrase segments.

The structural principle that emerges from viewing this interpretation may be designated *repetition with variation*. Almost every phrase in the whole chain of phrases at the central phrase level has melodic elements in common. All new phrases seem to emerge out of variation of previous phrases, still having different characteristics which make them separable from each other. The categorization of phrases into groups based on similarity offers many opportunities, since the overall similarity between phrases is salient.

The higher level structure, which emerges from the phrase analysis can be summarized as demonstrated in the following scheme using phrases at intermediate level:

Table 34. Summarized higher structure based on clustering of similar contiguous phrases at intermediate level. Phrase designations refer to the computer model output (fig. 107).<sup>106</sup>

First round (volta)	Second round (volta)
<u>A-section</u> A1, A2, A3, A2, A4	<u>A-section</u> A3, A2, A3, A5 (=A4)
<u>B-section</u> B1, B1	<u>B-section</u> B1, B1, A5 (=A4)
<u>C-section</u> B2 (B3+A4), B4 (The B- and C-sections could also be grouped into a single section based on salient general similarity)	<u>C-section</u> B3 (B3+A4), B8 (=B3)
<u>D-section</u> C1, C2	<u>D-section</u> C1, C1, A6 (≈ C phrase)
<u>E-section</u> D1, D2, D3	<u>E-section</u> D1, D2, D3
<u>D-section</u> C1, C3	<u>D-section</u> C1, C3
<u>B-section</u> B5, B6, B7	<u>B-section</u> B5, B9

As is noted in the table, an alternative section designation of higher level order based on phrase similarity would be to regard the B and C sections as one section due to general similarity. Since, they are separated by greater dissimilarity in the second round, however, they were initially by the section analysis categorized as section starts. This results in a round consisting of a chain of sections or macro phrases of A B C D E D B, which are identical for the two rounds.

A comparison with Levy's analysis of the form of the tune demonstrates the close resemblance between the above analysis and Levy's. There are nevertheless some inconsistencies in his designation of form elements (circuits) between identical passages in the first and second round which cause some minor differences.<sup>107</sup>

Table 35. Corresponding designated sections in Levy (1989 III:111-112) and the result of the current model. Corresponding sections displayed on the same rows. Levy designation of sections is kept in the table, stroked letter referring to the category floating sections (Levy 1989 I:255 etc.)

Current model	Levy (1989 III:111-112)
A	<u>A</u>

<sup>106</sup> Although there exists an identical phrase in the structure, some phrases have been assigned wrong phrase identity, primarily with regard to phrase integer. The correct identification is noted within in parenthesis after the original numbering.

<sup>107</sup> Levy's designation of sections based on circuits in the analysis of the current tune (Levy 1989 III:111-112, does strangely enough not comply with his own general listing of circuits appearing in Norafjells in Andres Rysstad's version (Levy 1989, I:99)

B	B (different section end)
C	C (different section start)
D	∅
E	∅
D	∅
B	C (different section start)

**“The story told in Norafjells” A closer look into the phrase chain**

A close look at the output of the analysis performed by the current model uncovers some interesting properties of the structure of the tune. In the figure below, the phrases identified by the current model have been cut out from the original result and ordered vertically. The phrases are displayed including the upbeats, reflecting the end-oriented interpretation. Note, however, that both upbeats and endings are included when implicated phrase-endings are interrupted or overlapped by forthcoming phrases. I have also added some sub-phrases (indicated by omitted phrase-ID integer) which are not displayed in the original output, to illustrate low-level similarities.

The figure displays three sections of a musical score, labeled A, B, and C, with various phrases and sub-phrases marked by brackets and IDs. Section A contains phrases A1, A2, A3, and A4, with sub-phrases A and A. Section B contains phrases B1 and A5, with sub-phrases B and A. Section C contains phrases B3, B, and A4, with sub-phrases B4 and B. The score is written in 6/8 time and includes upbeats and endings.

Chapter 6

The figure displays a musical score for 'Norafjells' with computer model output analysis. The score is organized into four sections, each indicated by a vertical bracket on the left. Each section contains two staves of music. Above the staves, boxes represent detected chords, with lines indicating their duration. The chords are labeled as follows:

- Section 1 (D):** C1, B, C2
- Section 2 (E):** D1, D2, B, D3, B, B
- Section 3 (D):** C1, B, C3
- Section 4 (B):** B5, B, A4, B6, B, A4, B7, B, A3

Figure 6-108. Norafjells. Output analysis by computer model for the first round: ordered vertically, including pickups. Sections are displayed by left vertical brackets.

The rather intricate phrase structure of the first round of *Norajfjells* will be explored below. The phrase and section designations refer to the above figure (fig. 108)

### **A section**

If we follow the course of events according to this disposition, the A section (A1-A4) exhibits minor variation of the phrases, which are generally of the same length and with the same beat length structure. However, the last phrase, A4, is actually interrupted, since the symmetrical ending which would imply a rhythmic structure in accordance with the previous A phrases, turns out to be the beginning of a new phrase and a new section.

This new section start can be interpreted in different ways, it begins either on the anacrusis (pickup beat) or as it is interpreted here. The reason for the model to treat the first three notes of the longest repeated sequence as an upbeat to the virtual start on the longer note, is that the following figures are closer in pitch; it involves a global continuous change of register beginning at the long note, together with a step-wise change of register to this note. (cf. Levy 1989 1:88)

However, the symmetrical implication of the previous phrases, together with the competing structural implication of overlapping repetition helps *nesting* the end of the previous phrase together with the start of the following. This is also emphasized by salient structural discontinuity on a global level.

### **B section**

The B section consists of a perfect sequence of two exactly repeated segments, the end of which contains the former A phrase in its last variation (A4). Thus the start indication of the A phrase is inhibited, since it turns into a phrase ending of a larger phrase. This shift of position of sub-phrases is quite common, not only in this particular tune, but in this repertoire at large. It also contributes to the nesting of phrases, new phrases emerging from the contents of the former.

### **C section**

The C section is comprised of a phrase which is a variation of the B section phrase. It is expanded in the way discussed in the previous section regarding *Sordölen* (section 6.3.4.2): - by the insertion of a new element of two beats following on the similar beginning, - but before the ending identical of the previous phrase concludes the phrase. Since, this ending; identical to the A4 phrase; links this section to the former sections and reinforces the structural principle of continuous variation.

### **D section**

The C section implicates a phrase ending after 12 beats of the second sequence half, but is interrupted by the D section. The D section is constituted by the C phrase which begins similarly to the A phrases and then end in similar way to the second measure of the B3 phrase of the previous section. Thus the start of the C phrase is identical to the first beat of the ending of the B3 phrase (the A4 phrase) and once again avoids the fulfillment of the implicated structure.

### **E section**

The second half of the D section (phrase C2) is also interrupted by the start of the E section, which once again is introduced by a new element at the beginning of the phrase. Here the ending of the first phrase within the section (D1) has a rhythmic structure similar to the A phrases, but, in the second phrase, the ending is more similar to the sub-phrase, which was introduced in the B3 phrase and contained in the ending of the C1 phrase (marked as B sub-phrase in the figure above). This connection becomes even stronger when the D3 phrase is prolonged by an exact replica of the sub-phrase B. The different endings make all of the D phrases different in length; D1 has a duration of 15/8, D2 12/8 and D3 18/8; Thus symmetrical implications are constantly inhibited. What remain constant is the smaller units, which emphasize the lower levels of the melodic structure.

### **C section - recurring**

The E section is followed by a return to the phrases of the previous section, the D section, with the introduction of phrase C1.

### **B section – recurring in varied form**

The repetition of the C phrase is once again interrupted by a 3/8 bridge over to a phrase, which is identical to the latter part of the B3 phrase of the C section. It bears close resemblance and is identical in length to the B1 phrase of the B section, thus forming yet another variation of the B section.

With the B section, the A phrase (A4 version) returns, as the ending of the B phrases. The B phrase is repeated three times, the third time with the typical augmentation by variation of the second beat.

But here everything changes once again. The implicated ending of the B7 phrase being the A phrase, now becomes the start of the next round, and starts a repetition of the same chain of sections with further elaboration of the phrase level.

### **Conclusive remarks**

According to this analysis, we can conclude that the sequence structure is the heart of the structural organization of the *Norafjells*; The variability regarding the number of times a sequence is repeated, the length of the sequence and the continuous forming of new sequences based on elements inherent in previous sequences creates an emerging structure based on low level phrase elements, which are combined to form segments on a primary phrase level.

As Levy demonstrates, the chain of sections is also subject to variation within the concept of the tune *Norafjells*. (Levy 1989:11 ff) He further demonstrates the emerging variable structure in several illustrative ways, including the presentation of the structure as a game or a circuit card (Levy 1989:101, 118 ff). Levy includes several aspects other than phrase structure in his study; for example tonal development within the tune seems to be of crucial importance for the identity of the melody.

Levy designates the repeat structures (sequences) as *circuits*, thus implying that these are circular structures with no beginning or end.

"I shall now introduce the concept of circuit as a name for, and a characterization of the repeat structures we have dealt with. The reasoning behind this concept is;

- 1) *As it is a question of immediate repetition, it is also in a sense a question of a return to the course first traversed (cf. the discussion of cyclic time, p.00). In that way the music moves in circuits similarly to the blood in the body, the current in electric circuits, or trains in railway circuits.*
- 2) *As long as the repeat structures, as they appear in a slätt, cannot be shown to indicate their own beginnings and endings, it would be a misplaced act of violence on the part of the describer to introduce such criteria into the music. Instead, the concept of circuits means a closed course, which does not in itself have a beginning or an end, and which is reached in practice by entering somewhere into the course, and leaving the course somewhere.*
- 3) *The concept of circuits, finally, makes good sense in connection with the observations made above concerning the nature of development in the music, where part of each repeat structure was carried into the next one. Thus it corresponds very well to the way electric part circuits can develop from each other by means of relays." (Levy 1989 I:96)*

This interpretation of the structural units of this music is intriguing and reasonable from the point of view of the extreme variability within the aforementioned repeated structures. The cyclic, streaming quality of this music is apparent; It is conceived as static, a constant ongoing without any starts or endings. But, as Levy himself demonstrates, there are constant successive properties regarding the chain of repeat structures of the tune, forming a significant course of events constituting the rounds of the tune. If such an order is recognized on one level, how can it be insignificant at another level? And if the phase of the period were not at all significant in the conception of the musical flow (which here cannot be reduced into an oscillation), how can the periodicity then be conceived – how can we determine when we have reached the return of the period? People are known to be chunking chains of events in time into conceivable patterns, implying impulse points even when they are not even physically present or are challenged by the musical structure (see e.g. discussion in London 2002:531ff). And why are balanced repeat patterns, which basically fall within the limits of duple symmetrical division described in section so common within the repertoire?

The *circuit* concept is attractive in many respects: The open and cyclic quality of the *repeat structures*, also emphasized by the analysis of the current model; the play with the creation of and inhibition of symmetrical expectations described in the above interpretation is a conceptual reality at least to me.

Thus I argue that since grouping of events in time, segmentation, is shown to be a spontaneous process in human cognition and is applicable to the scale and complexity of the repeat structures of this music, people are likely to induce *virtual starting points* or *impulse points* when entering a cycle, resulting in a segmentation of the large scale structure. This segmentation, however ambiguous and fluctuating, will, as a consequence of the repetition on which it is based, invoke expectations of symmetrical continuity. These expectations are challenged by the non-hierarchical, asymmetrical variation of repeated structures based on the free combination of sub-segmental elements, which is a general principle of variation in this style.

There are obviously many structural aspects regarding the “story told in *Norafjells*”, which are not touched upon in this limited description which focuses upon the surface structure of the melody. But as Levy acknowledges himself, these must be based on the identification of the surface structure elements. (Levy 1989 I:87 ff, II:554)

**A comparison with Levy’s analysis of structural units – circuits vs. sequences in *Norafjells***

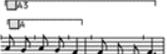
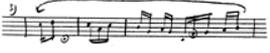
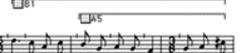
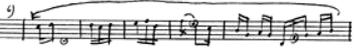
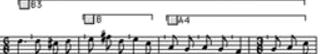
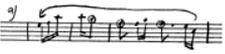
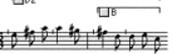
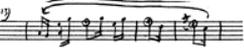
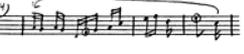
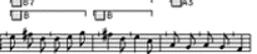
CIRCUITS (Levy 1989 I:99)	SEQUENCES (current model)
	
	
	
	
	
	
	
	
	
	
	
	
	

Figure 6-109. Circuits according to Levy (1989 I:99), with the corresponding sequence according to the output of the current model. (Note that 16<sup>th</sup> notes in Levy’s notation correspond to 8<sup>th</sup> notes in

*my notation. The d1-f1 figure at the beginning of the first circuit, end of second etc. is represented by a quarter note f1 in my notation).*

In order to evaluate the result of the model of phrase and section, it is interesting to see how well the results of this analysis manages to account for the structural units defined by Levy.

He has provided a circuit scheme for *Noraffjells*, based on another version of the tune played by Andres Rysstad. Fig. 111 below displays the circuits according to Levy and the corresponding sequences found by the current model. As can be seen in the above table, almost all of the circuits presented by Levy, are identified by the current model. Moreover, the designated starting points, which are non-significant according to Levy, mostly concur as well.

Thus, I interpret this result as demonstrating that the current model is able to generally account for the type of basic segmentation of a non-hierarchical, asymmetric structure, as the one performed by Levy.

What remains unimplemented in the current model, is the formation of higher order structure from the basic phrase.

### 6.3.5 Phrase structure challenging the metrical structure – examples from Baroque music and 20<sup>th</sup> century Western popular music

#### 6.3.5.1 The fortspinnung principle - Asymmetry in between global and local symmetry

gmollpresto.tot.mf

The musical score consists of 12 staves of music. Above each staff, there are chord diagrams represented by letters and subscripts, indicating the harmonic structure. The chords are: A1, A2, B1, B2, C1, D1, E1, E2, G1, G2, H1, H2, J1, J2, J3, K1, K2, L1, and L2. The music is in 3/4 time with a tempo marking of 'J=120'. The score illustrates the fortspinnung principle through its phrasing, which often extends beyond the expected metrical boundaries.

The image displays a musical score with ten staves of music. Each staff is annotated with various metrical phrase and section labels, represented by small boxes with text inside, connected to the music by horizontal lines. The labels include:

- Staff 1: L1, L1, J3, G2, M1, M1
- Staff 2: M1, M1, M1
- Staff 3: D4, B1, J1, B1, N1, C2, E1, A2
- Staff 4: A2, E2, A2, L2, H1, E3, B2, A1
- Staff 5: Q1, B3, L2, H2, E3, B2
- Staff 6: A1, Q2, B3, B3, I1, C2, D1
- Staff 7: D1, D1, D1, E1, E1
- Staff 8: D1, E1, E1, E1, G2, G2
- Staff 9: G2, G3, G2, G4, S1, G4

The annotations are placed above and below the musical staves, indicating the structure and organization of the piece.

The image displays a musical score for guitar, consisting of ten staves of music. Each staff is accompanied by a series of chord diagrams and fret numbers, indicating the specific fingerings and positions for each chord. The chords are labeled with letters and numbers, such as Q1, L2, J1, E4, A5, D2, C3, E5, A6, H3, F2, H2, F3, H2, B4, S1, A3, D1, S1, A3, D1, I1, I2, B2, I3, Q3, D3, E2, E2, D4, E2, E2, D4, J4, C3, F2, V1, E2, C4, V1, S2, C4, S2, X1, and J5. The music is written in a standard treble clef with a key signature of one flat (Bb) and a 4/4 time signature. The notation includes eighth and sixteenth notes, rests, and bar lines. The chord diagrams are placed above the corresponding notes on the staff, showing the fret number and the finger to be placed on that fret.

The image shows a musical score for the Presto movement of the Sonata G-minor Violin Solo by J.S. Bach (BWV 1001). The score is annotated with structural analysis labels. The labels are placed above and below the musical staves, indicating the boundaries of various structural units. A thick line is drawn across the score to indicate the second section/repeat. The labels include J6, K1, J5, F5, H3, K2, I2, G2, M1, A2, B2, J1, B5, and N2. The score is written in G minor and 3/8 time.

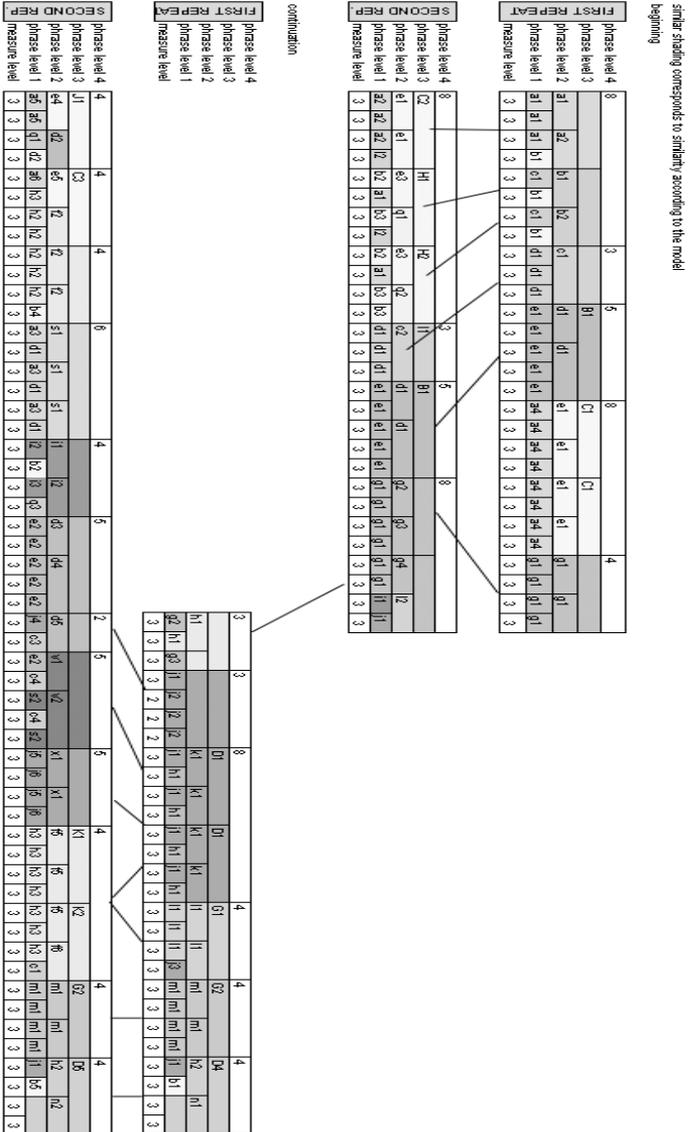
Figure 6-110. Presto from the Sonata G-minor Violin Solo by J.S. Bach (BWV 1001). Output analysis from computer model. Start-oriented interpretation. NB. Repeats and chords omitted. Second section/ repeat indicated by a thick line.

### Analysis of the Presto movement of the Sonata G-minor Violin Solo—general form structure

The music of J.S. Bach is often used in discussions of musical structure, probably because it is both well known to and appreciated by a Western audience, and offers a great variety of interesting structural problems. A comprehensive analysis of the complex structure of a work by Bach should, of course, involve analysis of polyphony and tonality aspects but this is not the objective of this model. The focus here involve only aspects of melodic surface structure common to musical styles of different origin. Thus, it is important to recall that the following analysis does not involve aspects of composition that would be required for a culturally learned understanding of a Bach piece, but is limited to the listener's basic conception of the sub-structural elements. Figure 110 displays the output analysis (start-oriented interpretation) of the *presto movement* from the *Sonata G-minor Violin Solo* (BWV 1001). This is actually a solo work, which makes it especially suitable for the application of the current model.

Even if the start-oriented solution probably feels a bit awkward in some respects to most Western listeners who are familiar with this piece, it is useful to demonstrate the fundamental structural relationships that are revealed by this analysis. Reduced to a schematic view with the music omitted the result can be viewed in Table 10 below. This scheme (and the output score displayed above) demonstrates a rather complex structure, generally consisting of phrases of one measure (3/8) at the lowest level, the level at

Table 36. Schematic overview of the output analysis from the computer model of the phrase structure of the presto movement from the Sonata G-minor Violin Solo by J.S. Bach. The phrase structure of each repeat section is displayed in parallel. Corresponding phrases in the two sections are connected by lines. Similar phrases are indicated by similar shading.



which the metrical structure is determined. On the next level, according to the output of the model, the phrases are generally two or three measures long, determined mainly by repetitions of pairs of the lower level phrases. On the third phrase level, the model does not offer a consistent interpretation, since mere clustering of lower level phrases are not yet included in the implementation. However, the structure that emerges is super-phrases of 3, 4 or 5 phrases, seemingly without any given periodicity.

In fact, the structure at the intermediate phrase levels resembles more or less the structure of the Norwegian *gangar* melodies, with the free combination of lower level phrases into higher order phrases, with no obvious higher order symmetry. Moreover, the phrases contain elements of one another in new combinations, which also contributes to the impression of an evolving phrase structure. On the other hand, the global repeated sections make it more related to the Swedish *polska* melodies, with the asymmetric intermediate phrase structure based on symmetry on both local and global levels.

However, the culturally experienced listener may conceive a higher order structure, which is not covered by the analysis performed by the current model, namely the structure determined by the development of tonality and harmonic patterns. This follows the typical pattern of moving from the key of the tonic to the key of the dominant at the end of the first repeat, while the second repeat section shows the reversed development. Such higher order periodical symmetry regarding tonality cannot be found in e.g. the Norwegian *gangar* melodies.

Even if we were to incorporate tonality evolution in the analysis, we would still find the asymmetry on intermediate levels. The repeat structures - the sequences - which determine the asymmetry in the analysis performed by the model are also the vehicle for the transformation of tonality in this piece.

The lack of consistent symmetry on intermediate phrase levels, in combination with symmetrical implications by occasionally repeated longer phrases, creates an ambiguous and vague structure at higher levels. Thus, the melody may be conceived basically as a chain of sequenced phrases comprised of similar elements, sometimes tail-biting each other like the Norwegian *gangar* melodies, sometimes implying higher order symmetry.

This creates a focus on the low level structures, latticity, the feeling that the combination of low level phrases is arbitrary and accidental – free. The mixture of symmetrical implications at mid level and the latticed structure is a difficulty for the model to interpret.

The image shows two staves of musical notation in 3/8 time. The first staff has four measures of music. Brackets above the staff group the measures into four phrases labeled A, A, A, and B. The second staff has four measures of music. Brackets above the staff group the measures into four phrases labeled C, B, C, and B. The notation includes treble clefs, a key signature of one flat, and various rhythmic values such as eighth and sixteenth notes.

Figure 6-111. Beginning of Presto movement. Two different interpretations of higher order phrase structure indicated by brackets. Lower level phrases identified as similar by the model indicated by capital letters.

## Latticity by sequence phase ambiguity

Consider the very beginning of the movement as displayed above in fig. 111.

The lower phrase interpretation is the one chosen by the model, while the upper is an alternative interpretation. Both these interpretations are possible for me to conceive, and probably there are even other possible groupings to find among listeners acquainted with the piece. The problem for the model as well as the listener is that the sequence of B and C sub-phrases is structurally significant but has two possible phases. Further, the segment A A A B can be conceived as a sequence pair with different endings A A, A B, but is latticed due to the lack of structural integrity of the first half. The segment A A A is on the other hand discrete, but is not supported by the continuing structure, which emphasizes a group size of two measures.

The cues for determining the phase of a sequence within the model are mainly:

1. the structural discontinuity of sequence boundaries;
2. the structural integrity of the sequence, based on similarity, latticity etc.;
3. the prominence of the first repetition, based on the assumption that the first repetition in an array of overlapping repetitions is more likely to be recognized than a subsequent start;
4. the delineation of the sequence with regards to other sequences.

Here the reduction of sequence prominence value due to the lack of integrity by latticity and weak local mid-boundary indication is not enough to stop the AA AB sequence from prevailing because it has a discrete start, and more importantly, is succeeded by a sequence CBCB, which has a discrete ending because of the beginning of a different sequence and high similarity. This divides the initial part into two phrases of equal length (four measures each) with perfect symmetry. The model uses in other words, the context to determine the solution.

My intuition leads me to the alternative (upper) interpretation, probably because of either: (1) the integrity of the start at B, which emphasizes the conception of the start of a new event; (2) BCBC is the first discrete sequence-start; (3) the relative distance after the first note of C which makes it an ending or; (4) because of the tonal tension employed in the repetition of the B sub-phrase. However, I end up with a feeling that the sequence BCBC was being interrupted at the implicated repetition by the next sequence, thus with a more complex structure than what is predicted by the model, 3+2+2+1 instead of 2+2,2+2.

Obviously, this structure is ambiguous at the two-three-measure phrase level. At least according to this introspective reasoning<sup>108</sup> it creates a latticed structure that emphasizes the more consistent lower level grouping. A part of a phrase can be repeated and sequenced to form new phrases, which does not need to be in a symmetrical relationship to the previous phrase segment, since the conception of the structure is not built on higher order symmetry, a general melodic principle known as *fortspinnung*.

This was but one example of numerous examples of ambiguous sequencing in this music. In this case, it was above the threshold for the model to make a prediction that was not even the dominant interpretation of the creator of the current model. It is worthwhile to stress that even if the model is designed not to give hierarchical interpretations when the structure is not discrete, it has to handle structural implications and will thus give results that are highly ambiguous depending upon the ambiguity of the melodic structure. This could be reflected in

---

<sup>108</sup> The ambiguity of melodic structures are, however, not specific to me, as is indicated by the experiments referred in section 6.3.7.

the implementation of the computer model by output of different basic structures and ratings of salience, but this has not yet been implemented.

### Higher order structure by quasi repetition

One feature of this piece which emphasizes the structural ambiguity, is that it is a *moto perpetuo*, almost the entire movement consists of notes of equal length. The highly latticed structure would easily have made it become just 'a lot of tones' for the listener, a load of tones with no higher order, only short figures piled upon each other. This is also probably true for some listeners who are new to the piece.

But when considering the summarized form schemata, certain structural features recur, forming a unity of the piece besides the previously mentioned tonal structure. In general, the second repeated section contains the same phrases at different levels as the first repeated section in about the same order. It is not an exact repetition, but all the phrase elements of the first repeat do actually occur in the second repeat in more or less varied form. (Corresponding phrases are indicated by connected lines.) The perfect repetition of the B1 macro phrase and the next to perfect repetition of the very ending contributes to the *motivic* unity of the piece.

The form scheme also demonstrates the *evolving* structure, in which prior phrase end elements become the start elements and vice versa. Thus, the b and c sub-phrase elements consistently change place within the higher order phrases, alternating as start- or end-elements. The sub-phrases general similarity is consistent, and, as in the Norwegian *gangar* music, new sub-phrases can be regarded as slight variations of prior phrases which makes the categorization of the phrases ambiguous. The identification of phrase similarity here is contextual; it depends on successive identification where later phrases are thought to relate to previous and the limit of similarity difference between two members of the same category is determined mainly by how well it reflects the segment structure. In this particular case, the structural integrity of most sub-phrases is so low that categories overlap greatly and cause some inconsistencies in the identification. This shows e.g. by gaps in the succession of root letters.

One particular form of quasi repetition, which originates from the composition process, is the inversion of a melodic phrase. Consider the following phrase segments, taken from the second major repeated section.



Figure 6-112. Sequencing by inversion. From Presto movement of Sonata G-minor Violin Solo by J.S. Bach (measures 101-104)

This is a typical example of an inversion in this music, it is not a perfect inversion, but is slightly altered to fit into the tonal pattern. An important question is whether this structural, compositional principle is significant for the conception of the listener?

The assumption within this method is that properties of pitch change, not only with regard to global successive register change, i.e. melodic direction, but also in terms of change of interval structure, may be significant in the process of conceiving a melody. Thus similarity with regard to interval structure can be significant. For the current passage, this could be

described as a motion in thirds in one direction for three beats, followed by a reversed motion in steps, which is equal to the two parts of the division. Consequently, the change of pitch interval structure is assumed to function as the discontinuity mark that is necessary for the sequence to be discrete.

Such similarity regarding structural properties regardless of similarity of melodic direction is generally assumed to be of weak structural significance within the method and is covered by the X-sequences (see section 6.2.3.3). Such sequences generally require either already identified start- and mid-positions by previous sequences or discontinuity at global level to be recognized. But this kind of inversion is a special case in which inverted sequences are identified if they exhibit similarity of inversed local melodic direction at sequence type NIL 3 and NIL 4. (See section 6.2.3.3). The indication is, however, generally weak, since it does not possess close similarity in terms of actual melodic direction and can thus easily be overridden by a closer similarity. This kind of similarity therefore assumed to be recognized if the position of the start of the sequence is concurrent with the established sequence structure or with global structural breaks.

In the models analysis of this piece, inversion is also recognized as determining the similarity between the start of the first and second major repeat section (see fig. 113 below). The start of the second major repeat section is evidently determined by the global repeat. It is emphasized also by the global change of interonset interval which marks the end of the first major repeat section. However, the consistent similarity with regards to inner pitch interval structure and change of melodic direction, which can be described as a near to perfect inversion of the first eighth measures of the first major repeat section, is recognized by the model as a pertinent similarity. This reflects the assumption that such a similarity may be significant for the segmentation of the melody.

Inversed sequence identified by the model is indicated by the box (0 162 X 2 17)

Start of first repeat section (beat position 0)



Figure 6-113. Beginning of first and second repeat of the Presto movement from the Sonata G-minor Violin Solo by J.S. Bach

## Decomposition of sub-phrase units – establishing and challenging the metrical structure

The metrical structure of this piece is determined by the repeat structures, the sequences, which create a general periodicity of grouping of six sixteenth notes at measure level. The beat-grouping is rather ambiguous and weak, but the complex metrical analysis model ends up with a grouping in 2/16 at beat level instead of the alternative 3/16 or 6/16 at beat level. This is due to more consistent sub-sequencing at the 2/16 level than on the 3/16 level and because the 6/16 groups contain too many elements to be at the ideal beat level and also are not

marked consistently by primary grouping indications. Once again, the lack of contrast regarding the rhythmic/*interonset* structure makes the melodic structure generally ambiguous; there is no “music off”, only “music on”.

Thus, the grouping creates the meter in this piece and the meter influences grouping by metrical implication. But in the music by J.S. Bach<sup>109</sup> the metrical structure is typically challenged by the grouping structure, Consider the end-oriented interpretation by the model of the first major repeat of the Presto movement (Fig. 114 below).

---

<sup>109</sup> This was evidently a more or less general feature of the style of composition at the time

gminorpresto.part1.mf

The figure shows a musical score for the first repeat of the Presto movement from the Sonata G-minor Violin Solo by J.S. Bach. The score is presented in 14 staves. Above the musical notation, various chord labels are placed to indicate phrase analysis. The labels include A1, A2, B1, C1, D1, E1, F1, G1, H1, H2, J1, J2, K1, B1, and H1. The music is in G minor and 3/4 time, starting with a tempo marking of 120. The analysis shows a complex structure of phrases and sub-phrases, with some labels appearing multiple times across different staves.

Figure 6-114. First repeat of the Presto movement from the Sonata G-minor Violin Solo by J.S. Bach. Phrase analysis output from computer model. End-oriented interpretation. (The inconsistencies regarding the identity of the phrases in comparison with the analysis presented above (fig. 113) is due to the different input, in the present figure only the first major repeat was analyzed)

The end-oriented interpretation which is based on the start-oriented, ‘moves’ the actual phrase boundaries to the point of largest discontinuity or distance within the seven closest events and in relation to the neighboring sequence boundaries and the sequence to which it belongs.

The crucial difference between the end-oriented and the start-oriented interpretations is that the latter does not have to conform to the metrical grid. End-oriented interpretation is thus fundamentally non-metrical. It reflects the view of the listener as if there were no meter present locally, where as the influence of the metrical structure is indirect through the relationship to the *virtual starts* which are identified by the start-oriented phrase analysis. Hence end-oriented interpretation regards basically discontinuity and similarity between events, not beats.

The level of end orientation is obviously different among different listeners, and probably very much influenced by cultural learning.<sup>110</sup> The model gives so far a weighted interpretation, which basically aims to demonstrate the principle than to provide the full variety of possible interpretations.

In the beginning there is no difference between the end- and start-oriented interpretations, since the points of greatest structural distance coincide. But when the sub-phrase C1 enters, the relative proximity between the notes following the pivot note moves the start to the second note of the virtual start in the end-oriented version.

The smaller change of pitch between the segments of sub-phrase sequence D1 moves the end-oriented start back to the position of the virtual start. In the next sequence section, once again the consistent greater pitch distance after the note at the virtual start moves the end-oriented start to the second note of the original phrase, while the sequence determining the next sequence section has its greatest distance before the virtual starting point.

If we move on to the H1 and H2 sequences at sub-phrase level the phrase structure determined by end-oriented interpretation is clearly in conflict with the metrical structure. It reveals discrete primary groups of four sixteenths each, which are identified by the metrical analysis as a strong indication of primary grouping with metrical significance. This grouping originates from the segment H1, the continuation of which is sequenced in conflict with the established metrical structure.

Do we then start to conceive a 2/8 meter? Probably not, people seem to be very consistent in maintaining pulse and time conception once a metrical structure has been identified, at least according to the results of the experiments 2 (see *Metrical Analysis* section 5.3.4.3, *Discussion*). But maybe it would be conceived as a temporary inhibition of the metricity, as a temporary stop or confusion of time.

The slurring of the original notation, which is concurrent with the above grouping in 2/8, indicates that Bach also intended this phrasing.<sup>111</sup> A slur creates a stronger accentuation of the first note of the slur, creating an impulse effect, hence metrical implication.

This seems to challenge the basic assumption of congruency of phrase structure with metrical structure based on the assumption of primacy of metrical structure. My interpretation

---

<sup>110</sup> The relationship between preferred phrasing and cultural learning would be possible to systematically study by using the interpretation rule system developed by Sundberg, Friberg and Frydén, in which the different parameters of typical phrasing in Western music are possible to control.

<sup>111</sup> The groupings of the end-oriented version do generally comply with the groupings indicated by slurs in the Bach's original notation.

is, however, that it is rather a matter of consistency of metrical structure; According to the fundamental assumptions about the nature of metrical conception, metrical structure may be conceived in some parts of a melody, while it may be absent in other parts. That is why the higher levels of metricity, at measure level, have to be analyzed on basis of the grouping structure.

To conclude, it seems that indication of discrepancy between metrical structure and grouping structure, is very common within the repertoire of Western classical art music compared to the dance music styles have been discussed earlier. Is it a manifestation of the music being freed from the obligations of the function of dance music? Or is it a result of the possibility of the ensemble to let the metric structure be covered by the accompaniment?

### 6.3.5.2 20<sup>th</sup> century Western popular music – separation of meter and phrase structure

#### Phase displacement of metrical grouping and phrase structure

The question that concluded the previous section hypothesized the influence of the ensemble on the relationship between segment structure in melodies and metrical structure. Obviously the existence of an ensemble, by which the metrical structure can be maintained regardless of the structuring of a melody, opens a number of possibilities for the metrical function of the melody. It can as in e.g. solo sections within some Balkan and Jazz music, be entirely without metrical structure creating a parallel grouping structure. It may challenge the metrical structure by indicating a parallel meter, which does not conform to the metrical structure. Or it can contribute to the metrical structure by period congruency but different phase.

The last type, which is extremely common in different types of metrical accompaniment in a variety of musical cultures with ensemble music tradition but also in e.g. 20<sup>th</sup> century Western popular music melodies, creates a drive, enhances the metrical flow by the interaction between impulses generated by different layers within the musical structure. The complex metricity with parallel pulse layers which is typical for e.g. the Norwegian *gangar* and *balling* music sometimes indicates a different pulse from the dance meter, which makes it impossible to trace the dance meter from the melody alone. But this is an exception within European dance music melodies. As has been indicated by the success of the current method tested on different repertoires, the melodies generally contain the metrical information in these repertoires.

With genuine ensemble music, the situation is different and the monophonic limitation of the current method becomes apparent. The model does not receive the necessary information to track the entire structure.

But still, even when the accompaniment is apparently missing, the current model can extract information about the phrase structure, which makes it possible to trace higher-level metrical structure because of congruency between phrase structure within the melodic structure and metrical structure.

Fig. 115 contains an example of such a melody taken from the repertoire of 20<sup>th</sup> century popular music, the Indiana Jones theme by the American film music composer John Williams. It is questionable to use the term melody to characterize this piece, as in the case of the Mozart examples used in section 6.2.6, since it is essentially an ensemble composition. The

proposed melody is abstracted from the score, but I am hereby making the assertion that it is possible to conceive this theme part of this composition as a one-layered melodic structure. This is indicated by the frequent use of reduced versions of multipart compositions as monophonic melodies as e.g. cell phone signals, as school music or in popular arrangements where the complex structure is reduced.

A. Start-oriented interpretation

**indianaj.theme.mf**

B. End-oriented interpretation

**indianaj.theme.mf**

Figure 6-115. Indiana Jones theme melody, abstracted from a composition by John Williams for the motion picture "Raiders of the lost Ark" (Paramount pictures 1981). Computer model output. A start-oriented and B end-orientation.

Note that the inconsistent identification of phrase similarities in the end-oriented version is due to its origin in the start-oriented interpretation.

This is a case where the start-oriented version becomes absurd. Due to the deficiencies of the graphical implementation of the model, which not allow phrase boundaries to be displayed within note events, the phrase markings are intertwined. In addition, the indications of phrase endings of the A phrases are so salient, marked by long notes exceeding more than two beats; still the model in the start-oriented version will interpret the dotted eighth note-sixteenth note figure as an upbeat to the main impulse point, which becomes a salient start due to its relatively long duration and the position in the sequence.

Most Western listeners would, however, probably conceive the melody more in the manner of the end-oriented version, regarding the upbeat figure as more of a start than an end of the phrase. That is not to say that they would conceive a different meter, since the virtual phrase start would be concurrent with the start-oriented version. On the contrary, the consistent up-beat of the reduced melody will rather emphasize the accentuation and impulse quality of the virtual starting point. The start-oriented version will, therefore, serve as the basis for the analysis of meter at higher levels, while the end-oriented version will display the predicted conception of phrase structure.

The model provides both outputs, but is designed to favor the end-oriented version when it is indicated by strong means such as the interonset distance in the current example.

The similarity between this example and e.g. the Mozart G minor symphony example provided in section 6.2.6, is obviously not arbitrary. The current example belongs to the central Western art and popular music tradition, the common-practice repertoire of Western music to use a term proposed by Temperley (2001), which has been an exponent of the 19<sup>th</sup> and 20<sup>th</sup> century societies.

### **Fascinatin' rhythm – phrase structure in conflict with the metrical structure at measure level**

The next example is taken from the repertoire of the early 20<sup>th</sup> century American popular song, designated the Tin Pan Alley repertoire (Middleton 1990:38). The development of this repertoire was strongly influenced by and intimately connected with the emergence of jazz music, where the metrical accompaniment became a more prominent part of the ensemble playing than in European 19<sup>th</sup> century popular music. With some exceptions, this has been an ongoing development, with a parallel decrease of the importance of melody (and ultimately also functional harmony) as a significant element within the musical structure.

## A. start-oriented version

**fascinating2.mf**

Figure 6-116. Reduced melody of *Fascinatin' Rhythm* by George and Ira Gershwin (transcribed from performance by Ella Fitzgerald, Verve records 1960). Lyrics, accompaniment etc. omitted. Computer model output. A start-oriented version.

This song, *Fascinatin' Rhythm* by George Gershwin, (Fig. 116) (Gershwin 1924, see also below under *A Foggy Day*), , uses stylistic elements borrowed from jazz and related popular musical styles of the time, these include the rhythmic division of beats, in combination with the stylistic elements of the European popular music, such as the significance of the melodic and harmonic patterns to constitute the song.

Also in this case the use of the melody line alone is questionable relating to the entirety of the composition. But since this model only considers melody, it is valuable to consider its behaviour when important parts of the musical structure are missing.<sup>112</sup> Here the model provides an output, which is based on the identification of sequences determined by similarity. The sequences identified by the analysis are these:

```
(0 16 NIL 4 16)
(0 7 C 3 5)
(0 4 NIL 3 3)
(7 4 Y 0 4)
(16 7 C 3 5)
(16 4 NIL 3 3)
```

<sup>112</sup> In this case I have used my own transcription as input to the analysis, mainly because the traditional notation leaves certain important features of the melodic interpretation unnoticed, such as the anticipation of beat onsets and the close to triple beat division. Since these properties of the performance are often impossible to trace in typical traditional notations, I have chosen 12/16 as the notated time rather than the traditional 4/4 notation.

*Chapter 6*

(23 4 Y 0 4)  
(32 16 C 3 10)

The sequence analysis finds two major sequences dividing the melody into two sections, each determined by a 16 beat sequence, the first of which exhibits perfect type 4 similarity. The second 16 beat sequence is defined by more general similarity (type C), but rather salient through a very similar first half of the sequence.

This second major sequence starts within the rest event, which is possible within the model at the last beat event due to the rhythmic structure (see section 6.3.3.3). The start position is here implicated by the preceding sequence, which gives an implied start indication at the beat event whether it is a rest or not.

Below this level, the structure starts to become more complicated. The first major sequence is decomposed into a sequence of 7 + 9 beats through the (0 7 C 3 5) sequence, implying five beats of general melodic contour similarity, which is in turn subdivided through the (0 4 NIL 3 3) sequence, subdividing the first 7 beat sequence half into two segments of 4 + 3 beats, by symmetrical implication also inherited to the second half of the 7 beat sequence, implying the (7 4 Y 0 4) sequence. The sequence structure thus implies a three level phrase hierarchy in the first part of the melody with consistent, slightly asymmetrical, duple division within a general framework of perfect symmetry.

The second major sequence (32 16 C 3 10), does not become subdivided because the melodic events are all gathered in the first part of the sequence half. Even if there is a small repeated figure in the end of the first half of the sequence, the long note that marks the ending of the phrase does not imply any melodic activity that could form the basis of a perfect duple division. And the possible sub divisions of the first half of the sequence half are too ambiguous to be interpreted by the model.

Obviously, the perfect symmetrical subdivision of this latter sequence into two symmetrical parts of 8 beats, which can be conceived when listening to the entirety of the song and is determined by e.g. harmony is not signified through the melodic structure in other terms than lack of melodic activity. The hierarchic subdivision of the first part of the melody obtained by the model – two segments of 16 beats subdivided on the next level into 7 + 9 beats and on the next level into 4 + 3 + 4 + 5 beats – conforms to the composition of the

**fascinating2.mf**

Figure 6-117. The phrase structure of the models analysis of Fascinatın' Rhythm, used as input to the transformation of metrical structure from start-oriented interpretation resulting in a re-analysis of the metrical structure at measure level.

lyrics of the song, which thereby corroborates the cognitive relevance of the interpretation of the phrase structure performed by the model. This would lead to a metrical substructure of 4 + 3 + 4 + 5 beats, as is indicated in the fig. 117 above.

Does this interpretation of the metrical structure based on the phrase structure, which has proved to be successful in the previous examples, really reflect a cognitive reality, a listener experience? Or, is this an example of phrase structure of melodies – grouping - without any relationship to and relevance for metrical experience?

Judging from the accompaniment, it bears no sign of any change of meter, rather the opposite; it indicates a consistent 4-beat metrical grouping, creating a periodicity of 12/16 time (or 4/4, depending on the mode of notation).

However, the lyrics, which complying with the indicated metrical interpretation of the melody as implied by the phrase analysis and underlined by the title of the song, indicate that the phrase structure may be conceived to be implying a metrical grouping conflicting with the periodicity implied by the accompaniment; creating a temporary confusion of meter at measure level. This is possible within a consistent symmetrical meter through the separation of melody and accompaniment.

## Chapter 6

According to this interpretation, the metrical implication of the phrase structure of the melody is significant and essential for the understanding of what is *fascinating* in the rhythm of this piece.

The image displays a musical score for the song "A Foggy Day" in G major, 4/4 time. The melody is written on a single staff in treble clef. The lyrics are written below the staff, with numbers 1 through 33 indicating the measure numbers. The lyrics are: "A fog-gy day \_\_\_ in Lon-don town \_\_\_ Had me low- \_\_\_ and had me down \_\_\_ I viewed the morn-ing with a-larm- \_\_\_ The Bri-tish Mu-se - um had lost its charm \_\_\_ How long I won - dered, could this thing last? - But the age of mir - a-cles had - n't passed - for sud - den-ly - I saw you there - And through fog - gy Lon-don town the sun was shin - ing ev' - ry - where." The melody consists of quarter, eighth, and sixteenth notes, with some rests and ties.

Figure 6-118. "A Foggy Day" by Ira Gershwin and George Gershwin (1937), from Middleton (1990:184)

### A Foggy Day – comparison with segmentation based on paradigmatic analysis

In the book "Studying popular music", the musicologist Richard Middleton applies the paradigmatic analysis developed by Ruwet (Ruwet 1987) to another song by Gershwin, "A Foggy Day in London town" (Middleton 1990:183-189). (see fig. 118) As the melody in this song is closely connected to the harmonic structure, a monophonic melodic analysis leaves important aspects of the musical structure largely overlooked. Nevertheless, Middleton applies paradigmatic analysis based on melody structure alone to demonstrate how the method may be used to segment the piece. This makes the analysis interesting in relation to the current method. Middleton's application of the paradigmatic method results in a segmentation within three parallel hierarchic levels, described as follows:

- Level 0: X No strictly repeated units. But using criterion of transformation . . .
- Level I:  $A + X + A^1 + Y$  ( $A^1 = A$  with slight rhythmic variants), or rather  $A + B + A^1 + C$  (all units same length, or very nearly).  
But  $B = b + c$ ;  $C = b^1 + d$  (see next level). Moreover, B and C are equivalent in length and end on held notes. Thus  $C = B^1$ . Hence:  $A + B + A^1 + B^1$ .
- Level II:  $A, A^1 = a + b$  (words assist segmentation here)  $b = a^1$  (transformation through transposition). Hence  $A, A^1 = a + a^1$ .  
Similarly,  $B = b + c, B^1 = b^1 + d$  ( $b^1 = b$  at higher pitch with slight changes in some interval sizes).
- Level III:  $a = x_1 + y_1$  (words again assist).  
 $a = a_1 + b_1$  (same length).  
 $a = a_1 + a_1^1$  (transformation through expansion of 3rd to 5th).  
Similarly,  $b = b_1 + b_1^1$  (sequence).  
 $c = c_1 + c_1^1$  (expansion and inversion).  
But all these are variants of  $a_1$  (see Ex. 6.5).  
 $d = d_1 + d_1^1$  (next level) +  $e_1$ .

Figure 6-119. Paradigmatic analysis of "A Foggy Day" (Middleton 1990:185-187).  
Description of hierarchic levels, top-down

Ex. 6.4

Figure 6-120 Illustration of Levels 0-II

Figure 6-121 illustrates the transformations between segments at level III. The segments are listed on the left, and the transformations are listed on the right. A bracket on the right groups the transformations 'transposition + inversion' and 'transposition + expansion' under the label 'contraction'.

Figure 6-121. Illustration of level III

The analytic criteria used in the segmentation described in fig. 121 above resemble the segmentation by similarity criteria used in the sequence analysis with implicative segmentation by symmetry and hierarchic consistency within the current model. This suggests that the current model will obtain a result by using these criteria for segmentation, rather than segmentation by discontinuity. And indeed, the result given by the sequence analysis corresponds with the identified segments in Middleton's analysis to a high degree.

A. Start-oriented interpretation

**foggyday.mf**

A1 A1 A2 A3 A4  
B1 B1 B2 A3 A5 B3  
A2 A4 A6 A7 A5 A8 A4  
B2 B2 B2 B4 C1 C1 D1

B. End-oriented interpretation

**foggyday.mf**

A1 A1 A1 A2 A2 A3 A4  
B1 B1 B1 B2 A3 A5 B3  
A2 A4 A6 A7 A5 A8 A4  
B2 B2 B3 B4 C1 C1 D1

Figure 6-122. Output from analysis of “A Foggy Day” by the current model. A start-oriented and B end-oriented interpretation

The result of the major section analysis, which divides the melody into two segments A1 and A2 (A1+B1, A2+B2), is not included in the fig. 122, due to the limitations of the graphical interface, which allows only three levels to be displayed simultaneously.

The lowest phrase level in the analysis obtained by the current model corresponds to Middleton's level III, the middle phrase level corresponds to level II etc.

Middleton discusses certain weaknesses of the paradigmatic model of analysis, in particular the question of which criteria should be used to determine equality. He argues that the use of criteria will influence the result. The current model handles this problem by the gradation in levels of similarity, reflected in the different levels and types of sequences, and the contextual use of similarity as a ground for segmentation, rather to assign a certain fixed level of importance to different structural parameters. If the structural contrast generally is low or the structure exhibits great variability, more general similarity may be significant. In the opposite case, more general similarity will be overridden by more precise, high level similarity between arrays of events.

Another problem discussed by Middleton, which is interesting in the current context, is the delineation between similar and contrasting elements. At which point of structural distance will two contiguous segments be regarded as contrasting elements or similar elements?

This is definitely an interesting question in relation to the current model. The identity of segments obtained by sequence and discontinuity analysis is in the current model based on a) the successive, contextual determination of sequences by similarity and discontinuity and b) categorization of relative levels of similarity according to the principles of relative similarity (see section 6.2.4.4, *Symmetrical implication*).

Successive, contextual identification of similarity is based on the result of sequence analysis: Sequences, exhibiting similarity above the X level – similarity by general structural properties, are basically treated as defining a category of similarity, while segments identified by division by structural discontinuity are treated as belonging to different categories. The b) categorization implies that the root ID of a certain segment depends on the relative similarity to other segments; if three segments are similar but the similarity difference between two of the segments and the third is greater than 1/2 the total similarity, then they will be regarded as different categories - contrasting elements.

In the above example, this procedure leads to a differentiation between segments A and B at the higher phrase levels. These segments originally determined by the R-sequence (2 16 R 3 8), which implies melodic pitch contour similarity at low beat-to-beat level, and (2 16 X 0 8) which implies compatible general pitch change and duration change (inversion) of the first 8 beats are reclassified from A segments to B segments due to the significantly higher similarity between segments with the falling melodic pitch contour than the raising melodic pitch contour. In another context, they might have been regarded as variants of one another, designated by the same root ID. The handling of contrast and similarity within the current model is thus basically contextual.

The general problem addressed by Middleton, regarding contrast and similarity as defined by different structural parameters, is not addressed in the current model. But since the model identifies similarity on different levels and with regard to different parameters, it would be possible to designate the identity of segments by a more elaborate identification code involving e.g. three dimensions (X x 1). Thus it would be possible to demonstrate even non-

syntagmatic properties of the structure such as the general rhythmic similarity between all a and b segments in the above example.

### 6.3.6 Sequenced and non-sequenced phrase structure within asymmetrical metrical structure

#### 6.3.6.1 Balkan dance music with asymmetrical beat structure – Transformation of phrase identity within a symmetrical framework

**07.alfanska.racenica.mf**

The musical score is titled "07.alfanska.racenica.mf" and is written in 6/8 time. It features eight staves of music. Above the staves, phrase segments are identified with labels such as A1, B1, C1, D1, E1, F1, G1, B2, C2, D2, D3, G2. Brackets indicate the duration of these segments across the musical notation. A tempo marking of 180 is shown at the start of the first staff.

The image displays a musical score for guitar, consisting of eight staves of music. Each staff is accompanied by guitar chord diagrams. The diagrams are represented by boxes with letters and numbers inside, indicating the fret and string for each finger. The music is written in a key with one sharp (F#) and a 4/4 time signature. The chord diagrams are as follows:

- Staff 1: B3, E1, F1, D1, C2, E1, G2, B3, E1
- Staff 2: E1, F1, C3, E1, G3, E1, J1, J2
- Staff 3: E2, C3, E1, G3, E2, J1, K1, E1, C3, E1
- Staff 4: G3, E1, J1, J2, E3, C3, E1, G3
- Staff 5: G1, J1, M1, G1, H1, T2, I3, H2, T2
- Staff 6: E4, C4, D1, N1, N2, E1, J3, J2, E5, C4, N1
- Staff 7: N2, B4, J3, F1, E4, C4, N1, N2, E1, J3
- Staff 8: E5, C4, J2, N1, N2, B4, J3, F1, E6, C5, E3

The image displays a musical score for the traditional Bulgarian dance melody 'Alfanska racenica'. The score is presented in a single system with eight staves of music. Above the staves, various musical phrases and sections are identified with alphanumeric labels enclosed in boxes. These labels are connected to the corresponding notes on the staff by thin lines, indicating the model's analysis of the melody's structure. The labels include:

- Staff 1: E4, C6, E5, E4, E6, C5, E3, E4, C6, E5
- Staff 2: H1, C5, E5, J4, C7, E6, Q1
- Staff 3: H2, C5, E5, J4, C8, E6, F1, J1, F1, I5
- Staff 4: J1, I5, E7, J1, F1, I5, E7
- Staff 5: J1, I5, F1, J1, K1, C7, R1, B5, C8
- Staff 6: J2, K1, C7, R1, F2, C8
- Staff 7: J1, K1, C7, R1, B5, C8, C10, J9, F1
- Staff 8: J2, K1, C7, R1, F2, C8, F1, E4, C4, NT
- Staff 9: E5, C4, N1, N2, J3, B4, J3
- Staff 10: E4, C4, N1, N2, E1, J3, J2
- Staff 11: E5, C4, N1, N2, J3, B4, J3, F1

Figure 6-123. *Alfanska racenica*. Traditional Bulgarian dance melody. Computer model analysis.

The asymmetrical beat structure of much Balkan dance music determines quite unambiguously the lowest phrase level in this repertoire. As the measure level is consistently and implies periodic, a repetition of a discrete, concurrent beat and primary grouping pattern, may be one of the factors influencing the often perfectly symmetrical phrase structure at middle levels of Balkan dance tunes, perhaps a reflection of a periodicity in dance patterns. The dance forms are also reflected in the section structure of many tunes, which can be compared to suite form structure. Many tunes consist of a chain of sections without any obvious super-structure, where the phrase elements of one section are formed by transformation of a phrase of a previous section. The phrase elements are in many tunes varied consistently even within sections through ornamentation and start-variation.

Consider the melody in fig.126, a traditional Bulgarian tune called “*Alfanska Racenica* “. This is a typical example of the above features of this repertoire.

One of the most typical features of this and other tunes of the same repertoire is the evolving phrase structure where new phrases grow out of elements of the former phrases, creating successively more and more overlapping categories (see also *Norwegian gangar*, section 6.3.4 and the J.S. Bach example, section 6.3.5.1.)

This becomes evident when looking at a schematic overview over the phrase and section structure.

Consider, for instance, the  $\epsilon$  phrases at the one measure level (level 1), which constitute the ending of the a2 phrase at the two measure level (level 2), but also the start of the closing of the d1 phrase at level 2 in the first section. It then becomes the start phrase of the second section, which is hence based on the ending of the first section. This start phrase on the second level, designated  $\epsilon$ , then transforms by keeping its one measure ending phrase through a transformation of this phrase thus forming the basis of the succeeding sections by constant variation.

The ongoing transformation of the phrases may be said to turn the phrase elements into gestures, rather than identifiable phrases on the one measure level. I doubt that the average listener can really keep track of this development of the lowest phrase level because relatively small number of pitch and rhythm changes which constitute the phrases and the overlapping similarity between categories. The similarity between transformed phrases of the higher levels would probably be easier to trace. The most obvious general repetition is probably the repetition of the entire sixth section. Also the transformation of phrases in the F1 and F2 sections would probably be salient enough to be recognized by listeners.

The figure 124 below displays the transformed phrases at level 1 categorized as belonging to the same root. The structural difference between phrase trees is not discrete, but overlapping, mainly due to successive variation between contiguous segments which define the variability within a category and the small number of unique events which define each root phrase. One example is the rhythmic similarity between E4 and F1 phrases that is greater than the rhythmic similarity within the E phrase category. Likewise the rhythmic and melodic contour similarity between the B1 and C1 phrases are equally as large as within in the C phrase category.

This reflects the lack of structural integrity at this phrase level. When the root identity overlaps become too frequent, in an ambiguous structure results, and makes the phrase structure less salient as melodic syntactic element.

Table 37. Scheme of the phrase and section structure in Alfanska Racenica from the analysis by the current model. (Due to deficiencies in the current computer implementation the root and number naming has gaps)

similar shading corresponds to similarity according to the result of analysis by the model

FIRST SECT.	section level	A																																
	phrase level 3	a1				b1				a1				b2				a1				b1				a1				b2				
	phrase level 2	a1	a2	a1	a2	d1	d2	c1	d1	d3	a1	b1	a2	c1	d1	d2	c1	f1	a1	b1	a2	c1	d1	d2	c1	f1	a1	b1	a2	c1	d1	d2	c1	f1
	phrase level 1	a1	b1	a2	c1	d1	d2	c1	d3	a1	b1	a2	c1	d1	d2	c1	f1	a1	b1	a2	c1	d1	d2	c1	f1	a1	b1	a2	c1	d1	d2	c1	f1	
	measure level	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
SECOND SECT.	section level	B																																
	phrase level 3	b4				b4																												
	phrase level 2	c1	b2	c1	b2	c1	b2	c1	b2																									
	phrase level 1	c2	g1	c2	f1	c2	g1	c2	f1																									
	measure level	3	3	3	3	3	3	3	3																									
THIRD SECT.	section level	C1																																
	phrase level 3	d1				d1																												
	phrase level 2	c2	e1	c2	e1	c2	e1	c2	e1																									
	phrase level 1	a1	g1	e1	f1	e1	g1	e1	f1																									
	measure level	3	3	3	3	3	3	3	3																									
FOURTH SECT.	section level	C2								C2																								
	phrase level 3	e1				e2				e1				e3																				
	phrase level 2	c3	e1	c3	e2																													
	phrase level 1	e1	g3	j1	j2	e1	g3	j1	k1	e1	g3	j1	j2	e1	g3	j1	m1																	
	measure level	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3																	
FIFTH SECTION	section level	D																																
	phrase level 3	g1																																
	phrase level 2	h1				h2																												
	phrase level 1	j2	i3	i2	o1	j2	i3	i2	o1																									
	measure level	3	3	3	3	3	3	3	3																									
SIXTH SECT.	section level	E								E																								
	phrase level 3	e4				e5				e4				e5																				
	phrase level 2	c4	b4	c4	e1																													
	phrase level 1	n1	n2	j3	j2	n1	n2	j3	f1	n1	n2	j3	j2	n1	n2	j3	f1																	
	measure level	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3																	
SEVENTH SECT.	section level	F1																																
	phrase level 3	e6																																
	phrase level 2	c5	c6	c5	c6	c5	c6	c5	c6																									
	phrase level 1	e3	e4	e5	e4	e3	e4	e5	e4																									
	measure level	3	3	3	3	3	3	3	3																									
EIGHTH SECT.	section level	F2																																
	phrase level 3	h1				h2																												
	phrase level 2	c5	c7	c5	c7	c5	c7	c5	c7																									
	phrase level 1	e5	i4	e6	g1	e5	i4	e6	f1																									
	measure level	3	3	3	3	3	3	3	3																									
NINTH SECT.	section level	G																																
	phrase level 3	i1																																
	phrase level 2	f1	j1	f1	j1	f1	j1	f1	j1																									
	phrase level 1	i5	e7	i5	f1	i5	e7	i5	f1																									
	measure level	3	3	3	3	3	3	3	3																									
TENTH SECT.	section level	H																																
	phrase level 3	j1				j2																												
	phrase level 2	k1	b5	k1	b5	k1	b5	k1	b5																									
	phrase level 1	c7	f1	c8	i9	c7	f1	c8	i9																									
	measure level	3	3	3	3	3	3	3	3																									
ELEVENTH SECT.	section level	E								E																								
	phrase level 3	e4				e5				e4				e5																				
	phrase level 2	c4	b4	c4	e1																													
	phrase level 1	n1	n2	j3	j2	n1	n2	j3	f1	n1	n2	j3	j2	n1	n2	j3	f1																	
	measure level	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3																	

Chapter 6

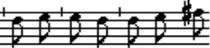
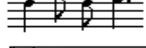
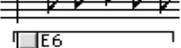
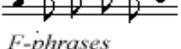
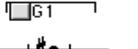
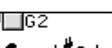
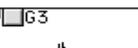
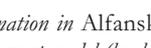
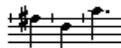
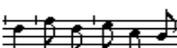
<p><i>A-phrases</i></p> <p>□A1</p>  <p>□A2</p>  <p><i>B-phrases</i></p> <p>□B1</p>  <p><i>C-phrases</i></p> <p>□C1</p>  <p>□C1</p>  <p>□C2</p>  <p>□C7</p>  <p>□C8</p>  <p><i>D-phrases</i></p> <p>□D1</p>  <p>□D2</p>  <p>□D3</p> 	<p><i>E-phrases</i></p> <p>□E1</p>  <p>□E3</p>  <p>□E4</p>  <p>□E5</p>  <p>□E6</p>  <p><i>F-phrases</i></p> <p>□F1</p>  <p>□F1</p>  <p>□F1</p>  <p>□F1</p>  <p>□F1</p>  <p><i>G-phrases</i></p> <p>□G1</p>  <p>□G2</p>  <p>□G3</p> 	<p><i>I-phrases</i></p> <p>□I2</p>  <p>□I2</p>  <p>□I3</p>  <p>□I4</p>  <p>□I5</p>  <p>□I9</p>  <p><i>J-phrases</i></p> <p>□J1</p>  <p>□J2</p>  <p>□J3</p>  <p><i>N-phrases</i></p> <p>□N1</p>  <p>□N2</p> 
--	---	--

Figure 6-124. Phrase transformation in Alfanska Racenica according to the successive similarity identification performed by the current model (level 1 phrases – sub-phrase level)

### 6.3.6.2 Non-sequenced asymmetrical segment structure within a symmetrical general framework –south Indian classical music

#### General problems

South Indian classical music is, in several ways, very different from the repertoire we have been examining so far. First, even if the analyses of especially folk music repertoires such as the Norwegian and Balkan music examples are based on very simplified score representation, in particular regarding ornamentation, the traditional score notations of Indian classical music are extremely simplified in relation to the sounding music. The relevance of score notation as input for analyses of structure of pieces within this repertoire hence can be seriously questioned, since those notations cannot, in any respect be compared to a transcription. Still, I have found it valuable to use as a test material of a number of reasons; First, the notations represent a culturally acknowledged meta-representation of the music. Thus it can be claimed that the most pertinent within-cultural features of the piece are included in the notation, in order to distinguish it from other pieces within the same repertoire. Another reason is that traditional Indian notation can be compared to Western score notation, which in its most simple form includes only information of relative duration and relative pitch. However, the fundamental problem of whether the notation contains the perceptually necessary information for a listener to conceive the melodic structure remains.

The most interesting features of notation of classical South Indian music in the current context is that it allows notation of primary grouping and involves notation of segments at phrase and section levels.

The analysis of primary grouping is in this method of analysis essentially a part of the metrical analysis, which makes use of the analysis of grouping to extract the metric structure at beat level and to some extent also at measure levels. Hence the primary grouping structure – at central pulse/tactus level - in the case of south Indian classical music is generally input to the phrase and section analysis.

Since primary grouping can be very varied with regards to patterns of group size, one might question the level of metricity of this grouping; can meter, as defined by periodicity, be *aperiodic*? The current model assumes that the accentuation pattern, achieved by primary grouping at beat level to be part of the metrical domain, allowing quasi-periodicity if the group sizes are discrete and exclusive (see Chapter 5, section 5.2.4.7 *Asymmetrical meter at primary grouping level*) and there is true periodicity at beat atom level. Many pieces of South Indian classical music both support and challenge this assumption, since repeated patterns of regular grouping in the beginning of the pieces, create the expectation of a repeated metrical grid, which later is challenged and distorted by more and more *aperiodic* primary grouping towards the end of the pieces. What remains is periodicity at measure and beat atom level (*Tala* period and beat atom), but weaker indication of periodicity/meter at central pulse level, thus emphasizing the gestalt rhythm features of this level.

#### Example Raag Suddha Danyasi

Once again will we return to the example piece also used in the description of the metrical analysis, *Raag Suddha Danyasi* after a notation by the South Indian violinist K. Shivakumar.

Chapter 6

Though the primary grouping analysis by the current model differs from the grouping notation made by the performer (see Chapter 5, section 5.3.4.3, *Discussion trial 6*), it is here used as input for the analysis of phrase and section structure.

**Raag Suddha Danyasi**  
original grouping, with repetitions                      after K. Shivakumar

The image displays a musical score for Raag Suddha Danyasi, transcribed into common Western notation. The score is organized into two columns. The left column, labeled 'original grouping, with repetitions', contains six staves of music. The right column, labeled 'after K. Shivakumar', contains ten staves of music. The score is divided into sections: Pallavi (measures 1-10), Anupallavi (measures 11-16), and Chittaswaram (measures 17-34). The notation includes treble clefs, a key signature of one flat, and a 4/4 time signature. The music consists of a series of eighth and sixteenth notes, with some rests and ties. The original notation uses vertical lines to indicate groupings and repetitions, while the Shivakumar notation uses a more standard Western notation style.

Figure 6-125. Raag Suddha Danyasi (after Shivakumar, K. 1992). Original score notation, transcribed into common Western notation. (first sections)

The output analysis from the computer implementation of the current model manages to identify the central phrase level corresponding to the measure level and the super-phrase/basic section level corresponding to the two-measure period. It also manages to identify the beginning of the *Charanam* section, since it is indicated by a “rondo-like” repetition of a two-measure super-phrase. But the sections *Anupallavi* and *Chittaswaram* are not identified by the model, since they lack identifiable structural integrity. Moreover, towards the end, certain overlapping implications make the structural analysis ambiguous.

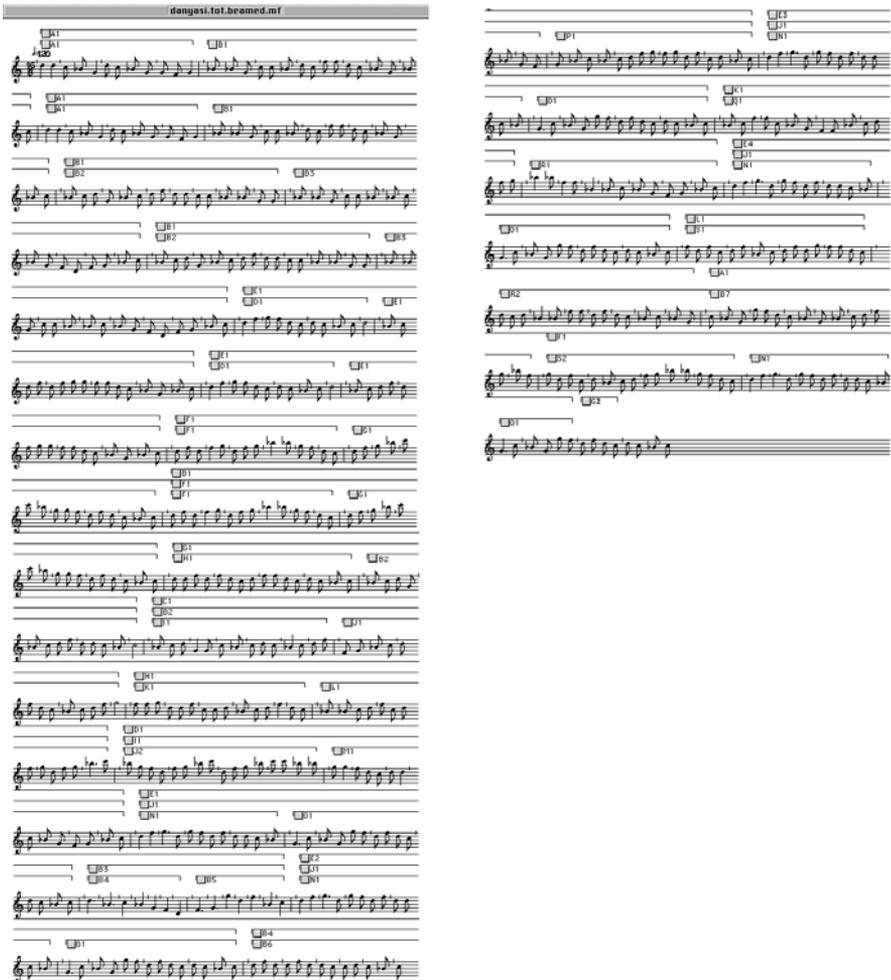


Figure 6-126. Computer model output of analysis of Raag Suddha Danyasi. Start-oriented interpret. (NB. Only higher phrase levels are displayed)

The differences in primary grouping from the original interpretation obviously influence the identification of phrase families and variants. The central structural levels, however, are identified by the analysis performed by the current model.

What is most striking about the output of this analysis in comparison with many of the previous examples is the low level of repetition and categorical similarity between non-adjacent segments on lower segmental levels. The number of phrases determined by sequence similarity is actually very low, considering the entirety of the melody (see below). Or put another way, the actual melodic contour is subject to variation all throughout the piece. Fig.

127 displays the output from the analysis with actual repeats omitted, which makes it possible to display also the sub-measure phrase level.



Figure 6-127. Output from computer model analysis of Raag Suddha Danyasi. Since repeats here are omitted, higher phrase levels have different phrase ID, than in fig. 127 above.

The great majority of the segments in this analysis are obtained by local phrase-boundary analysis (see section 6.2.5), based on discontinuity and symmetrical relationships. The number of sequences is very limited considering the length of the piece and, in relation to previous examples, are mostly based on rhythmic similarity (T-sequences):

- (4 7 T 3)
- (11 7 T 6)
- (33 2 T 3)
- (53 3 T 1)
- (83 22 NIL 5 9)
- (91 7 T 6)
- (124 19 NIL 5 10)
- (132 2 T 2)
- (155 7 T 6)

The variation of melodic contour further creates a great number of segment categories, as shown by the many different phrase IDs. The overlapping similarity between phrase segments makes categories defined by general similarity indiscrete, which causes the model to identify them as belonging to different categories, despite the limited melodic material. A great deal of the segments shows general similarity, but almost no segments can be categorized into



## *Chapter 6*

The incorrect analysis at section level, which divides the third general section into three parts and makes an inconsistent division of the final section is due to deficiencies in the computer implementation of the method. However, in spite of these errors, the evaluation shows that the most central elements of the general structure in the piece, according to the culturally informed score notation, can be obtained by the current method.

## 6.3.7 Evaluation by listener tests

### 6.3.7.1 Purpose of the tests

The purpose of the tests referred to in the following was mainly to get an *indication* of to which extent the current model can predict /account for the conceptions of phrase structure in melodies made by the test group. In other words it was not intended to be a thorough test of the conception of phrase structure with any claim of generality regarding *dominant* conceptions, since this would require a study in its own, beyond the scope of the current study.

The claim is rather that **the model is indicated to be cognitively valid if it can account for the main variability in responses for the test group**. The relevance is obviously limited to the cultural, social and individual limitations of the test group, but the results of some of the other tests, which contained a considerably higher number of test persons, (see Chapter 5, section 5.3.4.3) have indicated that the variability of answers in a larger group in relation to the variability of answers in a smaller group is not directly proportional to the size of the group. Thus the indication of relevance can be assumed to be valid even for larger groups given the same cultural and social background of the test group and given certain variability within the test results.

One of the serious problems regarding the interpretation of the test results is evidently the limited distribution of cultural affiliation/background among the test persons. To my knowledge no cross-cultural tests of conception of phrase structure have been reported in the literature. Such tests would obviously be very interesting in the current context. The cross-cultural aspect is, however, partly involved in the current experiment by the use of test melodies of different cultural origin. One can obtain some indication of the influence of familiarity of a musical style on the results by comparisons of results from (1) the test melodies which were culturally and stylistically unfamiliar to the entire test group, (2) the melodies that were stylistically familiar to only some members of the test group and (3) melodies that were familiar to the entire group.

Because of the limited claims of the test result, the composition of the test group has been more or less arbitrary. The test group consisted of 18 persons, though only 12 persons completed the entire test in all its parts. There were equal numbers of men and women and ages ranged from 18 to 72 years.

Due to the experimental task (relating to score notation) knowledge of common Western notation was required, which limited the possible choice of test persons. Thirteen of the participants were music students at the Royal College of Music in Stockholm, a majority of those studying folk music, the rest studying jazz, classical music and popular music. The other participants were music amateurs and professional musicians/composers.

### 6.3.7.2 General methodology

As has been mentioned in the section 6.3.2, *listener tests*, two different methodologies have been employed for testing conception of phrase structure in melodies: notation of phrase structure in scores by subjects and choice between different interpretations of a melody representing different structural conceptions.

### Experiment 3

Most of the tests presented in the following section use the score notation method. The stimuli in these tests were deadpan performances of the melodic excerpts created in Finale 2000/2002 software on Macintosh computer, transformed into audio files fixating the sound source to be the “Grand piano” sound patch of the synthesizer module Roland JV-1010. The *velocity* parameter was set to a moderate level (44) in order to facilitate grouping conception. The majority of the tests were performed in a computer classroom, where the participants listened to the test melodies on headphones plugged into the computer during two separate sessions for a maximum of 3 hours. However, a few of the test persons (3 persons) listened to the melodic excerpts via a tape recording identical to the one used in the classroom trial.

The instructions were as follows:

1. Listen to the recording without looking in the score
2. Listen to the recording and try to follow the music through the score
3. Mark out in any way you find suitable (rings, brackets, beams etc.):
  - pulse
  - time / measures
  - phrases

The subjects were allowed to listen to the stimuli several times, but were instructed to notate the first structural conception to come in their minds and move on, even if they were not sure that this was the best or final solution. The decision to allow this arbitrariness regarding exposition for stimuli was based on the objective of the study, which was to measure people’s cognition of phrase structure, and the fundamental assumption (see Chapter 3) that cognition of phrase structure is a learning process, which may take different time for different listeners. Scores were designed not to indicate any anticipated test results by visual cues. All notes were therefore equally spaced in the scores where both meter and phrase

Test ex 02

*poco* puls, takt, fras och grundton

GRUNDTON: (mest avslutande ton) C igen.  
 (Vet inte om man styrs av fjärra tonen, men en G-bordlun värdes definitivt som en levnt.)

Figure 6-128. Example of test score and subjects notation of result (subject 1) conforming to the anticipated dominant interpretation of phrase structure

structure were to be noted and in the tests, where the basic beat structure was given, beats were equally spaced. The breaks between systems were also made at points not coinciding with an anticipated phrase or beat structure.

The general reaction of the subjects to the sound stimuli was that it was unmusical, machine-like and some of the subjects noted a general tendency that all sound examples sounded more or less like 4- or 2-grouped because of the square impression of the sound stimuli, manifested in comments such as “everything sound as 4/4” (subject k). In future studies, such reactions are worth taking into account, the use of alternative phrasings may prove a better methodology.

## **Experiment 2**

This experiment involved the use of phrase and metrical alternatives and has been described in the Chapter 5, section 5.3.4.3 *Procedure (trials 8 and 9)*.

The results are interpreted below from the point of phrase and section analysis.

### **6.3.7.3 Results**

#### **Test 3a-b Relationship between conceived pulse and cognition of phrase structure**

The object of this test was to investigate to what extent perceived central pulse level and metrical structure influenced the recognition of sequence similarity as a factor of phrase structure. The experimental hypothesis was that when basic metrical structure (central beat level) is recognized by listeners, indications of phrase structure that interfere with metrical structure will be disregarded by subjects *as a basis for phrase cognition - segmentation*, while such indications may be forceful factors in the conception of melodic structure when conforming to the basic pulse perception.

Based on the experimental hypothesis two different variations of the same melodic structure were prepared, both designed to indicate a triple division of beats at central pulse level, 3/8 pulse within a 9/8 time (alt. 3/4 and triplets). To achieve this, sub-sequencing at primary level was used in connection with pitch discontinuity and continuity concurring with the indicated metrical structure. The duration of each eighth note was set to 250 ms, (240 MM), in order to enhance pulse conception based on grouping of individual notes.

In the first part of both melody variants (see fig. 132 below), the melody was designed to indicate a perfect symmetrical division of the four 3-beat measures into two equivalent phrases. In the second part of both test melodies, a repeated sequence was inserted, which in the first test melody did not concur with the anticipated beat conception.

The second test melody was constructed so that the inserted sequence would concur with the implied beat structure, but on the other hand imply an asymmetric division of the second part into two (or three) segments. This asymmetric division did not concur with the phrase length implied in the initial part of the test melody, thus implying a change of phrase length and possibly meter at measure level. (see also discussion in section 6.1.4)

Chapter 6

Experiment 3a (stimulus version a)

Experiment 3b (stimulus version b)

Figure 6-129. Melody examples 3a and b, with the implied pulse and metrical conception noted. The repeated sequence noted by brackets. Implied primary grouping (beat-grouping) and phrase structure in the initial part is indicated by slurs.

The results show alternative solutions, with respect to meter for both melodic stimuli, that all conform to the central pulse predicted by the experimental hypothesis. The measure level meter was interpreted by some of the subjects differently from the expected interpretation, with a preference for grouping the beats in two and four instead of three beats per measure; the former interpretations actually were slightly more frequent than the expected grouping in three beats per measure. The main prediction was confirmed by the results, where none of the responses exhibited a phrase segmentation complying with the interfering sequence in the first version (a). Instead the symmetrical division of the first part was continued in the second part for the main part of the subjects. (see fig. 130)

Figure 6-130. Different interpretations of phrase structure and meter for melody 3a. Metrical interpretation showed by brackets below system and phrase interpretation by brackets above system.

Figure 6-131. . Different interpretations of phrase structure and meter for melody 3b. Metrical interpretation showed by brackets below system and phrase interpretation by brackets above system.

Table 39. . Results exp. 3a

<i>Metrical interpretation</i>	<i>Frequency</i>
pulse 3/8	18
measure a	7
measure b	6
measure c	3
measure = pulse 3/8	2
<i>Phrase interpretation</i>	<i>Frequency</i>
phrase a1 (18/8)	9
phrase a2 (9/8)	5 (3 also within a1)
phrase b	2
phrase a1+b (first a, second b)	3
phrase c1-c2	2

The interpretations of the second test melody demonstrated the predicted difference regarding the recognition of the sequence repetition as a basis for segmentation. (Table 40)

Table 40. Results exp. 3b

Metrical interpretation	Frequency
pulse 3/8	18
measure a+a2	5
measure a+d2	2
measure b	5
measure c+c2	2
measure c+d2	1
measure d	1
measure = pulse 3/8	2
Phrase interpretation	Frequency
phrase a1 (18/8 etc.)	9
phrase a2 (9/8 etc.)	4 (all but 1 within a1)
phrase b	3
phrase c-c1	4
phrase c-c2	1

Since the segmentation task regarding the second melody followed immediately after the first melody was segmented and the beginnings were identical, it is not surprising to find the same metrical interpretation of the first part being withheld for the majority of the group. However, the metrical interpretation changed markedly in the second part of the melody, where the beat-grouping at measure level was for the most part of the test group changed in compliance with the noted phrase structure. Obviously, the repeated interfering sequence was recognized and, since it was in compliance with the basic metrical structure at beat level, some of the subjects conceived it to be of metrical significance, changing the metrical structure at measure level.

This result can be interpreted as to indicate the validity of the assumption of primacy of metrical structure at beat level, being a fundamental level of organization to which phrase conception relates. (see section 6.1.4, *Assumption no 12*). Further, it indicates the validity of the assumption that repeated sequences are significant for the conception of phrase structure when they are in compliance with the conception of fundamental metrical organization. On the other hand, it cannot be said to account for all the responses of the test group, since some of the responses did not concur with the repeated sequence. The result also indicated the validity of the assumed complementary nature of start- and end-oriented grouping preference, since the phrase interpretation *b* of both test melodies is phase shifted by one event in relation to the *a* interpretation, hence concurring with the larger pitch distances in the melody. (*Assumption no. 6*).

The start- and end-oriented solutions of the model thus accounts for the major tendencies of the results regarding the conceived phrase structure within the test group, even if the metrical interpretation of the first trial melody was not supported by a majority of the subjects.

A) test melody 3a start-oriented interpretation

B) test melody 3b start-oriented interpretation

Figure 6-132. Computer model solution for test melodies 3a and 3b.

It is worth noticing that the computer model makes an erroneous subdivision of the second major sequence A3-A4 in the second test melody into 2 + 3 beats instead of the 3+2

beat subdivision, which is indicated by the prior segment structure. This is due to incomplete implementation of the model.<sup>113</sup>

### Test 3c-d Further test of relationships between metrical structure and phrase structure

The next pair of test melodies was also used in the tests of cognition of metrical structure, related in section 5.3.4.3, *trial 1 and 2*. Here identical pitch material is used in both test melodies, while the pitch order was different in the two test melodies.

The experimental hypothesis, in this case, was also that the conceived phrase structure would be congruent to the conceived metrical structure. The first of the two melodies was constructed to indicate triple division of beats at central pulse level, while the second was constructed to indicate duple division.

#### Test melody 3c

The figure displays three systems of musical notation for Test melody 3c, each with a different metrical interpretation. Above each system, brackets indicate phrase structures assigned by subjects, labeled with letters and numbers in parentheses.

- System 1 (Metrical interpret a):** Shows a melody in 3/8 time. Brackets above the staff indicate phrase structures for subjects m. (1), i. (1), b. (7), and a. (2).
- System 2 (Metrical interpret b):** Shows the same melody in 3/8 time. Brackets above the staff indicate phrase structures for subjects p. (1) (6/8), d. (1), and c. (2).
- System 3 (Metrical interpret c):** Shows the melody with changing time signatures: 3/8, 2/4, 3/4, 3/8, 5/8, 3/8, and 3/8. Brackets above the staff indicate phrase structures for subject h.

Figure 6-133. Result of notation of meter and phrase structure for test melody 3c for subjects a-p. The phrase structures assigned by different subjects are displayed above the system by brackets, labelled by a letter representing a subject who presented that particular solution. The number of subjects giving the same response is displayed by numbers within parentheses<sup>114</sup>. Different metrical interpretations are displayed in different systems, clustering metrical interpretations with identical pulse notation regardless of different time signatures.

<sup>113</sup> The behaviour of the model regarding start- and end-oriented interpretation will be discussed in relation to other tests.

<sup>114</sup> Minor differences between phrase indications, when every indicated segment concurs, are disregarded in order to simplify the presentation of results.

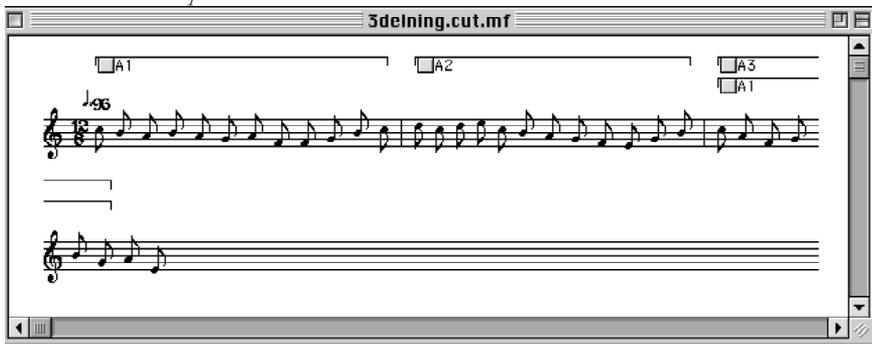
Since there are actually a number of rather odd and complex solutions present in the results both in this and a number of the following tests, - for instance metrical interpretation c above, - in what can seem like very straightforward and simple examples, it can be worth considering the reasons for this.

As have been observed in other similar tests (See Höthker et al 2002), there is considerable variability of responses even in a relatively small group of test persons. According to the comments made by the participants in these tests this variability relates to the machine-interpreted stimulus where all articulation and tempo variations are equalized, which makes the structure confusing. Some of the subjects actually express that they feel musically offended by the lack of interpretation, which is interesting considering the high average skill in reading music within the test group. This could be interpreted as to demonstrate, that the information given in these tests is not sufficient to invoke a conception of phrase and metrical structure; But since responses in general are not arbitrary, I am rather interpreting the reaction as to demonstrate that a performance of a melody – without any deviation from mechanical timing and variation in articulation – is an extreme and unnatural musical situation, which is also demonstrated in a number of studies (e.g. Bengtsson and Gabrielsson 1980, 1983). This assumption is also corroborated by the results of the experiment 2, (section 5.3.4.3, *trials 1-2*) regarding metrical structure, presented in section, giving more consistent results when stimuli involved choice between different metrical interpretations.

The results for test melody 3c show a preference for the phrase structure assigned to subject b in metrical interpretation a, which is identical to the phrase structure assigned to subject p in metrical interpretation b. The second most dominant solution is represented by the phrase structure assigned to subject a, which is partly indicated also in the response of subject i.

These two phrase structures are in compliance with the start- and end-oriented interpretations of the computer implementation of the current model:

*A. Start-oriented interpretation*



B. End-oriented interpretation

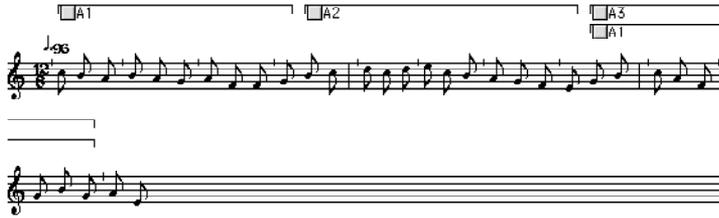


Figure 6-134. Computer model solution for test melody 3c<sup>115</sup>. A.start-oriented and B end-oriented interpretation

In summary, 70.2 % of the segments noted by the test group are accounted for by the current model, and 80.1 % of the individual segment boundaries. These figures are based on counting every subject's segmentation, even if it is a duplicate of a segmentation made by another subject. Hence it is a weighted ratio based on the total responses given by the group<sup>116</sup>, where more common responses have greater influence on the ratio than the more odd ones. The start-oriented interpretation seems here to be favored by the test group in relation to the end-oriented interpretation.

Test melody 3d



Figure 6-135. Result of notation of meter and phrase structure for test melody 3d for subjects a-p. (For explanation see fig. 133, test melody 3c)

<sup>115</sup> The ending of the last phrases should be placed after the last note of the melody, but due to deficiencies in the graphical interface phrase endings are often placed too early.

<sup>116</sup> All ratios/percentages given within this section are based on the same calculation procedure

The results also in this case indicate the correlation between conceived metrical structure and phrase structure, since practically all phrase boundaries are being placed on conceived beats by the subjects. Analogically, the metrical interpretation at measure level can be interpreted as being related to the conceived phrase structure.

Figure 6-136. Computer model solutions for test melody 3d. Start- and end-oriented interpretation

In this case the models solutions (identical for start- and end-oriented versions) accounted for 72.1 % of the segments and 84.7 % of the individual segment boundaries.

The tests of this pair of melodies evidently indicated the above mentioned relationship between cognition of phrase structure and metrical structure, thus corroborating the assumption of the primacy of metrical structure at beat level (see section 6.1.4, *Assumption no 12*). It further supported the assumption of the inter-relationship between metrical structure at measure level and phrase conception.

The model managed to account for the subjects responses better than chance for both test melodies.

### Test melody 3e – Phrase conception in relation to asymmetrical beat pattern 1: Balkan additive beat-grouping

This test melody was inspired by typical metrical structure in Balkan and Turkish dance music, e.g. *Bulgarian Dajcevo* or Turkish *Karsilama dances*, where the meter at division level (beat atom level) is periodically grouped into 2+2+2+3 beat atoms, resulting in an asymmetrical meter at central pulse level. (cf. Section 5.3.4.3, *trial 3*)

To indicate this metrical grouping repetitions of a pitch pattern was used (cf. *Rondo a la Turk* 5.2.4.7 *Further successive evaluation*), on both implied beat and measure level, which in combination with a relatively high tempo (250 ms per eighth note)

test ex 03  
puls, gruppering, takt, fras



Figure 6-137. Provided score for notation of phrase and metrical structure in test melody 3e

The test results demonstrates that the implied metrical structure was recognized by a majority of the subjects.<sup>117</sup>



Figure 6-138. Results of the notation of metrical structure and phrase structure of test melody 3d. NB Seven out of nine subjects assigned to the interpretation a. did not note the sub-phrase segmentation. The phrase designation provided only by subject f. is noted within the brackets representing the segments. One of the subjects assigned to f. did note consistent 9/8.

This result once again demonstrates the connection between perceived metrical structure and phrase structure. Subjects l and e, who differed from the rest regarding the segment boundaries, have also noted a different metrical structure. The difference regarding metrical conception at central pulse level between interpretations a and c, on the one hand, and f, on the other did however not result in a different phrase interpretation, since the phrase boundaries coincide with both the metrical interpretations.

<sup>117</sup> The similarity between this melodic example and e.g. the Rondo a la Turk theme may have influenced this result, which another version of the test melody was used in the experiment concerning metrical structure

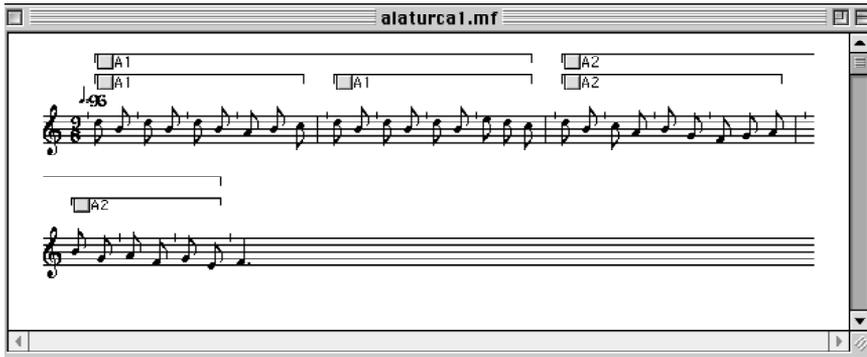


Figure 6-139. Computer model solution for test melody 3e. Start- and end-oriented interpretation identical. (The last ending phrase bracket is also here placed at the wrong graphical position).

The model does also, in this case manage to account for the majority of the responses given by the test group: 71.5% of the segments were accounted for by the model result and 84.3 % of the phrase boundaries.

### Test melody 3f - Phrase conception in relation to implied asymmetrical beat pattern 2: Swedish polska

The next test melody also concerned the conception of phrase structure in relation to asymmetrical meter. This melody was a section of the Swedish *polska* tune, which was also used in the experiment regarding metrical structure, *Polska* played by Gössa Anders Andersson (see Chapter 5, 5.2.2.3)

This can be interpreted as having an identical beat asymmetry at central pulse level as the quasi-balkan melody 3e. In this case, however, there is not an apparent meter at division level (beat atom level) implied by the melodic structure or in the traditional performance style. It was played at the same tempo as the previous example melody, 250 ms per eighth note.

One interesting aspect of the current test melody was that the melody style (and in some cases the actual melody) was culturally familiar to about half of the test group (folk musicians) while it could be expected to be more foreign to the other half (with classical, jazz, popular music background). This different background could be expected to show in the results.



## A. Start-oriented interpretation

## B. End-oriented interpretation

Figure 141. Computer model solutions for test melody 3f. A start-oriented and B end-oriented interpretation.

The start- and end-oriented interpretations given by the computer model could account for the majority of the responses, regardless of different notated metrical structure: 86.0 % of the segments were identified by the model and 97.9 % of the phrase boundaries.

It is interesting to see that, contrary to in some of the previous examples (e.g. test melody 3c), the end-oriented interpretation was most popular. The start-oriented interpretation was only proposed by test participants with a background in Swedish folk music.

The majority of the responses also had a corresponding phrase and metrical structure. Some of the responses, however, could be interpreted as if an ambiguous beat structure was derived from a more evident conception of phrase structure.

### Test melody 3g – Phrase conception in relation to asymmetrical beat pattern 3: “Kjempe Jo”

The *gangar* melody “Kjempe Jo” played by Norwegian fiddler Salve Austenå was also used in the tests regarding conception of metrical structure (see section 5.3.4.3, *trial 5*).

Its complex and ambiguous metrical structure is interesting in relation to conception of phrase structure. As have been noted earlier (see section 6.3.4), the phrase structure in the *gangar* melodies can be conceived and have been described as intertwined, phrase endings becoming starts of succeeding segments etc.

This ambiguity was also indicated by the responses by the test group. A few of the subjects actually demonstrated this by making consistently overlapping phrase boundaries.

The melody was, to my knowledge unknown, to all of the subjects, none of them being a specialist in Norwegian fiddle music. However, the Swedish folk musicians were considerably more acquainted with the style than the other subjects.

Here, I used a somewhat distorted version of the melody, to obtain an even higher degree of ambiguity than in the original version used earlier. The tempo was, also in this case, chosen to stimulate a complex pulse conception, 272 ms per eighth note (MM 220).

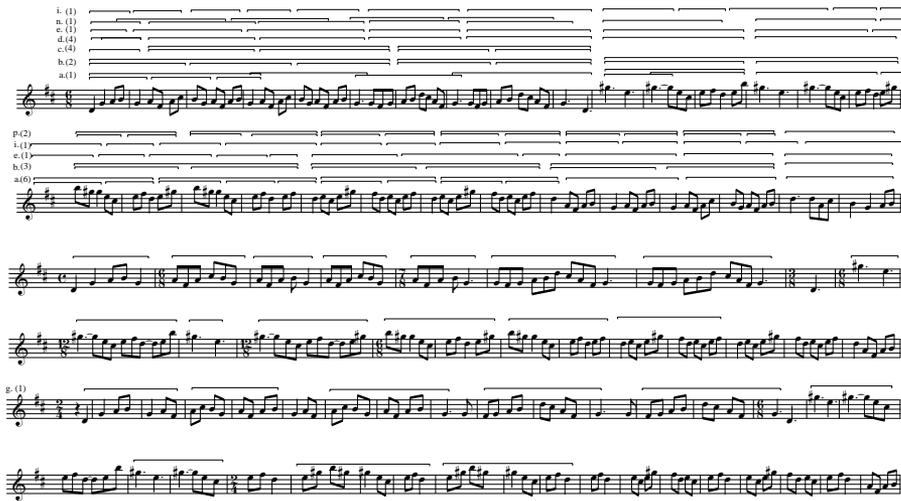


Figure 142. Results of notation of metrical structure and phrase structure in test melody 3g.

As can be seen in figure 142, some of the subjects demonstrated the intertwined nature of the phrase structure through consistently overlapping phrase boundaries, as e.g. subject a.

These different interpretations can be summarized as in the following figure leaving out both the differences regarding metrical interpretation and the differences regarding overlapping phrase boundaries which follow the more consistent segmentations (fig. 143).



Figure 143. Summarized result for test melody 3g.

Thus, the result can be interpreted from the view of start- and end-oriented grouping preference. In the beginning of the melody start-oriented grouping 1 uses the first longest sequence repetition starting on the note g in the second measure, while start-oriented grouping 2 disregards this sequence start assuming the sequence to be dividing the section into three parts of equal length.

The “end-oriented” interpretation in section A regards the quarter note in the beginning of the second measure as an ending, implying a start on the succeeding eighth note, hence determining the phase of the sequence. This grouping preference may thus rather be regarded start-oriented grouping 3 in the A section, since there is no upbeat or anacrusis to the starting point<sup>118</sup>. On the other hand, in the section designated B above, the difference between start- and end-oriented typically involves the inclusion of an upbeat in the end-oriented grouping. The preference for start and end-oriented interpretation changes within the melody for the test group. In section A end-oriented interpretation is strongly favored, in the B section slightly favored, while in the C section the start-oriented interpretation is dominant. The tendency can be summarized in the following way:

*End-oriented grouping is favored when there are consistent differences between point of greatest impulse (start) and point of greatest discontinuity (end), while start-oriented grouping is favored in this test group when these points coincide consistently within the melodic structure.*

The two different interpretations as performed by the current model can account for a majority of the responses given by subjects: 86.6 % of the segments were accounted for by the model and 88.2 % of the segment boundaries. The model does not account for e.g. the start-oriented grouping 2 in the summarized result above and not for the more odd results given by some individual subjects. The end-oriented interpretation given by the model also makes a division of A segments at the lowest level. This interferes with the beat structure and is not recognized by a single subject (A1-A2 phrase boundary). This division is obviously an over-interpretation of distance based grouping which should be disregarded. But, in the majority of the responses, the two interpretations seem to incorporate most of the variability. However, in this case and similar once, it might be suggested that the model could be developed to combine the start- and end-oriented into the intertwined interpretation suggested by some of the subjects, with overlapping segment starts and endings

<sup>118</sup> The metrical interpretation in fig. 144 does imply exactly this interpretation, by the assignment of measure start at this point.

A. Start-oriented interpretation

kjempejo.fin98.mf

♩<sup>110</sup>

Chord diagrams shown in the score:

- A1
- A2
- A3
- C1
- C2
- B1
- E1
- B2
- B3
- A2
- A1
- A3
- F1
- D1
- B1
- B2
- B3
- A1
- A2
- A4
- B1
- H1

B. End-oriented interpretation

**kjempejo.fin98.mf**

Figure 144. Computer model solution for test melody 3g. A start-oriented and B end-oriented interpretation.

### Test melody 3h – Phrase structure in relation to implied asymmetrical beat pattern 4: Carnatic Raga melody

The fourth test of phrase conception in a melodic structure which can imply an asymmetrical beat pattern concerned the south Indian raga melody “*Raag Suddha Danyasi*” (see also section 6.3.6.1).

The interesting objective using this melody was to use a material which was culturally foreign to all subjects.

The results of experiment 2, trial 6, regarding metrical cognition, indicated that listeners were sensitive to the melodic cues of asymmetrical beat-grouping at central pulse level, which is contained in the structure of this melody and recognized in culturally informed notations.

However, considering the method of the current experiment, which did not imply any alternative accent patterns in the stimulus, one would expect the subjects to favor metric patterns from their own cultural experience. Comparing the results of the two tests, this expectation was confirmed, though not to the level that was expected.

(Varying time signatures 2/4, 4/8, 4/4, but with generally the same beat grouping as below)

c. (1)  
a. (8)

(Varying time signatures: 16/8, 8/8, 4/4, but with generally the same grouping as below)

b. (4)

d. (1)  
e. (1)

f. (1)

Figure 145. Results of the notation of metrical structure and phrase structure of test melody 3h. Tempo: 250 ms/ eighth note

The first of the metrical interpretations, given by a majority of the subjects, can be said to conform to the main cultural experience of the subjects, since it implies perfect periodicity at all metrical levels. The other subject interpretations (6 out of 15) did actually recognize the culturally relatively foreign structural cues within the melody resulting in an asymmetric beat-grouping at central pulse level. This may be influenced by prior experience of this type of

music but definitely not for all of the subjects who chose this structure. Another explanation is that asymmetrical beat patterns of this kind (e.g. 3+3+2 grouping) do exist in musical styles known to the Western listener and the melodic cues make subjects relate to this experience.

Regarding the identification of phrase structure the responses were very concordant and in very much in accordance with the main phrase structure indicated by the original notation of the melody.

The solution of the current model did also, in this case, account for the dominant alternative segmentations given by subjects (see fig. 126, section 6.3.6.1)

The result for this test melody was very good, the model being able to account for 92.8 % of the segments and all segment boundaries (100 %) given by subjects.

The repetition of segments and the structural discontinuity cues were obviously clear here evident to a degree to which the models' prediction and subjects agreed to a high level. It is, however, worth noting that neither the model nor the listeners agreed to such a high level on the culturally informed notation of beat-groups, while the phrase segments in the notation coincided totally with the listeners judgment.

### **Test melody 3i - Ravel's Boléro: Asymmetrical segment structure at phrase level interfering with metrical structure**

The piece "*Boléro*" by M. Ravel belongs to the common Western repertoire well known to the members of this test group, independent of their stylistic background. This piece is frequently used in different popular music and "light classical" settings, as film music and is frequently performed in concerts. Even if it is originally an orchestral piece, its structure is determined mainly by melodic line, which makes it possible to use the melody without the accompaniment.

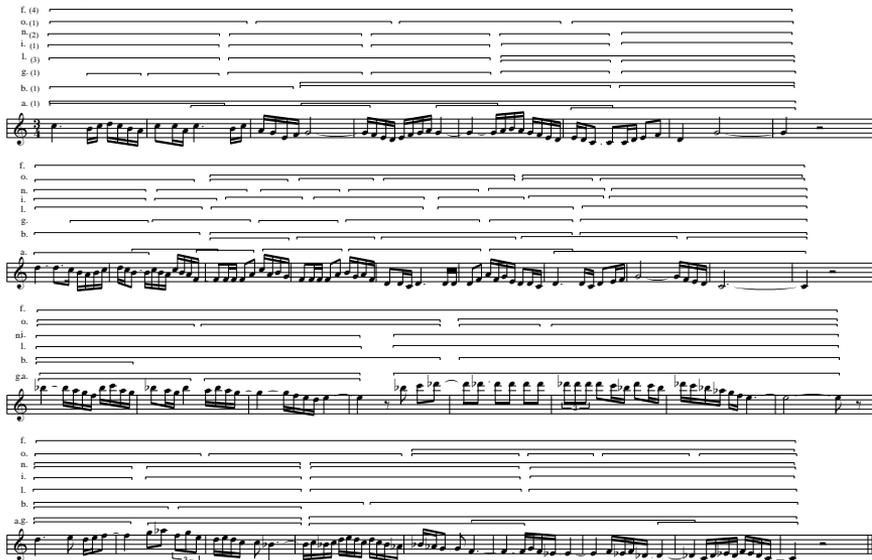


Figure 146. Results of the notation of phrase structure of test melody 3i. Tempo: 625 ms/quarter note.

Chapter 6

The interesting question regarding this melody was whether the test group, who obviously knew the melody and probably would be able even to imagine the accompaniment, would have a clear and concurrent conception of the melody structure?

The beat structure was given by the notation, but the time-signature and bar lines were omitted. The beat structure was assumed to be known to the subjects from earlier experience of the piece. The variability regarding phrase notation was quite remarkable, considering the popularity of the melody. On the section level, almost all subjects agreed about the segment boundaries. But on phrase level, there were almost no identical solutions among the subjects. In order to get an overview over the main traits in the subjects responses, the individual results are summarized below in fig. 147. The super-imposed boxes above the systems refer to the number of subjects appointing the corresponding note to be a segment start, each box representing an individual subject. Shaded boxes refer to notated sections, involving phrases at lower levels, while lower level phrase starts are indicated by transparent boxes.<sup>119</sup>



Figure 147. Summarized result of notation of phrase structure in test melody 3i. (For explanation of figure see above)

As can be seen, the results can be summarized into four different interpretations, two start-oriented and two end-oriented. The main end-oriented interpretation was preferred by the subjects. All four interpretations can be traced to features of the melodic structure, the interpretations thus are related to structural cues and not arbitrary.

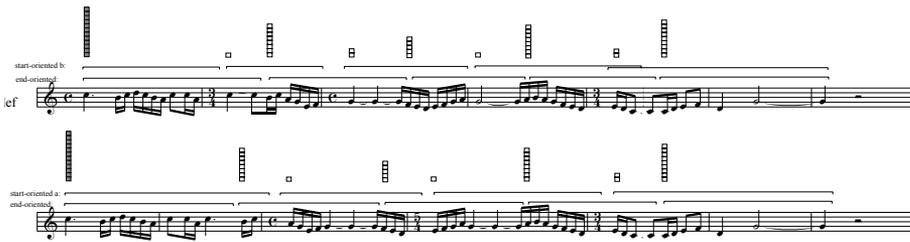
Why are these results so confusing on a test melody that is well-known to the subjects?

One feature of the structure, which might enhance the ambiguity is the lack of periodicity regarding similar structures - similar local structural changes do not occur periodically throughout the melody.

<sup>119</sup> This display was adapted from Höthker et al 2002

Consider the rhythmic structure of the very beginning, characterized by the start at a note of relatively long duration symbolized by a dotted quarter note followed by shorter notes. This rhythmic gestalt appears at beat 0, beat 4, beat 7 and beat 11 within the first section, hence *aperiodically*. Two of the subjects still used this rhythmic gestalt as a cue to phrase structure.

The *aperiodicity*, with regards to the extremely long notes in the first section and the salient segment repetitions in the second section in combination with the frequent suppression of onsets at metrically prominent positions throughout the melody, creates ambiguous structural cues, which may be the reason for the many different interpretations even within the small test group. It is obvious that even the start-oriented solutions do not consistently concur with the metrical structure implied by the accompaniment. If one judges from the dominating responses given by subjects they might imply an asymmetrical metrical structure of the melody at measure level, with alternating 3/4 and 4/4 measures, like in fig. 147.



*Figure 148. Two alternative metrical interpretations of test melody 3i based on notations of phrase structure, start-oriented solutions*

The melodic structure can thus be interpreted as interfering with the consistent metrical structure implied by the accompaniment, enhancing the conception of structural ambiguity.

The current model chooses a start-oriented phrase structure, which basically concurs with the start-oriented segmentation *a* above (see fig. 148), while the correspondent end-oriented segmentation corresponds to the dominant interpretation within the test group: 79.7 % of the segments are identified by the model and 94.2 % of the segment boundaries.

A. Start-oriented interpretation

**bolero.263.new.mf**

The musical score is presented in ten staves. The first staff begins with a tempo marking of ♩ 96. The score is annotated with various labels in boxes, each connected by a horizontal line to a specific measure or group of measures in the music. The labels are: A1 (first staff), B1 (second staff), A2 (second staff), A3 (second staff), D1 (third staff), A2 (fourth staff), C1 (fourth staff), E1 (fourth staff), F1 (fourth staff), A4 (fifth staff), A5 (fifth staff), A3 (sixth staff), A1 (sixth staff), A6 (sixth staff), G1 (seventh staff), H1 (seventh staff), A2 (seventh staff), I2 (eighth staff), A3 (eighth staff), J1 (eighth staff), K1 (ninth staff), L1 (ninth staff), M1 (tenth staff), and N1 (tenth staff).

## B. end-oriented interpretation

**bolero.263.new.mf**

The figure displays a musical score for a piece titled "bolero.263.new.mf". The score is presented in a single system with ten staves. Above the staves, various phrase labels are indicated by brackets and boxes, representing a computer model's analysis of the music. The labels include:

- A1 (multiple instances)
- A2
- A3
- A4
- A5
- A6
- A7
- A8
- A9
- A10
- A11
- A12
- A13
- A14
- A15
- A16
- A17
- A18
- A19
- A20
- A21
- A22
- A23
- A24
- A25
- A26
- A27
- A28
- A29
- A30
- A31
- A32
- A33
- A34
- A35
- A36
- A37
- A38
- A39
- A40
- A41
- A42
- A43
- A44
- A45
- A46
- A47
- A48
- A49
- A50
- A51
- A52
- A53
- A54
- A55
- A56
- A57
- A58
- A59
- A60
- A61
- A62
- A63
- A64
- A65
- A66
- A67
- A68
- A69
- A70
- A71
- A72
- A73
- A74
- A75
- A76
- A77
- A78
- A79
- A80
- A81
- A82
- A83
- A84
- A85
- A86
- A87
- A88
- A89
- A90
- A91
- A92
- A93
- A94
- A95
- A96
- A97
- A98
- A99
- A100
- A101
- A102
- A103
- A104
- A105
- A106
- A107
- A108
- A109
- A110
- A111
- A112
- A113
- A114
- A115
- A116
- A117
- A118
- A119
- A120
- A121
- A122
- A123
- A124
- A125
- A126
- A127
- A128
- A129
- A130
- A131
- A132
- A133
- A134
- A135
- A136
- A137
- A138
- A139
- A140
- A141
- A142
- A143
- A144
- A145
- A146
- A147
- A148
- A149
- A150
- A151
- A152
- A153
- A154
- A155
- A156
- A157
- A158
- A159
- A160
- A161
- A162
- A163
- A164
- A165
- A166
- A167
- A168
- A169
- A170
- A171
- A172
- A173
- A174
- A175
- A176
- A177
- A178
- A179
- A180
- A181
- A182
- A183
- A184
- A185
- A186
- A187
- A188
- A189
- A190
- A191
- A192
- A193
- A194
- A195
- A196
- A197
- A198
- A199
- A200
- A201
- A202
- A203
- A204
- A205
- A206
- A207
- A208
- A209
- A210
- A211
- A212
- A213
- A214
- A215
- A216
- A217
- A218
- A219
- A220
- A221
- A222
- A223
- A224
- A225
- A226
- A227
- A228
- A229
- A230
- A231
- A232
- A233
- A234
- A235
- A236
- A237
- A238
- A239
- A240
- A241
- A242
- A243
- A244
- A245
- A246
- A247
- A248
- A249
- A250
- A251
- A252
- A253
- A254
- A255
- A256
- A257
- A258
- A259
- A260
- A261
- A262
- A263
- A264
- A265
- A266
- A267
- A268
- A269
- A270
- A271
- A272
- A273
- A274
- A275
- A276
- A277
- A278
- A279
- A280
- A281
- A282
- A283
- A284
- A285
- A286
- A287
- A288
- A289
- A290
- A291
- A292
- A293
- A294
- A295
- A296
- A297
- A298
- A299
- A300
- A301
- A302
- A303
- A304
- A305
- A306
- A307
- A308
- A309
- A310
- A311
- A312
- A313
- A314
- A315
- A316
- A317
- A318
- A319
- A320
- A321
- A322
- A323
- A324
- A325
- A326
- A327
- A328
- A329
- A330
- A331
- A332
- A333
- A334
- A335
- A336
- A337
- A338
- A339
- A340
- A341
- A342
- A343
- A344
- A345
- A346
- A347
- A348
- A349
- A350
- A351
- A352
- A353
- A354
- A355
- A356
- A357
- A358
- A359
- A360
- A361
- A362
- A363
- A364
- A365
- A366
- A367
- A368
- A369
- A370
- A371
- A372
- A373
- A374
- A375
- A376
- A377
- A378
- A379
- A380
- A381
- A382
- A383
- A384
- A385
- A386
- A387
- A388
- A389
- A390
- A391
- A392
- A393
- A394
- A395
- A396
- A397
- A398
- A399
- A400
- A401
- A402
- A403
- A404
- A405
- A406
- A407
- A408
- A409
- A410
- A411
- A412
- A413
- A414
- A415
- A416
- A417
- A418
- A419
- A420
- A421
- A422
- A423
- A424
- A425
- A426
- A427
- A428
- A429
- A430
- A431
- A432
- A433
- A434
- A435
- A436
- A437
- A438
- A439
- A440
- A441
- A442
- A443
- A444
- A445
- A446
- A447
- A448
- A449
- A450
- A451
- A452
- A453
- A454
- A455
- A456
- A457
- A458
- A459
- A460
- A461
- A462
- A463
- A464
- A465
- A466
- A467
- A468
- A469
- A470
- A471
- A472
- A473
- A474
- A475
- A476
- A477
- A478
- A479
- A480
- A481
- A482
- A483
- A484
- A485
- A486
- A487
- A488
- A489
- A490
- A491
- A492
- A493
- A494
- A495
- A496
- A497
- A498
- A499
- A500
- A501
- A502
- A503
- A504
- A505
- A506
- A507
- A508
- A509
- A510
- A511
- A512
- A513
- A514
- A515
- A516
- A517
- A518
- A519
- A520
- A521
- A522
- A523
- A524
- A525
- A526
- A527
- A528
- A529
- A530
- A531
- A532
- A533
- A534
- A535
- A536
- A537
- A538
- A539
- A540
- A541
- A542
- A543
- A544
- A545
- A546
- A547
- A548
- A549
- A550
- A551
- A552
- A553
- A554
- A555
- A556
- A557
- A558
- A559
- A560
- A561
- A562
- A563
- A564
- A565
- A566
- A567
- A568
- A569
- A570
- A571
- A572
- A573
- A574
- A575
- A576
- A577
- A578
- A579
- A580
- A581
- A582
- A583
- A584
- A585
- A586
- A587
- A588
- A589
- A590
- A591
- A592
- A593
- A594
- A595
- A596
- A597
- A598
- A599
- A600
- A601
- A602
- A603
- A604
- A605
- A606
- A607
- A608
- A609
- A610
- A611
- A612
- A613
- A614
- A615
- A616
- A617
- A618
- A619
- A620
- A621
- A622
- A623
- A624
- A625
- A626
- A627
- A628
- A629
- A630
- A631
- A632
- A633
- A634
- A635
- A636
- A637
- A638
- A639
- A640
- A641
- A642
- A643
- A644
- A645
- A646
- A647
- A648
- A649
- A650
- A651
- A652
- A653
- A654
- A655
- A656
- A657
- A658
- A659
- A660
- A661
- A662
- A663
- A664
- A665
- A666
- A667
- A668
- A669
- A670
- A671
- A672
- A673
- A674
- A675
- A676
- A677
- A678
- A679
- A680
- A681
- A682
- A683
- A684
- A685
- A686
- A687
- A688
- A689
- A690
- A691
- A692
- A693
- A694
- A695
- A696
- A697
- A698
- A699
- A700
- A701
- A702
- A703
- A704
- A705
- A706
- A707
- A708
- A709
- A710
- A711
- A712
- A713
- A714
- A715
- A716
- A717
- A718
- A719
- A720
- A721
- A722
- A723
- A724
- A725
- A726
- A727
- A728
- A729
- A730
- A731
- A732
- A733
- A734
- A735
- A736
- A737
- A738
- A739
- A740
- A741
- A742
- A743
- A744
- A745
- A746
- A747
- A748
- A749
- A750
- A751
- A752
- A753
- A754
- A755
- A756
- A757
- A758
- A759
- A760
- A761
- A762
- A763
- A764
- A765
- A766
- A767
- A768
- A769
- A770
- A771
- A772
- A773
- A774
- A775
- A776
- A777
- A778
- A779
- A780
- A781
- A782
- A783
- A784
- A785
- A786
- A787
- A788
- A789
- A790
- A791
- A792
- A793
- A794
- A795
- A796
- A797
- A798
- A799
- A800
- A801
- A802
- A803
- A804
- A805
- A806
- A807
- A808
- A809
- A810
- A811
- A812
- A813
- A814
- A815
- A816
- A817
- A818
- A819
- A820
- A821
- A822
- A823
- A824
- A825
- A826
- A827
- A828
- A829
- A830
- A831
- A832
- A833
- A834
- A835
- A836
- A837
- A838
- A839
- A840
- A841
- A842
- A843
- A844
- A845
- A846
- A847
- A848
- A849
- A850
- A851
- A852
- A853
- A854
- A855
- A856
- A857
- A858
- A859
- A860
- A861
- A862
- A863
- A864
- A865
- A866
- A867
- A868
- A869
- A870
- A871
- A872
- A873
- A874
- A875
- A876
- A877
- A878
- A879
- A880
- A881
- A882
- A883
- A884
- A885
- A886
- A887
- A888
- A889
- A890
- A891
- A892
- A893
- A894
- A895
- A896
- A897
- A898
- A899
- A900
- A901
- A902
- A903
- A904
- A905
- A906
- A907
- A908
- A909
- A910
- A911
- A912
- A913
- A914
- A915
- A916
- A917
- A918
- A919
- A920
- A921
- A922
- A923
- A924
- A925
- A926
- A927
- A928
- A929
- A930
- A931
- A932
- A933
- A934
- A935
- A936
- A937
- A938
- A939
- A940
- A941
- A942
- A943
- A944
- A945
- A946
- A947
- A948
- A949
- A950
- A951
- A952
- A953
- A954
- A955
- A956
- A957
- A958
- A959
- A960
- A961
- A962
- A963
- A964
- A965
- A966
- A967
- A968
- A969
- A970
- A971
- A972
- A973
- A974
- A975
- A976
- A977
- A978
- A979
- A980
- A981
- A982
- A983
- A984
- A985
- A986
- A987
- A988
- A989
- A990
- A991
- A992
- A993
- A994
- A995
- A996
- A997
- A998
- A999
- A1000

Figure 149. Computer model solutions for test melody 3i. NB The lowest phrase level is not displayed in the output due to the limitation of three simultaneous levels. A start-oriented and B end-oriented solution.

**Test melody 3j – Fascinatin’ rhythm: Asymmetrical segment structure at phrase level interfering with metrical structure**

As was demonstrated in section 6.3.5.1, the song “Fascinatin’ rhythm” by Gershwin exhibits a phrase structure which is clearly interferes with the general metrical structure proved by the accompaniment of the song. The interfering phrase structure is not only cued by the melody structure, but also underlined by the lyrics of the song.

The experimental hypothesis was that the asymmetrical phrase structure would show in the responses of the test group, since the members of the test group were assumed to be acquainted with the song.

The test result corroborated the hypothesis giving a rather unambiguous indication of the melody structure in comparison with the previous test melody.



*Figure 150. Results of notation of phrase structure of test melody 3j for subjects a-p. Tempo: 469 ms/ dotted eighth note (NB Four of the subjects in the group did not make this notation, result based on only 12 subjects)*

The majority of the test group did also, in this case, favor an end-oriented grouping. However, it is interesting that at least three subjects did notate the second major section (second note system above) a begin within a rest, which is the first beat in the measure in the original setting. Thus these test persons indicated the possibility of phrases to begin at ‘empty’ beats, which is postulated in the current model.

The dominant interpretations correspond very well to the end- and start-oriented solutions given by the current model. The sub-segmentation of the second major section is not accounted for by the model.

A. Start-oriented interpretation

**fascinating2.mf**

B. End-oriented interpretation

**fascinating2.mf**

Figure 151. Computer model solutions for test melody 3j. End-oriented interpretation (NB The highest level of the second major section is not displayed correctly according to the analysis, due to deficiencies in the graphical interface)

The models interpretations accounted for 69.2 % of the segments and 77.4% of the segment boundaries notated by subjects.

### Test melody 3k – “Jesu bleibet meine Freude”: Latticed, multi-layered phrase structure

The *obligato* melody part from the piece “*Jesu bleibet meine Freude*” by J.S. Bach has been used as a recurrent example of a complex multi-layered phrase structure constituted by varied repeated segments. This melody, like “*Boléro*” by M. Ravel, belongs to the corpus of popularized classical melodies, known to most Western listeners even outside the community of classical musicians.

One question of particular interest regarding this melody was whether the test group would conceive the general structure of the piece or if they would conceive the melody as a latticed structure, a variation piece without principal structure. The hypothesis was that the principal section structure would be disguised to the majority of the subjects due to the frequent repetitions within the sections, the *latticed* structure.

The general experimental hypothesis was that the test group would recognize the similarity in spite of variation and use it as a basis for the conception of melodic structure.

n. (section level (attr. to subject n et al) also implied by some other subjects)

l. (2)

f. (4)

b. (2)

b. (3)

a. (6)

Figure 152. Results of notation of phrase structure of test melody 3k for subjects a-p. Tempo: 250 ms/ eighth note. NB A shortened (and transposed) version of the melody was used, where the interpretive choral sections were cut out, and replaced by very long notes to mark the end of sections. As not all subjects completed the whole piece, only the part analyzed by all subjects is included.

The experimental hypothesis was generally confirmed by the results, which demonstrated segmentation based on melodic similarity.

Also in this case one can find a complementary end-oriented and start-oriented grouping within the responses of the test group. However, in this case the end-oriented grouping was not as dominating as in several of the previous test melodies. The tendency to favor end-oriented grouping seemed to have been strongest when the test melodies contained interonset changes or offset-onset changes not coinciding with metrical impulse points for this test group. In the melodies with low rhythmical contrast, the tendency towards end-oriented grouping was much lower.

In conformity with e.g. the test melody 3g, the Norwegian *gangar* "Kjempe Jo", some of the subjects (subjects a, l and m) suggested overlapping phrase boundaries. This points to a structural similarity regarding the nesting of phrases by phrase end and start overlaps common to the two melodies and perhaps also to the two melody styles in general.

The principal section structure this was, in fact, noted by some individuals within the test group (5 persons out of 15), which may reflect a closer relation to the piece or a general experience to identify higher order structure. This was surprising in relation to the hypothesis, but one can hardly draw any general conclusions regarding the possibility of decoding the higher order structure from the small number of musically relatively skilled persons who participated in this test.

The phrase structure obtained by the application of the computer model shows that the start- and end-oriented solutions in this case also can account for a majority of the segmentations made by the test group. The model accounted for 84.5% of the segments and 95.7% of segment boundaries.

A. Start-oriented interpretation

jesuall.333.mf

120

A1  
A1  
B1  
C1  
D1

A2  
A1  
B1  
D1  
E1

B2  
E1  
F#2  
F#2

A3  
A1  
A1  
D1

B1  
D2  
A1  
A1  
A1

B1  
C1  
D1

A1  
A1  
B1  
D3  
E1

B2  
D#2  
D1

A4  
A1  
A1  
B1  
D4

A1  
A1  
A1  
B1  
C1

D1  
A2  
A1  
A1  
B1

The image displays a musical score with ten staves of music. Above and below the staves are various chord annotations, each enclosed in a small box. The annotations include:

- Staff 1: B1, B3
- Staff 2: A7, A1, A1, B1, A4, B1
- Staff 3: B1, A2, A8, A1, A1
- Staff 4: B1, B2, A2, A9
- Staff 5: A10, A1, A1, B1, D1
- Staff 6: D1, E1, B2, D1, D2
- Staff 7: D1, D3, A1, A1, A1
- Staff 8: B1, B1, C1, D1, B1
- Staff 9: A11, A1, A1, B1, E1

The music is written in a single melodic line on a treble clef staff. The annotations are placed above or below the staff lines, often with horizontal lines indicating their span across multiple measures.

B. End-oriented interpretation

jesuall.333.mf

The image shows a musical score for a piece titled "jesuall.333.mf". It consists of ten staves of music in treble clef, with a tempo marking of ♩ = 120. The music is written in a key with one sharp (F#) and a 3/4 time signature. Above each staff, there are guitar chord diagrams represented by boxes with letters and numbers. The chords are: A1, B1, C1, D1, A2, A1, B1, D1, E1, F1, B2, A3, A1, A1, B1, D2, A1, A1, B1, C1, D1, A1, A1, B1, D3, E1, J1, A2, A2, D1, A4, A1, B1, D4, A1, A1, B1, B1, C1, A2, A1, B1.

The figure displays ten staves of musical notation for test melody 3k. Each staff is annotated with chord labels and brackets indicating phrase boundaries. The labels include A1, B1, A4, B3, A7, A8, A9, D1, E1, B2, D3, and C1. The notation features treble clefs, a key signature of one sharp (F#), and various rhythmic values such as quarter, eighth, and sixteenth notes, along with rests.

Figure 153. Computer model solutions for test melody 3k. A start-oriented and B end-oriented interpretation

### Experiment 2, trial 4. Test melody 4: Metrical ambiguity at measure level by sequence contra-metrical to the established meter

This test is described in the *Evaluation of metrical analysis*, (section 5.3.4.3, *Experiment 2, Trial 4*), and involved the method of choice between different alternative metrical interpretations. There were predominantly different participants in this test than in experiment 3, and experiment 2 was performed a half year later.

The current melody, was composed in quasi baroque/classical style, in order to create an expectation of symmetrical structuring of the melody common to these styles. The melody was generally designed to test two different but related aspects of metrical conception in relation to phrase structure, consistency of metrical conception and significance of sequence phase in relation to sequence identity as structural determinants.

The melody was constructed as to indicate an even meter at measure level, 2-beat or 4-beat measures in the beginning of the melody. This melodic structure is interrupted in the mid part of the melody by a sequenced structure with a period of 3 beats at measure level and, in the third part of the melody the theme with the 2/4-beat structure returned, with a short 3-beat sequence indication at the end. The object of this construction was to test whether the test group would recognize the change of melodic structure to be of metrical significance (see fig. 154 below).

Yet another conflict was built in the melody, by using a typical ending melodic phrase as the start in the first 2/4-beat melodic structure, overlapped by a sequence starting at the longest note within the phrase, making the first phrase sequenced by end-similarity. At the return of this structure in the third part of the melody, the structure was repeated parting of overlapping sequence, at the longest note (see below). This design was used to test factors determining phase of sequence, such as preference for start-after-long-note (*separation start*) vs. start-on-long-note (*grave accent start*), with reference to perfect vs. first presented sequence preference. Further, it involves the question of the relationship between phase and melodic content – will the same melodic structure be consistently conceived identically regarding phase throughout the melody.

The figure shows a musical score for Test Melody 2:4, consisting of three staves of music. The first staff (measures 1-10) is annotated with 'fem. seq.', 'basic group:', 'masc. seq.', and '3-beat mel. seq.'. The second staff (measures 11-16) is annotated with 'cont. 3-beat grp. (beat division)', '4-beat seq. (fem.)', and '4-beat seq. (masc.)'. The third staff (measures 17-20) is annotated with '3-beat seq.'.

Figure 154. Structural construction of test melody 2:4. Sequences and structural units marked by brackets.

The test stimulus involved four different metrical conceptions indicated by percussion accompaniment, corresponding to four alternative possible metrical interpretations related to phrase structure:

1. Consistent grouping in four beats per measure throughout the melody with start on the initial beat, *separation-start* interpretation. This corresponded to the alternative hypothesis that listeners would use the first identified grouping structure as input to metrical conception and would stick to it in principle “no matter what happens” as long as the central pulse prevails. This implies that the metrical conception of the initial melodic phrase will change throughout the melody.
2. Grouping in four beats per measure, but consistent *grave-accent-start* interpretation of the 4-beat sequence, with the first complete measure starting at the third beat from the beginning and a correspondent phase shift at the recurrence of the theme in the third part. This interpretation corresponds to the alternative hypothesis that listeners will use the same interpretation of a melodic material throughout the melody (content identity based grouping) and that the consequent *grave-accent* sequence with start-similarity would override the overlapped *separation-start* sequence defined by end-similarity.
3. Change of metrical conception in accordance with the 3-beat sequence in the mid part of the melody corresponding to the alternative hypothesis that the repeated sequence would be salient and persistent enough to break the initial grouping 2/4 beat-grouping. In other respects, it could be interpreted according to the interpretation 1, implying that the sequence based on end similarity will be more salient due to start on the first event of the melody and the *separation-start* being more perceptually prominent than the *grave-accent-start* because of the separation involving two beats.
4. Change of metrical conception like in 3., but with consistent *grave accent start* interpretation, like in 2.

The result of this test (see section 5.3.4.3, *Experiment 2, Trial 4*) demonstrated the ambiguity of this test melody and the possibility of the above alternative hypotheses. The dominant interpretations were interpretation 1 and interpretation 4, represent consistent general metrical conception at measure level based on the first presented pattern and varied metrical conception at measure level based on sequence patterns respectively.

Both these interpretations demonstrate the relationship between phrase structure and metrical conception; the metrical conception at measure level is based on conception of phrase structure, but to a different degree. The metrical interpretation 1 can be described as if the subject takes the first presented phrase structure as input to a metrical conception that is withheld in spite of changes in the melodic structure as long as the basic metrical structure, i.e. the central pulse/*tactus*, prevails. The metrical interpretation 4 shows a more direct relationship between conceived phrase structure and metrical conception, where the metrical grouping at measure level is drawn directly from the phrase structure and can be continuously revised.

The problem in interpreting these results is evidently if the phrase structure used as a basis for metrical interpretation 4, was at all recognized by the subjects preferring the metrical interpretation 1? This question cannot be answered from the methodology of this test. However, some of the other tests performed with the same test group showed significant sensitivity to changes in melodic structure involving use of melodic sequence as a basis for

Chapter 6

metrical conception. Therefore, one cannot draw the conclusion that melodic sequence was not a significant structural feature for these test persons. Hence, of to what extent the conceived metrical structure reflected a conceived phrase structure for these test persons remains an unanswered question.

For the other half of the test group, the direct relationship between phrase structure and metrical structure was evident by the results.

The current model provides a phrase structure, which is consistent with the metrical interpretation 3, recognizing the change of sequence structure but disregarding consistent grouping based on phrase identity. The overemphasis of this interpretation in relation to the interpretation consistent with metrical interpretation 4 is, however, relatively weak, which makes the major section interpretation non-consistent.

A. Start-oriented interpretation

assfras1a.ny.mf

The image shows a musical score in 4/4 time with a treble clef. The score consists of six staves of music. Above the first staff, there are two boxes labeled 'A1' and 'A2'. Above the second staff, there are four boxes labeled 'B1', 'C1', 'D1', and 'C2'. Above the third staff, there are three boxes labeled 'D1', 'E1', and 'F1'. Above the fourth staff, there are four boxes labeled 'E1', 'H1', 'E2', and 'C2'. Above the fifth staff, there is one box labeled 'F1'. The annotations are represented by horizontal lines with small square boxes at the beginning and end, indicating the start and end of phrases.

B. End-oriented interpretation

assfras1a.ny.mf

Figure 155. Computer model solution for test melody 4, exp. 2. A start-oriented and B end-oriented interpretation

A measure of the extent to which the model accounts for subjects responses is pointless in this case because of the difficulty of interpreting the results. However, the start-oriented version, which is more consistent with the metrical interpretation than the end-oriented version, is not consistent regarding phrase similarity throughout the melody. This makes it less consistent with the most preferred of the two alternative versions with phrase related metrical structure.

This failure of the model may be considered an implementation error but still reflects the ambiguity of the melodic structure; none of the structural cues being prominent enough to provide a consistent interpretation.

The results of this test thus indicate that

- Meter at measure level is primarily derived from and related to phrase structure;
- The first presented grouping pattern is most important for the conception of meter at measure level;
- Changes in implied periodicity in phrase structure influences to a variable degree the conception of meter at measure level;
- Consistent metrical conception of identical melodic structures cannot be taken for granted, hence pointing to the limited influence of similarity as a structural principle between non-adjacent segments

## Experiment 2, trials 8-9. Test melodies 8-9. Phrase and metrical ambiguity at measure level due to overlapping sequences

If the previously related test within experiment 2 was difficult to interpret regarding phrase conception because of the construction of stimuli, the trials 8-9 are more closely related to the question of phrase conception.

The methodology here was to demonstrate the different alternative conceptions of the melodic structure by inserting rests between and alter tempo within implied phrases (see section 5.3.4.3, *Experiment 2, Trials 8-9*). This has been shown to be common means of expression of a phrase conception in melodies (see e.g. Bengtsson & Gabriellson 1983). The subjects were asked to rank the alternative interpretations with regards to how well they fit with the structure of the melody. The test melodies were composed to indicate asymmetrical patterns based on overlapping sequences, implying a latticed structure of similar though not identical melodic segments.

Figure 156. Composition of test melodies used in experiment 2, trials 8 and 9. An alternative sequence implication is noted below. The sequence designation is described in section 6.2.3.3.

As can be seen from this overview of test melodies 8 and 9, the main difference between them is that test melody 8 involves perfect repetition, NIL 5 sequences, while test melody 9 involves transposition, NIL 4 sequences. Moreover, overlapping sequences begin consistently on the last beat of the previous sequence.

The central question behind this construction was whether the segment boundaries implied by an earlier sequence would override the implications of an overlapping sequence of a different period? If this were to be true, the motif similarity between adjacent sequences would have to be disregarded by subjects, as in the alternative sequence noted below the systems. If the overlapping sequences were recognized as significant, this would indicate that the melodic gestalt was regarded as significant as a basis for the conception of the structure.

On the basis of the results of the pre-test four alternative interpretations were made, each demonstrating different segmentation strategies:

1. Consistent metrical grouping in four beats (4/4) based on first sequence implication (see above).

2. Grouping only by structural discontinuities, i.e. pitch distance (lowest level above)
3. Asymmetrical metrical grouping by adjacent overlapping sequences
4. No particular consistent grouping due to structural ambiguity (indicated by inclusion of a consistent grouping in three beats (3/4).

The results of the two trials showed a statistically significant preference for asymmetrical grouping based on sequence structure (strategy 3) above, with overlapping sequences.

This was the dominant interpretation in both test melodies, hence no particular difference was shown between transposition and perfect sequential repetition.

The problem of interpreting the results is in this case rather the opposite of *trial 4*, concerning the metrical implication of the results instead of the implications regarding phrase structure. The design of the stimulus material here uses phenomenal cues typically connected with phrase conception rather than metrical conception. However, considering the strong connection between conceived metrical structure at measure level and phrase structure, the latter giving the fundamental input to the former. It seems reasonable that this phrase structure can be assumed to have metrical implication.

A. Test melody 8.

**assfras4a.ny.mf**

B. Test melody 9

**assfras4b.ex.mf**

Figure 157. Computer model solution for test melody 8 and 9, exp. 2. NB Start and end-oriented interpretations identical.

The computer model solution gives a result, which is concurrent with the most preferred alternative interpretation by the test group. The result of the test thus supports the

assumption of overlapping sequences defined by start-similarity to be significant in the determination of melodic structure (cf. interpretation of *experiment 2, trial 4* above).

Therefore this result supports the crucial role of sequence structure for melody conception which is assumed in the current model.

#### 6.3.7.4 Summary of results of listeners tests

In general, the results show that a combination of the start- and end-oriented interpretations given by the current model accounts for a majority of the responses given by the test groups.

The variability of responses was generally considerable in spite of the limited cultural background of the participants; a larger test group and cross-cultural tests would be useful in determining the generality of the model. However, the methods employed in the experiments are not optimal, and further developments of test methodology regarding phrase structure and the correspondent relationship to metrical structure would be required before the generality of the model could be tested more thoroughly.

The considerable variability in responses indicates that the conception of melodic phrase and section structure is not unambiguous even among people with similar background, which is also indicated by the test performed by Hötkher et al (Hötkher et al 2002). However, the variability seems to be restricted to a few strategies, since the majority of these results can be accounted for by the two different complementary strategies depicted by the current model.

The results thus support the model by indicating the cognitive reality of the model's predictions. In summary the results indicate:

- That there is considerable individual difference in the conception of phrase structure among people given a equalized non-interpreted stimulus regarding duration changes, articulation and tempo variations.
- That people can abstract certain culturally acknowledged features of a melodic structure even when confronted with culturally relatively foreign melodies distinguished by unfamiliar structural features, without cues given by the cultural context. Thus the result indicates that the general cognitive principles which govern the current model to a certain degree are relevant for melodic cognition in some relatively distant musical cultures and musical styles where the concept of melody is applicable.
- The primacy of metrical structure at beat level as a time grid for conception of phrase structure.
- The primacy of global melodic grouping principles, i.e. similarity and dissimilarity between groups of events, before grouping by local discontinuity in the conception of melodic structure.
- The relationship between metrical structure at measure level and phrase structure is indicated by the results, where conception of phrase structure functions as input to metrical conception, which in turn may implicate phrase structure. However, the connection between conceived meter at measure level and phrase structure seem to vary in strength between different individuals.
- The complementary nature of start- and end-oriented grouping is interlinked in the conception of phrase structure in melodies with metrical structure. The results further indicate that there are individual preferences for start- and end-oriented grouping and

that people can experience them as simultaneously interacting within a structure. It seems that end-oriented grouping is preferred when the experimental question regards phrase structure.

- That similarity regarding melodic content, melodic parallelism, is a stronger structural factor when it can determine adjacent melodic structures than in the determination of distant melodic structures.
- That start-similarity is more prominent than end-similarity
- That the first identified/established pattern has a strong influence on the conception of continuing subsequent structures.

However, one has to consider that these are indications, not evidence, due to the limitations of the experiment. In general the results nevertheless support the cognitive relevance of the current model.

## **6.3.8 Comparison with other models**

### **6.3.8.1 Evaluation of Grouper and LBDM models**

There are especially two relatively new computer based models of cognition of melodic structure that are relevant in relation to the current model, the Grouper algorithm within the Melisma system developed by Temperley and Sleator (Temperley 2001) and the Local Boundary Detection Model (LBDM) together with the Sequential Pattern-Matching Algorithm (SPMA) within the General Computational Theory of Musical Structure developed by Cambouropoulos (Cambouropoulos 1998). (see also Chapter 2, *Related works*.)

These two models (obviously developed independently of one another, in parallel with the development of the current model) have many features in common with the current model. Both are rule-based systems, based on assumptions of the nature of human cognition of music. These assumptions are essentially based on applying gestalt principles to music cognition. Neither of the two models limits itself to a certain musical style, rather such limitations seem to be inferred from the implementations of the models for practical reasons. Both models aim, in this respect (in parallel with the current model), at modeling human cognition of musical surface structure, approaching musical structure from the listener's point of view.

The greatest difference between these models and the current model is that the former have a much wider scope regarding aspects of musical surface structure. Both include, at least partially, aspects of polyphony, tonality and harmony in the analysis. However, when it comes to segmentation, both models focus on monophony in the proposed implementations and do not include aspects of tonality, which make them comparable to the current model.

There are also some other important differences: Temperley and Sleator's algorithm Grouper disregards pitch information, which makes it less useful with regards to a great number of melodies in different musical styles that are characterized by low rhythmic contrast. Further, Grouper requires prior metrical analysis at measure level, which is problematic since phrase structure, as have been demonstrated, influences conception of metrical structure at that level (see section 6.3.7.3, test results). Cambouropoulos model of sequential pattern-matching does not involve successive evaluation of sequence-structure thus neglecting aspects of symmetrical implication and good continuation. Like a proposed model of involving

## *Chapter 6*

parallelism as a factor in metrical analysis by Temperley and Bartlette (2002), it is limited in its treatment of melodic similarity since it considers only note-to-note similarity concerning pitch and rhythm and does not formally handle metric structure in the categorization of melodic parallelism.

Höthker, Spevak and Thom (2002a, 2002b, 2002c) have evaluated the success of the two models mentioned above in relation to segmentations of melodies made by a group of listeners using a similar methodology as was used in experiment 3 (see section 6.3.7.2). Unfortunately, they have used only the LBDM module of Cambouropoulos model, which makes it a bit difficult to compare the results of this with the results of the current model. But nevertheless, this evaluation is the only independent evaluation of the performance of the two models I have come across, which makes it interesting to compare the results with the performance of the current model.

Table 41. Table displaying listener's segmentations in relation to computer generated segmentations.  
 From Hötkher, Spevak and Thom (2002a)

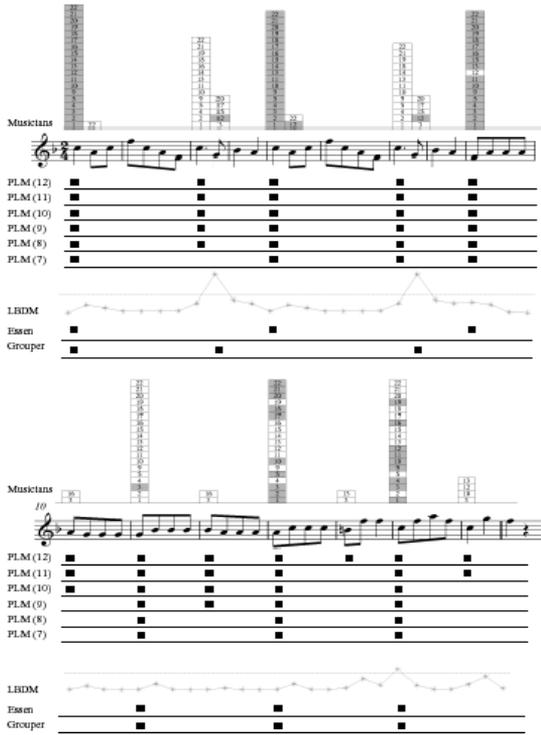


Fig. 1. Example 1: Folk song E0356, annotated with segmentation results from the musicians (upper histogram) and the algorithms (lower table). Each histogram count number identifies a particular musician. Grey squares identify the first note of both a phrase and sub-phrase; white squares correspond to sub-phrases only. The black squares refer to segmentations generated by the Position Length Model (Sect. 5.5). The last three rows show the boundary strength calculated by LBDM, the segmentation provided in the Essen collection, and Grouper's solution. The dashed line demonstrates how LBDM uses its threshold to derive  $s$  from  $w$ .

Unfortunately I have not been able to make this comparison with the entire test material used by Hötkher et al, but only the examples published in their papers. However, this comparison shows some interesting features of the behavior of the current model in relation to the two other models and regarding the results from their experiment.

As can be seen in Table 15, Hötkher et al basically compares the results of the listener segmentations with the algorithms performance in terms of coinciding individual segment boundaries. In their experimental task, they restricted the number of hierarchic levels of segments to two, designated phrase and sub-phrase levels, which is important when comparing the results with the current model. In general, Hötkher et al report that LBDM and

Groupers produce results, which do not account for the results of the experiment to a significantly high degree (Hötkher et al 2002b).

### 6.3.8.2 Comparisons of results

#### Folk song from the Essen collection

Hötkher et al, provides two examples of their results in the published papers, with annotations of segmentations by listeners and predicted segments by the two tested computer models.

The first is a folk song from the Essen collection of European folk songs, displayed in Table 15 above. The other is an excerpt of *Fugue XV* of “*Das wohltemperiertes Klavier*” (BWV 860) by J.S. Bach.

For the folk song the segmentations performed by the current model are displayed in fig. 159.

#### A. start-oriented interpretation

The image shows a musical score for a folk song in 2/4 time, starting at measure 160. The score is divided into three systems. The first system contains the first 12 measures, with segmentations labeled A1 (measures 1-6), B1 (measures 7-12), and A1 (measures 13-18). The second system contains the next 12 measures, with segmentations labeled B1 (measures 19-24), C1 (measures 25-30), and C2 (measures 31-36). The third system contains the final 4 measures, with a segmentation labeled C3 (measures 37-40). The notation includes a treble clef, a key signature of one flat, and a tempo marking of 160.

## B. end-oriented interpretation

The figure displays three staves of musical notation for folk song E0356. Above the first staff, there are two rows of brackets indicating segment boundaries. The first row shows segments A1 (covering the first two measures), A1 (covering the next two measures), and A1 (covering the last two measures). The second row shows segments A1 (covering the first two measures), B1 (covering the next two measures), A1 (covering the next two measures), and B1 (covering the last two measures). The first staff has a tempo marking of ♩160. The second staff has three rows of brackets: the first row shows B1 (covering the first two measures) and C1 (covering the next two measures); the second row shows C1 (covering the first two measures) and C2 (covering the next two measures); the third row shows C3 (covering the first two measures). The third staff has one row of brackets showing C3 (covering the first two measures).

Figure 158. Computer model solutions for folk song E0356. A start-oriented and B end-oriented interpretation

When comparing the interpretations made by the current model with the results obtained by Höthker et al from listener tests, the similarities are obvious. In statistical terms, the current model can account for 87.2 % of the segments given by the listeners in this experiment, and a great majority of segment boundaries, 96.1 %. Corresponding figures for Grouper is 24.9% of the segments and 55.8% of the segment boundaries and for LBDM 8.6% of the segments and 34.8% of segment boundaries.

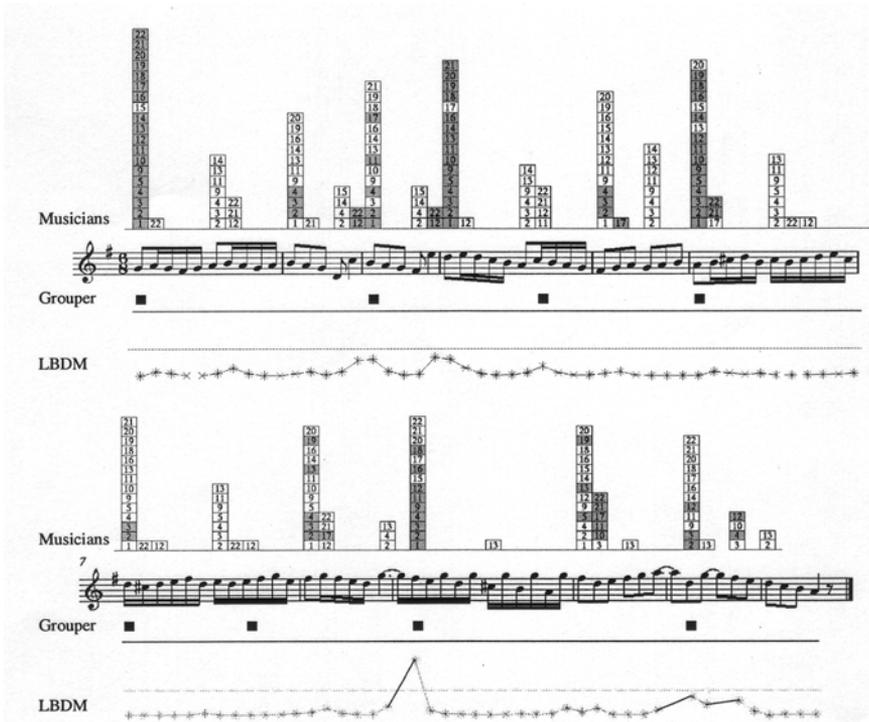
However, it has to be emphasized that with the inclusion of the sequential pattern matching algorithm (SPMA) proposed by Cambouropoulos the results for LBDM + SPMA would be probably more in the neighborhood of the results of the current model. It is also worth noticing that the better performance for Grouper than for LBDM is probably due to the metrical interpretation being input to the analysis.

The dominant interpretation given by the test group is most consistent with the start-oriented interpretation by the current model, which is not especially surprising considering the few possible upbeat interpretations within a structure which is determined basically by perfect sequences.

### Fugue XV of “Das wohltemperiertes Klavier” by J.S. Bach

The second example provided by Höthker et al is an excerpt from *Fugue XV* of “*Das wohltemperiertes Klavier*” ( BWV 860) by J.S. Bach, part 1. This example was chosen by the researchers because its complexity, which was predicted to give an ambiguous phrase interpretation by the test group. And indeed, the result displays a number of different segmentations:

Table 42. Table displaying listener's segmentations in relation to computer-generated segmentations in Fugue XV by J.S. Bach. From Hötkker, Spevak and Thom (2002a)



The segmentations made by listeners in this test melody display a similar level of variation of responses as for many of the test melodies in experiment 3 (see section 6.3.7.3).

They do, however, converge into a more limited number of compatible and complementary interpretations. As Hötkker et al point out, a great deal of the variability concerns granularity of the phrase interpretation, which hierarchical level is assigned phrase and sub-phrase level respectively.

The solutions provided by the current model, in this case, can also account for a majority of the phrase interpretations made by the listeners in the test by Hötkker et al.

A. Start-oriented solution

B. end-oriented interpretation

Figure 159. Computer model solutions for Fugue XV by J.S. Bach. A start-oriented and B end-oriented solution

The complementary results of the current model account for 63.9 % of the total of segments and 89.5% of the segment boundaries given by the test group. In this case, the

results of applying Grouper and LBDM to the test melodies are significantly lower: The segmentation by Grouper can be interpreted to account for 8.6% of the segments and 40.9% of the segment boundaries, while LBDM only can be interpreted to account for 2.0% of the segments and 11.9% of the segment boundaries.

This significant difference in the results can largely be explained by Grouper algorithm not taking pitch information into account, and the lack of the sequential pattern-matching in the use of LBDM in isolation. Moreover, there is no hierarchical phrase analysis employed in either of the models, which lowers the possibility account for hierarchical phrase analyses.

The most evident deficiencies in the solutions of the current model are the lack of the super-phrase levels, such as e.g. the segment (0 18) made up from the sub-phrases a1a1-B1-B1. This is expressed internally by the sequence (0 18 x 0 6), defined by global changes of rhythm and melodic direction, but is excised in the evaluation process since it does not display global structural similarity for more than half of its length.

To obtain a higher level structure one would have to combine the phrase elements into super-phrases and sections based on clustering of similar adjacent phrases, a feature of the model which is not yet included in the implementation.

## Discussion

In general, the current model performs considerably better than the two tested models in terms of prediction of the total of segments and segment boundaries noted by the test group in the two example melodies given by Höthker et al. This can be explained by the greater number of structural parameters used by the current model. This can also be regarded as an indication that these parameters are essential for the analysis of melodic surface structure: Since Grouper is limited to duration information and needs metrical structure as input it will probably give poor results for melodic structures with low rhythmic contrast, such as the *Presto movement* by J.S. Bach (see section 6.3.5.1). LBDM, which does not take sequence structure into account, will (as Cambouropoulos points out (1998:109)) perform poorly in melodies whose structure is determined by melodic parallelism. The poor results of LBDM alone can be said to corroborate the hypothesis of the prominence of sequence structure.

LBDM is, as mentioned, part of a larger model for computer-assisted music analysis, focusing on melody analysis, developed by Cambouropoulos (1998). This model, GCTMS, incorporates metrical analysis, sequential pattern matching (sequence analysis), phrase categorization and category based bottom-up hierarchical analysis of phrases. Thus, this model does, in many respects, resemble the current model and a more comparable evaluation of the model developed by Cambouropoulos in relation to the current model should rather relate to the more elaborate GCTMS, than a single module, such as LBDM. Cambouropoulos provides a few examples of the application of the model, which was not implemented at the time of the description. Because of the reliance of manual selection of results involved in the analysis process, the model is hard to evaluate, especially regarding selection of pertinent patterns.

One of the examples given by Cambouropoulos is the finale theme of Beethoven's 9<sup>th</sup> Symphony, which can serve as an example of the model design. From the application of the LBDM and the sequential pattern-matching algorithm (SPMA), he derives a metrical grid (cf. Chapter 5, section 5.2.4 *Complex Metrical Analysis*), which is used in the evaluation of sequence structure obtained by further application of the SPMA. Neither the result of the sequential

pattern analysis, nor the selection of pertinent patterns from this analysis is clear from Cambouropoulos example, but, according to his example it ends up in a segmentation at sub-phrase level, mainly coincident with the meter at measure level.



Figure 9.3 Segmentation of the finale theme of Beethoven's 9th Symphony

Figure 160. From Cambouropoulos: *Towards a General Computational Theory of Musical Structure* (1998:111)

From this segmentation, Cambouropoulos performs a categorization of the segments based on similarity by a module called the Unscramble algorithm. This algorithm performs a categorization based on different threshold values given by the analyst, resembling the phrase identity categorization within the current model (see section 6.2.4.5). The 'best' result is picked by the analyst.

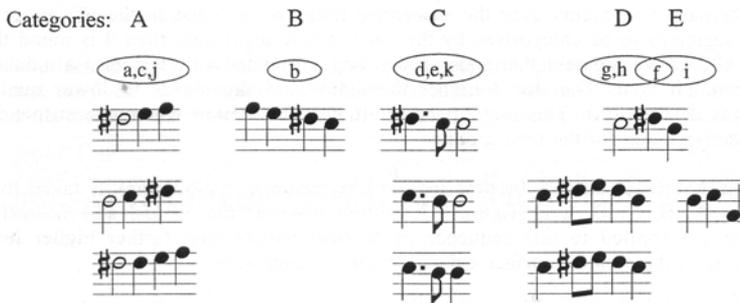


Figure 9.5 The 'best' classificatory description revealed by the *Unscramble* algorithm.

Figure 161. From Cambouropoulos (1998:113)

From this categorization higher levels are constructed, by the application of SPMA and a Selection Function module, resulting in a hierarchical analysis:

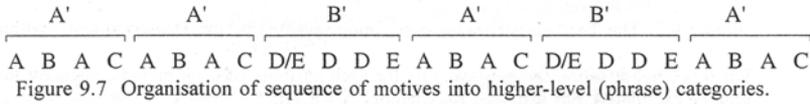


Figure 162. From Cambouropoulos (1998:114)

However, how the sequential pattern-matching algorithm and selection function can be applied to this meta-level analysis is not clear from the description provided by Cambouropoulos.

This analysis both resembles a structural analysis of the same piece by Lerdahl & Jackendoff (1983:124), which functions as the reference for Cambouropoulos segmentation example, as well as the solution which is the output of the current model when applied to this piece.

**beethoven9.fin2.mf**

Figure 163. Structural analysis of the finale theme of Beethoven's 9<sup>th</sup> Symphony by the current model (start-oriented interpretation).

Cambouropoulos approach is essentially bottom-up, where higher levels are constructed from lower levels. A prototype of the current model (Ahlbäck and Emtell1993) had a similar approach. The hypothesis that all higher-level segments can be constructed from lower level segmentation had to be rejected for a number of reasons:

- The categorization of identity of lower level segments, as Cambouropoulos himself points out, usually results in overlapping categories. (See e.g., to my mind the erroneous, but still reasonable A1-A4-A5-A6-B2 categorization of the lowest phrase level above). Since the analysis of higher levels depends on this categorization there will be no means of evaluating which of the overlapping similarities will be salient if the syntactic function of subsequent segments is not considered.
- Lower level segments can change their structural position within higher level segments, due to higher level structural relationships, as demonstrated in several examples (see sections 6.3.3 and 6.3.4). Thus, implication of lower level segment category identity will be ambiguous for the abstraction of higher level structures.
- Lower level segments are often defined by e.g. symmetrical implication of higher level segments. In the Beethoven example above, the numerous possible different lower level segmentations have to be evaluated in relation to higher-level similarities including evaluation by symmetrical implication.
- Generally, since the significance of a similarity between melodic structures is dependent on the number of similar events and duration of the similarity, it follows that local sequences and local discontinuity can be disguised by global similarities, which has been demonstrated in several examples (see e.g. section 6.3.4). Thus it follows that global segments defined by similarity cannot always be revealed by abstraction of local segments. Cambouropoulos model will thus run into problems when encountered with such melodic structures.

There are also some further general problems inherent in the model proposed by Cambouropoulos:

- The selection of the lower level segments identified by Cambouropoulos model seem to rely heavily on prior metrical analysis on measure level. As has been demonstrated, metrical analysis on measure level is, in many cases, highly dependent on phrase structure. (see e.g. section 6.3.7.3 *test melody 3b*). Thus, the syntactic analysis of melodic structure at phrase level has to be input to final metrical analysis at measure level.
- The selection of pertinent patterns is significantly related to the frequency of occurrence of patterns. This methodology, in opposition to the general design of the model, disregards the listeners view by omitting the temporal dimension from the conception of melodic structure. As has been demonstrated, the first presented pattern in a row of overlapping patterns seems to have a relatively strong impact on the conception of subsequent structures; repetition of segments of a certain length implies periodical continuation etc.

Höthker's et al main criticism of the two tested models regard what they call their deterministic design, both models essentially provide a single output without taking ambiguity and relative strength of the predictions into account.

This criticism can be applied to the current model, since it does not, in its present implementation, provide more than two alternative analyses or a fitness value designating the strength of the segmentation in relation to other segmentations.

The main reason for this is that such an output goes beyond the scope of the present study, which does not primarily concern the construction of an efficient segmentation algorithm, but the study of the cognition of melodic surface structure where the algorithm rather serves as a test of the proposed model of melody cognition. The test of the generality of the model is dependent on that the same factors are tested with the same weights in

different contexts. However, it would be possible to present further alternative segmentations, either based on different weights of selective parameters or by using alternative segmentations above a certain threshold to produce parallel syntactically coherent analyses. One has to consider, however, that the successive syntactic relationship between segments is the core of the analysis; a second choice of an initial segmentation may affect all subsequent segmentations. Thus, one cannot regard single segment boundaries in isolation, when they are not only determined by local discontinuity factors.

Höthker et al argue for a probabilistic methodology, in the line of Bod (2001), which has developed a model employing machine adaptation to a corpus of manual segmentations of melodies, using probabilistic grammar technique. This and other current models involving neural network methodology, may be very successful in modeling and predicting peoples responses to a certain repertoire given 1) a sufficient test material 2) a concurrent conceptions between the people who provide input to the program and those who judge the results. However, the apparent question in this context is what the analyst will learn about the nature of melody cognition from the program? Since this study has a different focus, these approaches are not relevant in the current context.

## 6.3.9 Conclusions

### 6.3.9.1 General success of the model

The fundamental hypothesis of this study is that melody cognition is based on perception of structural properties of melodies, which can be described in cognitive terms and based on the perceptual and cognitive capabilities we share as humans. From this it follows that melody cognition is not regarded as arbitrary in relation to melodic structure, nor is it regarded primarily a matter of cultural learning, even if the latter may have strong impact on the cognitive process.

The basic claim of the study is that a rule-based method of analysis, based on general assumptions of the nature of melody cognition and gestalt laws, can produce syntactic segmentations of melodic structure which conform to people's experience of melody structure better than chance in monophonic melodies.

Given this hypothesis, one may regard the results of the listening experiments as convincingly demonstrating that melodic cognition is really arbitrary in relation to the structural dimensions of melodies that were involved in the experiment, thus negating the hypothesis. The number of different segmentations given by a relatively small number of test participants with similar cultural background may well be interpreted as to indicate that conception of melodic structure is, in fact, individual, arbitrary and that concurrent segmentations are coincidental.

While a vast majority of previous music theorists have not been occupied by stressing the ambiguous nature of melodic structure (or musical structure in general), but rather have been trying to provide optimal structural descriptions, the ambiguity of melody structure has been emphasized in different recent studies. Temperley, who himself developed a computer-aided

method for structural analysis of monophonic melodies, does in fact question the value of such an undertaking:<sup>120</sup>

*“...it seems reckless to attempt a computational model of an aspect of perception in which it is so often unclear what the “correct” analysis would be”(Temperley 2001:65)*

However, as shown in the description of the results, the majority of the segmentations given by subjects are definitely not arbitrary (see section 6.3.7.3). The results demonstrate ambiguity in melodic conception, which here is regarded as a typical quality of cognition of melody structure. Since there is no need for a precise meaning to be conveyed to the listener, a melody must only be syntactically coherent to the limit that it will be possible to segment into conceivable units which can be grasped within the perceptual present and can be understood as parts of a comprehensible whole. The level of ambiguity of a melodic structure is according to this view, related to the number and strength of competing structural cues and lack of structural contrast.

Thus, the basic claim of the study can be viewed in the light of ambiguity, allowing for complementary interpretations of a melodic structure.

From this perspective, the results of the evaluation, involving comparison with expert analyses and listener tests, strongly supports the basic claim, since the majority of the segmentations can be accounted for by a rule-based model governed by general assumptions of the nature of melody cognition and gestalt laws.

Further, it supports the claim that, for a large corpus of melodies of different stylistic and cultural origin, information of event duration (IOIs, OOIIs) and relative pitch is sufficient to extract a melody structure, with relevance for people’s conceptions of melody structure.

### **6.3.9.2 Interpretation of the results in relation to the basic assumptions for the model**

The results generally support the *The General Model* presented in Chapter 1/3, which forms the basis of the current method of analysis.

#### **General assumptions about the nature of melody cognition**

The cognitive reality of at least two levels of sub-structural levels in melody cognition is corroborated by the evaluation, since both expert analyses and listener test segmentations exhibit these levels.

The cognitive reality of top-down processes in the conception of a melody, reflecting the view of melody cognition as a process that takes place in time and can be revised throughout the process, is corroborated by the evaluation. This is demonstrated by the prominence of large-scale grouping in relation to local grouping.

The influence of local structure on global structure – bottom-up processing, e.g. by determining phase of sequences, is corroborated by the results of evaluation.

---

<sup>120</sup> Temperley’s comment regards Western common-practice music. The reference example happens to be a perfect example of complementary start- and end-oriented grouping preference. However, Temperley considers Western folk melodies to be a repertoire “where these problems do not arise”. The current study as well as the study by Höthker et al, does not support this notion.

## The relationship between metrical structure and phrase structure

The dominant interpretations of phrase structure are, as demonstrated by listener tests and by expert analyses, congruent with metrical structure, which strongly supports the assumption that the primacy of metrical structure at central pulse/tactus level is the fundamental temporal grid in melodies with conceived metrical structure. The results indicate that structural cues for determining phrase structure, such as perfect sequence repetition, that have been shown to be significant in phrase conception are disregarded by listeners and in expert analyses when they do not concur with conceived beats at central pulse level.

The assumed interrelationship between meter and phrase structure is supported by the results; conception of phrase structure is the input for metrical conception at measure level, while metrical conception influences phrase structure by implication of continuity regarding phrase length and phrase phase. This is indicated by e.g. listener interpretations of metrical structure at measure level; it is congruent with notated phrase structure and the preference for a phrase structure that chunks the melody into contiguous phrases of equal length.

The assumed relative importance of metrical structure for the conception of phrase structure reduced for segments above measure level is corroborated by the evaluation. *Aperiodicity* regarding segment length for segments above measure level is much more common throughout the tested material than *aperiodicity* between segments at measure level.

## Start- and end-oriented grouping

The cognitive reality of the concept of complementary start- and end-oriented grouping preference is strongly supported by the results. A majority of the phrase notations provided by listeners in experiment 3 and other examples from the literature could be accounted for by the application of the two interrelated segmentation strategies.

The results indicate that the tendency to favor start- and end-oriented interpretation varies among listeners with similar cultural background. And, since there were almost no results that unambiguously supported one interpretation such a variation regarding conception of segment boundaries may be typical for melodic cognition. However, the preference for start- and end-oriented grouping respectively seem to be related to structural cues such as consistency regarding in-phase and phase-shifted segmentation cues.

The complementary nature of start- and end-oriented grouping was demonstrated by e.g. numerous notations of start- and end-overlaps, which support the hypothesis that both interpretations can coexist in the cognition of a melodic phrase structure. This tendency seems, in the current tests, to be related to the relative strength of the different structural cues; the cues are relatively equal in strength, more overlaps are noted.

These relationships between favor of start- and end-oriented and phrase overlaps may be used to make a single prediction of most favored phrase structure based on a valuation of the structural cues.

The prominence of start-oriented grouping as a cue for metrical conception was also indicated by the evaluation. Metrical notations in transcriptions could be accounted for by the result of the start-oriented segmentation and subjects tended to notate a metrical structure that coincided with start-oriented grouping, in- or out-of-phase with the notated phrase structure.

## Segmentation by similarity – the sequence concept

The significance of segmentation by similarity in melody cognition is demonstrated by the general success of the application of the current model, which is designed as to favor grouping by similarity before grouping by discontinuity. The results of the evaluation corroborate this assumption by the general agreement with the results provided by the model.

The significance of the sequence concept – the special status of contiguous similar series of events as structural determinants of melodic structure – is strongly supported by the results of the evaluation. Probably the most evident example is the transition of structural position of similar sub-structures within larger segments between non-adjacent phrase segments, while structural position of similarity has been shown to be significant for adjacent segments (see e.g. section 6.3.4). Since segments are delineated successively, a segment boundary defined by similarity will basically consider pair-wise relationship between adjacent series of events.

The influence of local and global discontinuity on sequence recognition is supported by evidence of similarity categorization of segments or more general similarity, while sequences of more specific similarity are disregarded or regarded subordinate in the same structure (see e.g. section 6.3.3).

## Means of melodic structure – pitch structure

The cognitive reality of the different aspects of pitch and rhythmic structure are supported by listener and expert segmentations and designation of similarity between segments. Generally the relevance of these concepts can be traced from the melody examples provided in the description of sequence types and in the examples presented in relation to the expert analyses evaluation. Clear general examples are, for instance, the categorizations of similar motifs provided by Levy (section 6.3.4) and Middleton (6.3.5), which exhibit most of the structural levels assumed in the model.

The primacy of pitch-change-similarity at the level of melodic pitch categories above specific pitch-similarity is corroborated by e.g. the results of experiment 2, test melodies 8 and 9, where identical phrasing is chosen regardless of transposition or perfect repetition. Absolute pitch similarity is also disregarded as structural determinant e.g. the segmentations of “*Fascinatin’ Rhythm*” and “*Jesu Bleibet Meine Freude*”, where subjects regarded similarity in terms of pitch change as more structurally significant than the dissimilarity regarding absolute pitch. Absolute pitch dissimilarity and more specific pitch change dissimilarity than on MPC level seem to be discriminating only when a segmentation requires this level to be recognized, which implies that MPC change is identical.

On the other hand e.g. pitch-set-similarity has proved to be structurally significant, such as in *Polska* after Per Danielsson (see section 6.3.3.2), when melodic contour differs between correspondent segments. General pitch-change-similarity can also involve inversion of melodic contour, such as in the *Presto movement* by J.S. Bach (see section 6.3.5.1).

The significance of the step-leap dichotomization emerges from the significance of type “NIL 3” sequences, which equalizes pitch change into the categories step, leaps and large leaps. “NIL 3” sequence similarity is frequently required for the identification of similarity between segments in a number of the tested melodies, such as “*Jesu Bleibet Meine Freude*”, the Svenska Låtar examples analyzed by Övergaard and the segment categorization of “*A Foggy Day*” by Middleton (see sections 6.3.7.3 *test melody 3e*, 6.3.3.2 and 6.3.5).

The cognitive relevance of successive and stepwise pitch change between groups of tones is corroborated by e.g. Övergaard's analyses, which frequently require global melodic direction and global pitch register change for the assigned similarity between phrases to be recognized. This also involves the significance of pitch at beat positions.

The contextual relativity regarding structural significance of different levels of pitch similarity is frequently demonstrated in the provided evaluation material. The very general level of pitch similarity that is structurally significant in e.g. the *Noraffjells* analysis (as performed by the current model and Levy) would not lead to any segmentation at all of the test melody 8 from experiment 2, since all possible segments exhibit pitch set similarity at the level which is determinative in the *Noraffjells* structure.

The exclusion of tonality relationships, such as functional harmony, as a means of pitch structure which is due to the general claim of the model to be style-independent and the assumption of phrase structure being of primary significance for the conception of tonality relationships, have, in general, not affected the results of the structural analysis. However, melodies whose structures rely entirely on harmonic accompaniment or harmonic preconception will, like melodies in which the structural cues are provided by polyphonic structure, rhythmical accompaniment, dance meter etc., surely be misinterpreted by the current model.

### **Means of melodic structure – rhythmic structure**

In parallel to pitch structure pattern of duration changes has proved to be more structurally significant than patterns of absolute duration. It is generally demonstrated by the low significance of tempo changes for the categorization of similarity between rhythmic structures.

The assumed limitations of categorical perception of rhythmic structures can neither confirmed nor rejected through the evaluation, since no melodies where successive rhythmical relationships with higher complexity than 3:4-4:3 are structurally significant, are used in the evaluation.

The assumed categorization of rhythmic change in relation to beat structure at tactus level is strongly supported by the results. In all repertoires, except the *moto perpetuo* style melodies, rhythmic similarity in terms of division, non-division and suspension of beats (tied beats) have been essential to reveal similarities, which are designated through notations and listeners. (see e.g. the *Raag Suddha Danyasi* 6.3.6.1, *Alfanska Ravenica* 6.3.6.1 and the *test melody 4*, *experiment 2*, section 6.3.7.3).

Global rhythmic changes, considering e.g. change of rhythmic density between groups of beats have also prove to be significant (see e.g. the *Fugue* example section 6.3.8.2).

### **Structural significance of rhythm and pitch structure**

The basic assumption of similarity and change that involve both pitch and rhythm being more prominent, than similarity regarding only one of the two dimensions, is supported by the general success of the model.

The relative structural significance of rhythm and pitch as competing structural determinants is assumed to be greater for rhythm at local structural levels and pitch at global structural levels. The theoretical consideration behind this assumption is that it is possible to discretely address more unique temporal positions by pitch than by duration relationships,

while onset information – tone or no tone – is the fundamental input to musical conception locally (see section 3.1). It is difficult to assess the general significance of this assumption from the evaluation of the method of phrase and section analysis, since no specific test of this assumption has been performed. However, the result of the segmentation of e.g. *Raag Suddha Danyasi* with regards to the listener segmentations in experiment 3h and the notated structure, would drop by around 30% if the depreciation of longer rhythmical sequences is excluded from the model.

The cognitive relevance of this assumption is also indicated by the evaluation of the metrical analysis, which demonstrates that an onset-based algorithm is generally sufficient for determining metrical structure at beat-atom level, while the influence of pitch structure must be considered in complex metrical analysis.

The consequence of the assumed relative weight of rhythm and pitch structure in relation to the length of the structure can be illustrated by the following example (see fig. 164). This example shows two distorted versions of the finale theme of Beethoven's *Symphony No.9* (Op.125). In the first example a nine beat rhythmical sequence is applied to the pitch structure, while in the second example a three beat rhythmical sequence is applied. The argument here put forward is that the longer rhythmical sequence will not have the same influence on the conception of the melody as the shorter sequence, since the shorter sequence can be conceived as a consistent metrical grouping of shorter duration than the metrical grouping implied by the pitch structure. The output of the computer implementation of the model for the two examples demonstrates the significance of this assumption:

*A Nine beat sequence*



*B Three beat sequence*



Figure 164. Two rhythmically distorted versions of the beginning of the finale theme of Beethoven's *Symphony No.9* (Op.125). Version a exhibits a nine beat rhythmical sequence, while b exhibits a three beat rhythmical sequence.

The results of the application of the model demonstrate that the different relative influence of pitch structure and rhythmic sequence are related to the length of the sequence. While the nine beat sequence does not seriously alter the phrase structure implied by pitch sequences, the analysis of the three beat sequence version results in an ambiguous and inconsistent analysis by the model.

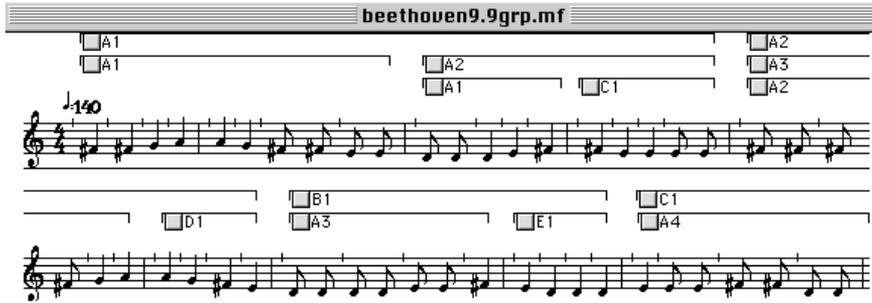


Figure 165. Computer model output of analysis of the nine beat sequence version a.

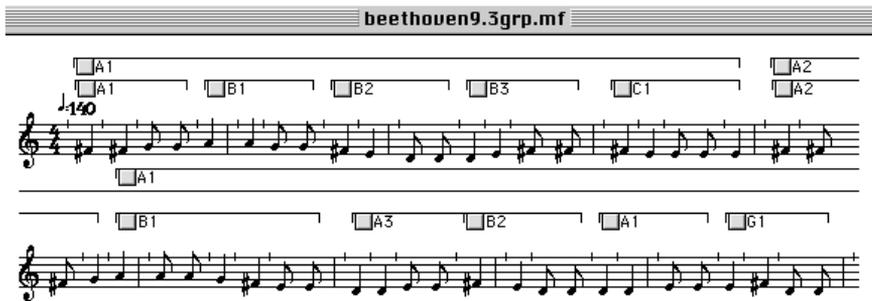


Figure 166. Computer model output of analysis of the three beat sequence version b.

This melody is probably recognizable in spite of the distortion; thus listeners who are very familiar with the theme will be able to categorize with regards to its pitch structure, even when presented with the three-beat sequence version. I argue however, that the level of distortion of the theme will probably result in a more ambiguous interpretation of sub-phrase levels.

### General assumption of the significance of grouping principles in cognition of phrase structure

The significance of regarding grouping-by-similarity and grouping-by-discontinuity as separate and interacting views of grouping is strongly supported by the findings. (cf. Cambouropoulos 1998:82) Grouping-by-similarity, manifested above all in the sequence analysis process, has prove to be essential for the segmentation of a large part of the tested material (see e.g. *Comparisons Folk song*, section 6.3.8.2). Likewise, grouping-by-discontinuity, manifested primarily through the local phrase boundary analysis, is essential for the segmentation of melodies whose structure is not built on sequencing.

The assumption of the relative structural significance of the two different principles, that similarity is a more influential structural factor than discontinuity, while discontinuity determines the recognition of similarity, is commented upon above.

This assumption is based on the theoretical concept that it is possible to discretely address more temporal events by similarity than by discontinuity, since the latter structural perspective concerns the boundaries while the former concerns the grouping.

The cognitive reality of this assumption is demonstrated frequently within the material, by the inclusion of structural discontinuities within sections and sequences. (See e.g. the section level analysis of “*Jesu Bleibet meine Freude*” or “*Fascinating rhythm*”, where the listeners’ notated conceptions of section structure either interfere or encompasses with major structural breaks). And since this assumption is the fundament of the evaluation of sequence structure, the results of the application of the model to the tested material in general support this assumption.

The general assumed relationship between primary, secondary and tertiary grouping principles is also confirmed by the results. Sequence implication by good continuation, periodicity and symmetry is dependent on primary grouping as input. The application of general symmetry or metricity to a structure without the relationship to structural cues inferred by the application of primary grouping principles would lead to segmentations inconsistent with either music analysts’ or listeners’ notated conceptions in all melodies which exhibit asymmetry at any level, hence the majority of the provided examples.

The role of the principles of articulation and structural pregnancy/integrity, as factors involved in the selection of structures rather than in the formation of structures, is evident from theoretical considerations.

### **The significance of secondary grouping principles**

The relevance of the primary grouping principles needs no further comment, since the results of the application of the model are apparently dependent on these principles.

The relative influence of the secondary grouping principles is, however, not as evident. Symmetry is primarily involved in two aspects of the method of analysis; the hierarchical evaluation of sequences and segments, and in the implication of sequences and segments by hierarchical and successive symmetry.

The relevance of these principles is evident from many of the given examples, such as the lower levels of *Polska* no. 83 (see section 6.2.4.5), which are identified by symmetrical implication (see also “*A Foggy Day*” section 6.3.5, Finale theme from Beethoven’s 9<sup>th</sup> *Symphony* section 6.3.8.2 Discussion).

The main influences of symmetry on the results of the analysis can actually be tested, since both the symmetrical implication module and the hierarchical segment evaluation can be turned off in the model. When the symmetrical implication module is turned off, the results for “*Svenska Låtar Jämtland*” (see section 6.3.3.2) drops with 18.2 % with respect to identified segments relating to Övergaard. This difference is considered significant, since it means a reduction of the result with about 1/5. Even more important, it reduces the number of identified measures to an even greater extent, since symmetrical implication is a more forceful factor for determining lower level segments than higher-level segments.

When the hierarchical analysis implying symmetry is turned off, the result for this repertoire drops even more, resulting in a total fit vs. Övergaard of 55%, which is just above chance level. This can be explained by the relative ambiguity of the discontinuity values assigned to the segments, which have strong influence of the structural prominence value.

The influence of these principles in different styles varies. However, these examples indicate that for some repertoires the influence of symmetry by segment implication and selection is crucial for the analysis of the melodic structure.

Regarding the special case of good continuation, which can be expressed as metrical implication, i.e. implication of segmentation based on periodicity of segmentation, the influence on the results for the Svenska Låtar, material is not as prominent, not resulting in any significant drop in the results (< 1%). This is mainly due to that such relationships overlap with symmetrical relationships determined by the sequence analysis.

The repertoire, in which segmental implication by metrical consistency is most influential in the current model, is the South Indian Ragas. These generally extensive compositions, are not characterized by a multi-level, hierarchic sequence structure as in many European melodies, but rather by a continuously changing grouping at primary level, in relation to constant periodicity at phrase/measure level. The periodicity of groupings at measure level thus becomes a very important factor when the structure gradually becomes more and more complex in the development of these compositions. (See e.g. *Raag Abbogi*, section 6.2.3.3 T-sequences, in which the periodicity regarding metrical sub-grouping is a crucial factor for determining the phrase structure). This implication is essentially equivalent to the complex metrical analysis (see Chapter 5, section 5.3.3.6 etc.).

It is important to consider that the significance of the different grouping principles cannot be assessed from their influence on the results by the application of the current computer implementation alone. It might very well be that their significance relates to deficiencies regarding the implementation, that certain principles of primary grouping are not included in the model etc. However, the general significance of principles of symmetry and periodicity in human cognition, e.g. in the cognition of visual objects, supports the claim that those principles are relevant also regarding melody cognition.

### 6.3.9.3 Concluding remarks and future development

As argued above, the results of the evaluation generally support the presented general model of melody cognition and the basic claim that cognition of melodic structure is basically related to and inferred from structural properties in melodies and not primarily imposed on a sounding structure without relation to the structure or arbitrary in relation to the structure. The evidence indicates that a great number of melodies of different cultural and stylistic origin contain structural cues at the level of relative duration and relative pitch, which are sufficient for a syntactically coherent segmentation of a melody to emerge. Thus, this study supports the significance of these dimensions as essential in melody cognition.

Hence, the results also support the claim that more style-dependent structural dimensions, such as conception of tonality (including harmony and other aspects of tonal relationships) is not a prerequisite for melody cognition in a great number of melodies of different cultural origin. Rather, the results point to the sub-structuring of a melody as a spontaneous and subconscious process where the structural cues are essentially common to humans. Saying this, one has to consider that the concept of melody is not relevant in all musical cultures and the possibilities of assigning different culturally determined weights to different structural cues are great.

This leads to the conclusion that a rule-based system may be developed along the lines of the current model, which can be successful in modeling human cognition of melodies and to study cultural and style-dependent differences in melodic structure.

The approach of the model, which regards structural difference and similarity as fundamentally relative, and bottom-up and top-down processes as recurring and interrelated must be considered generally successful.

There are, however, a number of limitations to the interpretation of the results which must be taken under consideration:

Most importantly, it will be no problem to find many melodies for which the current implementation of the model produces irrelevant interpretations in relation to listener conceptions of melodic structure. It may be, that the model will produce irrelevant results for a whole corpus of melodies. Most of the interpretations, which are clearly inconsistent with manual segmentations that have been encountered in the evaluation process (see section 6.3.3), can be explained by deficiencies in the implementation of the model. This implementation should not be regarded as a commercial software package, with extensive beta testing before the release. Rather it can be regarded as a prototype, which needs to be optimized to be useful as an analytical tool. Moreover it must be stressed that it has been developed in order to test the assumption of the model.

But, if misinterpretations due to implementation errors and deficiencies are omitted, there are still examples of melodies within the material, for which the model does not come up with a segmentation that is consistent with listener evaluation. The examples of such melodies within the tested material can be explained by either lack of evident structural cues or ambiguous structural cues. Interestingly enough, a very popular melody such as Ravel's *Bolero* seems to confuse both the computer implementation as well as listeners given the task to describe the sub-structure of the melody.

As previously mentioned, there is thus obviously a limit to which a structure can be interpreted from the internal structure alone. In vocal music, lyrics may easily serve as a more forceful structural cue than pitch and rhythmic structure. In dance music or ritual music, extra-musical structural aspects, such as body movement or ceremony structure, may be the determinative factors governing cognition of melody. And most commonly, in ensemble playing, the most important structural cues for the cognition of the melody may be only inherent in the accompaniment.

The preconception of melody structure, which is assumed in the general model, is for obvious reasons not included in the implementation of the model - it does not involve any machine learning across melodies, only within the melody. Thus, the computer model, so to say, listens to each melody in isolation as if it were the only melody. It is obvious that this is not the case with humans. Thus, there are certainly melodies which will be cognized significantly differently by listeners familiar with the common structures of melodies within the culture and listeners foreign to the culturally-defined style. Such melody structures would be suspected to appear more frequently in isolated cultures, developed during a long time as well as in small sub-cultures where important cues can be assessed by extra-musical means.

However, in the evaluation process, which has involved more than 200 melodies of different origin, I have not yet encountered one melody for which the failures of the interpretation can be explicitly related to the need for style-dependent codes to produce an interpretation reasonably consistent with the culturally determined interpretation. That is not

## *Chapter 6*

say that it does not exist, but that it is very common that the structural information needed for the cognition of a melody is either inherent or given by other musical or extra-musical means.

Thus, by the interpretation of the results of the evaluation of the current model, melody emerge as a 'language' we all understand, but in our own individual way.

The model put forward here obviously require more thorough testing in order to claim to hold for other melody styles than the ones included in the evaluation.

The implementation needs, as has been frequently mentioned, further development in many respects:

- The influence of different structural factors used in the analysis process should be more easily accessible in the output of the program in order to make it useful as an analytical tool.
- Structural ambiguity should be more consistently handled in the output of the computer program through the output of different possible syntactic melodic structures for one melody and the output of total quantifications of structural prominence for different alternative solutions.
- Redundancies in the implementation should be omitted
- The implementation should be more consistent with the model and, in order to optimize results, certain processes which now run sequentially, should be run in parallel.



# 7 Non-metrical Figure and Phrase Analysis

---

## 7.1 What is non-metrical structure?

### General concepts

*“Als sich allmählich die Melodien von den Körperbewegungen loslösten und nur um ihrer selbst willen vorgetragen wurden, lockerte sich auch die tanzmässige Straffheit des ursprünglich knappen Rhythmus. Der Rhythmus der Melodie konnte sich nun an den Rhythmus des ihr untergelegten Textes anpassen, ihre einzelne Töne konnten von den Vortragenden empatisch gedehnt werden”*  
(Bartók, Bela 1920:11)

### 7.1.1 Outline

The model of analysis of non-metrical structure will be described more briefly than the analysis of metrical structure chiefly because many of the important concepts used here have already been presented in the previous sections. Another reason is that the model is not evaluated as thoroughly as the structural analysis of metrical melodies. I will present the model mainly through a few handpicked examples and without any listener evaluation of the model. A third reason regards space. A presentation of the same size as the metrical analysis would outgrow the current study.

The disposition of this section replicates the previous chapters, starting with a discussion of the fundamental assumptions of the method, followed by a presentation of the method of analysis and finally some evaluation examples.

### 7.1.2 Non-metrical and quasi-metrical structure

Non-metrical rhythmic structure has generally received much less attention in the literature than rhythm in metrical music. In their influential book “The Rhythmic Structure of Music” Cooper and Meyer merely state that rhythm can exist independently of meter, but do not present any examples of this type of rhythmic structure.

*“First, rhythm can exist without there being a regular meter, as it does in the case of Gregorian chant or recitativo secco. That is, unaccented notes may be grouped in relation to an accented one without there being regularly recurring accents measuring metric units of equal duration.”* (Cooper and Meyer 1960:6)

Lerdahl and Jackendoff make a similar remark in “A Generative Theory of Tonal Music” (1983:18), also without going deeper into the subject;

*“Before proceeding, we should note that the principles of grouping structure are more universal than those of metrical structure. In fact, though all music groups into units of various kinds, some music does not have metrical structure at all, in the specific sense that the listener is unable to extrapolate it from the musical signal a hierarchy of beats. Examples that come immediately to mind are Gregorian chant, the alap (opening section) of a North Indian raga, and much contemporary music (regardless of whether the notation is “spatial” or conventional). (Lerdahl & Jackendoff 1983:18)*

Still other writers seem to question if rhythm does exist without meter. For instance, Gabriellson (1984:251), in a discussion of rhythm and non-rhythm, includes regularity in the definition of rhythm:

*“The rhythm experience is characterized by perceived grouping; the sounds in the sound sequence/music are not perceived as separate elements but form groups of sounds, usually called «rhythms» or «melodies». Furthermore, one or more of the sounds/tones in the sequence is perceived as accented (stressed) in relation to the others. There is also some kind of perceived regularity, for instance, regular occurrence of accents and a regular underlying pulse rate. Presumably there is always some kind of perceived motion as well as emotional characteristics [...] Non-rhythm would thus be characterized by the opposites to conditions. If there is no perceived grouping, accentuation, and regularity, you usually don't experience any rhythm, even if the sound sequence is within the perceptual present. [...] Of course, there are borderline cases between rhythm and non-rhythm in music. Gregorian chant may be an example, at least in certain parts, as well as many pieces in the contemporary Western art music.” Gabriellson (1984:251)*

As is evident from this citation, a definition of rhythm that excludes non-metrical structure calls for a wider concept, which includes grouping structure and significant temporal relationships in music with and without perceived pulse, which would make rhythm a special case of this broader category. To my knowledge, no such concept of temporal organization has gained widespread acceptance and rather than developing such a concept I am using rhythm in the sense of perceived grouping structure and perceived temporal relationships in music generally, in non-metrical and metrical contexts respectively. A popular term often used in relation to non-metrical structure is “free rhythm”.

Maybe the most obvious example of non-metrical rhythm is speech, which usually is not perceived metrically, but for the conceived structure of which perception of grouping and temporal relationships is highly significant.

I usually present the concept of non-metrical structure to students by asking them if they are aware of any music in which it would feel unnatural and forced to tap their feet along with the music – implying that the music does not evoke a conception of meter. In the discussion that follows it is common that a distinction is made between music in which no metrical conception arises at all and music to which you can tap your feet along if you are asked to, but that it would feel forced or unnatural to do so.

Musical styles that are identified as belonging to the former category are e.g. herding calls, some religious ritual music like prayer recitations, certain psalmody and religious chant, laments, recitatives, introductory sections such as the *alap* of the raga or the *taqsimi* in oriental

music, much contemporary art music and certain avant-garde jazz music.<sup>121</sup> This category can be labelled true non-metrical structure.

The second category is often identified throughout these discussions and gives sometimes rise to some disagreement among the participants. Styles and genres frequently mentioned in relation to this category are different kinds of lyrical folk songs, folk hymns and ‘lyrical’ instrumental music. Instrumental music of the era of romanticism is often mentioned in this context. The question arises as to whether these examples really are non-metrical – do not have any perceivable pulse – or if they are interpreted freely in relation to the pulse. A typical standpoint can be “- OK, I can force myself to tap along with the music, but it has never come to my mind and I cannot really hear it – is it really there? – while the other position might be “Can’t you hear that it is really just a flexible pulse?”. What becomes obvious in these discussions is that people can have different conceptions in this matter when relating to the same musical performance. “*Parlando*” and “*Rubato*” expressions have been used by scholars such as Bela Bartók (Bartók 1920:11) to designate such structures and will here be labelled quasi-metrical structure.

The relationships to call, chant and speech are evident in most of the above examples and it is obvious that non-metrical structure is especially frequent in styles with a strong relationship to vocal expression.



Figure 7-1. Example of non-metrical structure assigned by transcriber. Herding call after the performance of Karin Edvardsson, Transtrand, Sweden, compiled from original transcription by Jules de Vries (Moberg 1959:49)

---

<sup>121</sup> That avant-garde jazz is commonly categorized as non-metrical by the students, may rather reflect its structure generally being difficult to conceive than the relative frequency of non-metrical structures in avant-garde jazz.

**220. P S A L M**  
Gamla Psalmb. nr 381

Hem-lig stod jag en mor-gon Up-på en lön-lig stad,  
Med nycken gråt och vån-den Hör-de jag hu-ru bad Dö-den en ung-er man,  
Den ho-nom ha - de gri - pit Och bun - dit med starkt band.

\* Närmast *ess.*

Figure 7-2. Quasi-metrical structure a) Folk hymn variant from Sweden after the performance of Pers Karin Andersdotter, Mora, Sweden notated by Nils Andersson (Andersson 1922:130-131). The quasi metricity is here indicated by regular grouping into beats, while no time signature or bar lines appears – double bar lines indicate phrases. The aperiodicity at designated pulse level is indicated by fermatas.

**No. 1/b Song**

Parlando Quasi giusto  $\text{♩} = 76$  1) 2) simile 3) 4) 5) 6)

Cin ie-ram in vré-mea mea mă, Cin ie-ram in  
vré-mea me mă, Cin ie-ram in vré-mea me mă.

Figure 7-3. Quasi-metrical structure b) Folk song from Rumania, after the performance of Susana Crăciunescu, Hunedoara, Rumania, transcribed by M. Rabinovici (Dragoi 1959:19). Quasi-metricity indicated by tempo designation, by bar lines and beat grouping. No time signature and fermata indicates deviation from perfect regularity.

As is exemplified in the above, the boundaries between metrical structure and non-metrical structure are sometimes impossible to draw based on structural considerations only, since people have different tendencies to conceive metrical and non-metrical to the same musical stimuli. This tendency may be much influenced by individual musical experiences, thus culturally related. And even if I have not made a statistical study of the subject, I have noticed a clear difference between the singers in my groups, who are much more phrase-oriented (non metrical conception), and instrumentalists who are more beat oriented (metrical conception). It would be quite expected that singers have a natural focus on lyrics while instrumentalists, who mainly are occupied with metrical music, have a different tendency in their conceptions. Thus, when forming a model of analysis of structure of non-metrical music the floating boundary between non-metrical structure and metrical structure typical of the quasi-metrical category have to be considered. Further, it is not uncommon that there are sections of local metricity within a mainly non-metrical structure and vice versa. (see e.g. Moberg 1959:43, Temperley 2001:301). The figure below illustrates this relationship.

Fig. 4 illustrates the postulated relationships between metrical and non-metrical rhythmic structure. There are basically two categories, metrical and non-metrical, indicated by more shaded areas for metrical structure. The quasi-metrical sub-category has features of metricity but with such discrepancies regarding periodicity that it can be conceived both metrically and non-metrically. Within the categories there can be sections of the opposite structure, indicated by the mixture parameter, illustrated by boxes of opposite shading.

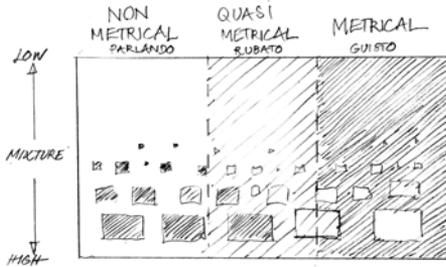


Figure 7-4. Schematic overview over the assumed cognitive categorization of non-metrical vs. metrical structure (for explanation, see above)

Even if the degree to which a melody structure can be said to exhibit structural metricity with possible perceptual significance cannot be described categorically, the conception of meter in the sense of a regulative periodicity is categorical: A general metrical grid is either conceived or not. A model aiming at predicting a possible conception of melodic structure must thus be categorical in the sense that it determines whether meter exists or not, involving the degree to which it influences the structure.

In the current implementation of the model, this categorization is determined by the metrical quantization and beat-atom analysis (see chapter 5, sections 5.2.2 and 5.2.3). If these two modules fail to find a common metrical division at beat-atom level, the analysis continues to the non-metrical analysis module.

The above reasoning can be summarized into two assumptions:

- (1) **There is no clear boundary between metrical and non-metrical rhythmic structure from a phenomenal point of view. The same structure may be conceived metrically and non-metrically by different people. There can also be a mixture between metrical and non-metrical sections within the same melody.**
- (2) **There is a categorical difference between metrical and non-metrical rhythmic structure, regarding the conception of meter being a regulative factor.**

### 7.1.2.1 Determining temporal relationships in a non-metrical context

The analysis of melody structure in non-metrical and quasi-metrical melodies is both more simple and more complex than analysis of melody with metrical structure:

Since grouping by meter generally can be disregarded and impulse quality can be suppressed, end-oriented grouping will be dominant on a global level. Factors such as good continuation in terms of periodicity, symmetrical implication, hierarchical symmetry is less influential since periods of time are not really measurable at phrase level. The level of

structural hierarchy thus can be expected to be lower when no symmetrical temporal divisions are conceived and there is no global temporal grid, which can be used to evoke expectations of temporal relationships.

But if there is no meter, how are temporal relationships conceived?

The basic assumption in the current study is that the fundament of musical rhythm is the perception of categorical pair-wise relationships between successive musical events within the perceptual present (c.f. Chapter 3, section 3.1). The cognitive reality of such categorical relationships have been demonstrated from different perspectives in the work of Fraisse and others from the point of view of perception of rhythmic relationships (e.g. Fraisse 1956, 1982) and by Gabriellson and others from the point of view of performance of rhythm (e.g. Bengtsson & Gabriellson 1980, 1983).

In this context, the number of simultaneous rhythmic relationships within a perceptual present, are assumed to be limited to the categorical present defined as  $7 \pm 2$  simultaneous categories (Miller 1956). The categorization is thus based on an estimation of successive relationships between events, the fundament of which is assumed to be the categorization of event durations into the equal or unequal durations (see further below). This implies two principal differences between non-metrical and metrical rhythm: One is the consistency of the categories, which cannot be measured in non-metrical context because there is no consistent time-grid, hence allowing for greater variation of the actual duration of categories. The second is the categorization into duration categories, which, in a metrical structure, can be based on the relationship to meter, while, in a non-metrical context, it must be based successive relationships between adjacent events.

Further, I assume that the basic means of categorization that have been demonstrated to apply for metrical contexts, is valid also in non-metrical contexts, but with less precision. This implies that categorization in duple and triple relationships would be valid also in non-metrical categorization of durations, but that complex rhythmic relationships such as complex syncopation which relates to pulse and the higher rhythmical 4:3 / 3:4 relationships which are allowed in metrical contexts are not assumed to significant in non-metrical structure. (see further below under quantization).

Because of the lack of a general, referential time-grid I assume that the assumption of categorical conception of symmetrical and hierarchical proportion, hence the categorical proportion rule, (see Chapter 3, Section 3.1 and Chapter 6, Section 6.2.4.3) applies for non-metrical rhythmical relationships. This rule states that the conception of proportions between atomic temporal events<sup>122</sup> is categorical and limited to a maximum of four categories; equal duration, duple division, triple division and quadruple division. This implies that two successive events are regarded equal with regards to duration when they are more similar than dissimilar, which is interpreted as being closer to 1:1 relationship than 1:2, i.e. differ with less than 1/3 of the longest duration. The application of this rule in the model of metrical phrase and section analysis, within symmetrical implication and hierarchical evaluation of segments, produced results that were confirmed by the evaluation of the results.

There is evidence that our discrimination of durational differences relates to absolute duration of events, suggesting that the discrimination is best is when the time span is around 600 ms (Fraisse 1967). This evidence further indicates that existence of meter makes

---

<sup>122</sup> That is, not being conceived as composites of a common divisor

discrimination of duration differences considerably better than for time discrimination without a metrical context. Taken together this evidence implies that discrimination of duration categories in non-metrical contexts must be considerably more liberal than regarding metrical contexts and that the peak of pulse salience around 600-700 ms (Parncutt 1994) should be included also in the categorization of temporal relationships in non-metrical rhythmical contexts.

The categorical proportion rule including levels of subdivision is therefore assumed to be applicable mainly to durations and groups of durations within the span of typical pulse salience, which implies that duration categorization has to be based on the identification of a central duration category, with brief super- and sub-categories. From the assumptions of the maximum number of simultaneous categories and the categorical proportion rule it seems reasonable to assume that the categorical conditions of a central pulse level (see chapter 3, *General Model*, section 3.1.) also applies for the standard/referential category in non-metrical structure. This assumption states that a referential/standard category should ideally not have more than three super-categories and three sub-categories. Events with durations below 1/4 of the central category or with durations below 150 ms in a context of durations above 600 ms or 100 ms in a context with maximum duration below 600 ms are regarded ornaments or articulative rests without rhythmical significance in non-metrical contexts. (The higher limit of 150 ms in relation to the threshold value of 100 ms used in the metrical analysis is related to the general assumption that duration discrimination ability is assumed to be less precise than in a metrical context, for which there is evidence for a lower limit of 100 ms, see London 2002:535).

It should be emphasized that the threshold values presented here are not considered absolute in any sense. Probably most people will not be able to simultaneously handle as much as seven duration categories in their mind throughout a non-metrical melody, maybe not even within a phrase. These values should thus rather be regarded as biases, necessary for the formal implementation of the model.

But this implies further that the level to which the analysis can predict a conception of structure depends highly on structural contrast; While sections of low structural contrast or ambiguous structural indication can be interpreted in metrical melodies by metrical, symmetrical or/and hierarchical implication the ambiguity of a structure in a non-metrical melody will depend (almost) entirely on the level of local structural contrast.

This further implies that principles of good continuation and symmetry, such as periodicity, successive and hierarchical symmetry can be regarded as ternary grouping principles in non-metrical structure at least at higher levels of grouping (see *Assumption* no 14 below), that is not being of implicative strength but rather of significance for the structural prominence of different grouping indication.

The lack of a general temporal measurement also makes the contribution to the analysis of analogy with similar structures, which is a core element in the metrical analysis, harder to estimate, since temporal implications are missing and the conception of rhythmic structure is primarily local. This means that similarity between structures becomes more obscure, since we have no means of estimating the temporal similarity. Further, the impulse or structural accent implied by similarity between contiguous sequences of events will be less prominent in a non-metrical structure than in a metrical where two contiguous series of similar events creates a temporal unity. This implies that similarity is a less prominent structural factor in the conception of non-metrical melodies.

This implies in turn that grouping within a non-metrical rhythmic structure will be conceived predominantly end-oriented (see chapter 6, section 6.1.2), which means that group boundaries will be perceived primarily at points of greatest discontinuity or structural distance.

To summarize:

- (3) Temporal relationships between musical events in a non-metrical rhythmic structure are assumed to be perceived primarily locally, based on relative and categorical perception of pair-wise relationships between durations, but influenced by global categorization of standard durations (see 6).
- (4) The categorization of durations in non-metrical contexts is assumed to be limited to the perceptual present, typically extending up to 5-6'', and the categorical present, typically allowing for  $7 \pm 2$  simultaneous categories.
- (5) The basic categorization of durations is assumed to be governed by the categorical proportion rule, which states that conception of proportions between atomic units in time ranges is limited to equal, duple, triple and quadruple division.
- (6) The global categorization of durations is also assumed to be influenced by the absolute duration of events and event groups, favoring the conception of a referential standard duration category, which is ideally concurrent with salient pulse duration and ideally in the centre of the category spectrum.
- (7) Complex rhythmic relationships based on the relationship to a meter, such as syncopation, and complex rhythmic relationships which require metrical subdivision are not assumed to be applicable in a non-metrical rhythmical conception.
- (8) The lack of a general temporal grid makes the conception of structure in non-metrical contexts highly dependent on structural contrast. The conception of grouping structure in non-metrical contexts with low structural contrast is thus assumed to be highly ambiguous.
- (9) The lack of a general temporal grid, influences the strengths of the factors determining melodic structure. The primary grouping principle of similarity, specifically regarding grouping by sequencing, is assumed to be generally subordinate to discontinuity. The secondary grouping principles of good continuation regarding periodicity, successive and hierarchical symmetry are assumed to be less influential than in metrical contexts. Thus are these grouping principles have been dethroned to ternary grouping principles with regards to non-metrical structure on higher levels, i.e. not being of implicative strength.

### 7.1.2.2 Levels of non-metrical grouping structure

The levels of grouping structure, which have been identified in melodies with metrical structure, includes primary grouping at beat level, sub-phrase and phrase grouping at measure level and above and super-phrase and section levels which include a substructure of phrase levels. The relationship to metrical units, which as have been demonstrated in chapter 6 are

defined to a large extent by grouping structure, makes the stratification of different levels in metrical music relatively evident<sup>123</sup>.

For example, the requirement of a pulse to be recognized within the perceptual present puts both category size and temporal restrictions on the primary grouping level. These limitations are also the basis of the sub-phrase and phrase-levels that correspond to measure levels.

But, since the basic conditions of the psychological and categorical present also apply also for non-metrical structures, these levels can be applied also to non-metrical contexts.

In a study of Swedish herding calls, the Swedish musicologist Carl-Allan Moberg tried to describe the structural levels of melodies without metrical structure (Moberg 1959:10-57). He identifies the phrase level as the most important and below that he groups shorter notes together with longer notes, the latter being designated "skeleton notes" (Swedish: *skelettoner*) adopting the term *Gerüsttöne* from the German musicologist Wiora (Wiora 1941). This can be regarded as the primary grouping level. Moberg questions whether higher grouping levels are significant within the repertoire but does however present form schemata, which implies the grouping of phrases into higher-level super-phrases. (Moberg 1959:42-57).

The phrase level is stressed by e.g. Moberg (1959) and others (c.f. Rosenberg 1986) as the central structural level of non-metrical melodies. It could be argued, in the line of Moberg, that higher level grouping, such as super-phrase levels, becomes cognitively more difficult to achieve in non-metrical melodies than in metrical music because the lack of a general time grid makes it difficult to measure the strength or weight of a structural indication of one phrase to another. This limits the possibility of a structural hierarchy of phrases. When we have limited possibilities of estimating which phrase is most prominently marked by e.g. temporal distance, and the possibility for clustering of phrases by similarity is limited because temporal similarity is not measurable, how can an elaborate hierarchy of structural prominence of phrases be conceived?

How the central phrase level is determined does not seem to be an issue to Moberg (c.f. Johnsson 1986). It is just taken as self-evident from the musical structure. Why is that? If we return to the melody example given in fig. 1 we can see that bar lines are inserted at points of great temporal distance and after rests, that is, at points of global discontinuity. However, some points of great temporal distance like some points of rests, are disregarded by Moberg, for reasons that are unclear from his analysis (Moberg 1959:37-49). In vocal music with lyrics, the segmentation at a central phrase level is often solved by using the structure of the lyrics as a guideline, as in the examples of songs provided above.

The purpose of this study excludes use of such extra-musical guidelines for segmentation; the means of extrapolating a phrase structure at central phrase level must be obtained by music structural means only. This requires not only consideration of temporal, onset and pitch discontinuities, but also conditions regarding the extensions of a melody phrase. I have not been able to find any comparative studies of phrase length in non-metrical melodies but, from some studies of Scandinavian folk singing focusing on quasi metrical melodies, it can be drawn that phrases determined by lyrics in these songs generally last between 4 and 10 seconds with a peak around 6-7 seconds (Rosenberg 1986, Ledang 1966, Jersild & Åkesson 2000). Sundberg (1980:29) claims that while phrases in the sense of time spans between

---

<sup>123</sup> Note however, the floating boundaries between the different levels that are demonstrated in chapter 6, especially regarding higher structural levels

breathing is usually around 5 seconds in speech, usually while phrases in song can often last for more than 15 seconds. However, musical phrases do not have to concur with breathing points.

The size of phrase length measured in Scandinavian folk song and what is normal in speech seem to fit in well with the estimations of the perceptual present, the range of which is generally estimated to between 3 and 8 seconds (Clarke 1999:476). Thus, I assume that a typical phrase length in non-metrical music would fall into the limits given by the perceptual present, supported by its close relationships to speech structuring, with a typical best fit between 5 and 6 seconds. Indications of phrase considerably shorter than this peak value would according to this assumption be regarded as indications of sub-phrase level and considerably longer phrase time spans between phrase indications will be regarded indication of super-phrase levels. In addition, I have chosen to couple the concept of the categorical present, implying that a central phrase should typically contain  $7 \pm 2$  primary groups. This implies that a phrase that contains e.g. only three notes would be regarded indication of a sub-level phrase.<sup>124</sup>

A similar stratification of a primary grouping level in relation to a central phrase levels, as is made by Moberg referred above is also typically indicated in the notation practice regarding Eastern European folk song originating from Bartók et al:

**Dirge**

Figure 7-5. Rumanian folk song sung by P. Ispas, Hunedoara, Rumania, transcribed by E. Comisel (From *Dragoi* 1959:27). No tempo designation, time signature. Bar lines used as phrase indications.

The primary grouping level is here indicated by means of beaming notes together in groups and by slurring notes that are connected. These means of indicate grouping in non-metrical melodies are extremely common in transcriptions. To beam notes indicates that what

<sup>124</sup> Note that I have not managed to find any empirical studies of durations of phrases in non-metrical music and that there exists examples of phrase lengths which seem contradict these indications, such as e.g. the *Drubhad* song style.

is notated is really beats – but since there is no pulse this could hardly be the case. Rather the groups indicated by beaming in such contexts could be designated *figures* the non-metrical equivalent to *beat-group*. (see Chapter 6.1.5.1)

If we consider the grouping indications at primary level in the notation above, it is evident that the *figures* above are between most articulated events, by long duration and articulation. That e.g. ethnomusicologists and transcribers, to a great extent, bothered to notate primary groups in non-metrical contexts in this manner – as if it was metrical – indicates that they might have conceived figures as related to beats.<sup>125</sup>

Even if this notation practice may be influenced by conventions of notations, there are really very good reasons for doing so, considering the floating boundary between metrical and non-metrical musical structure. If a conception of beats emerges from a certain point of regularity within the structure, where do these beats come from? If we consider the figure level as the primary grouping level in non-metrical contexts – a musical parallel to morpheme – it is logical to assume these primary groups are the fundament of beat identification.

This implies that the same fundamental principles for low level grouping applies to both non-metrical and metrical melodic structure, with the important difference that metrically implicated grouping by symmetry and periodicity does not apply to the same extent for a non-metrical structure.

This is important, since this implies that start-oriented grouping can be assumed to be applicable also for lower level non-metrical grouping, in the sense that groups can be formed by articulation and not just distance on this level.

From the point of view of a floating boundary between non-metrical and metrical contexts, Temperley questions if there is really any non-metrical structure (Temperley 2001:301).

*“However, there are some reasons to suppose that, in cases like figure 11.3, we continue to infer some kind of irregular metrical structure, rather than none at all. First of all, it seems clear that we do not simply “turn off” our metrical processing in hearing such music; some kind of metrical analysis is always in operation. The evidence for this is that, even when a piece has established a norm of being completely nonmetrical, we will immediately notice any suggestion of a beat or regular pulse. This would be difficult to explain if we assumed that no metrical analysis was taking place.”*

Even if Temperley, to my mind, makes a serious mistake in his deduction that a metrical analysis is taking place at the same time as meter is not conceived, hence implying that there can be no other structuring principle, the point that we are in general sensitive to periodicity – probably to a varying degree depending on musical experiences even looking for periodicity – and will be likely to recognize it when it appears points to the relationship between grouping structure in metrical and non-metrical contexts.

- (10) **The grouping levels in non-metrical melody may be divided into primary grouping level, sub-phrase levels, central phrase level and super-phrase/section levels.**
- (11) **Of these, the central phrase level seems to be the most conceptually salient level. One might rather designate non-metrical rhythmical structure as phrase-oriented structure.**

---

<sup>125</sup> Moberg criticizes this manner of transcription in his article but nevertheless use a similar notation in many crucial examples

- (12) The primary grouping level, here designated figure level, is assumed to be corresponding to grouping at beat level in metrical music, while the central phrase level is assumed to be basically confined to the limits of the psychological and categorical present. This makes the central phrase level comparable with e.g. 2-3 measure phrases in metrical music made up from typically  $7 \pm 2$  primary groups, while phrases at measure level in metrical contexts may rather be compared to sub-phrase level within non-metrical structure
- (13) It is here assumed that grouping at primary level gives rise to a sub-structuring in more structurally significant events defined primarily by articulation Gerüstttöne, and structurally less significant events which can be considered as embellishment of the structure made up by the more articulated events.
- (14) Conception of higher phrase levels and sections are assumed to be relevant in non-metrical contexts but to a lesser degree than in metrical contexts, since the means of establishing temporal hierarchical relationships are restricted by the lack of a general metrical grid.
- (15) The same structural principles are assumed to apply for the conception of figure level groups as for primary groups at beat level in metrical contexts, with the exception that the influence of secondary grouping principles such as periodicity and symmetry is less prominent. (see chapter 5, section 5.2.4.3 and chapter 6, section 6.1.7)
- (16) As a consequence of the assumption (1) that there is no clear structural boundary between non-metrical and metrical rhythmic structure, conception of periodicity is assumed to be of possible significance in an essentially non-metrical conception, without being a regulative factor with structural implication. This implies that in a melody that is essentially conceived as phrase-grouped, local periodicity can influence grouping at lower levels.

## 7.2 Method of analysis of non-metrical melodies

### 7.2.1 General design of the model

The non-metrical analysis is in the current model involved when the metrical quantization and beat-atom analysis (see chapter 5, sections 5.2.2 and 5.2.3) fail to find a common metrical division at beat-atom level. This reflects the general assumption that we are likely to recognize and conceive a metrical structure in the sense of regulative and implicative accent structure if the melodic structure exhibits periodicity at lower levels.

It must once again be emphasized that for analytical and practical reasons the computer implementation of the model follows the different steps in the analytical model of structural conception of a melody sequentially, and not in parallel, as is assumed by the general model of melody cognition. The analytical consideration behind this is to be able to study the interaction and function between different parts of the analysis. And, practically, it is easier to evaluate the effects of the different modules by sequential procedure. A future development would be to implement a parallel analysis in order to make the implementation mirror the theoretical model and to get more consistent results.

The first step in the non-metrical analysis is the identification of successive rhythmic and pitch relationships. (*Assumption no 3*). This implies that a quantization of durations has to take place, which basically discriminates equal durations from different durations and implies a categorization of durations into classes within the perceptual present and categorical present. (*Assumption no 3-6*). This is based on an initial quantization based on the category discrimination rule, the determination of a standard category based on ideal figure/pulse period duration and the constraints of the categorical present.

This means that basically the same conditions of a standard category of duration apply for non-metrical contexts as for the central pulse/tactus level in metrical structure, which states that a standard category should ideally be grouping a maximum of three subordinate categories and dividing a maximum of three superordinate levels. (see chapter 5, section 5.2.3). In this context of non-metrical structure the prominence of less grouped levels has higher significance.

From the analysis of a standard category, the successive events of the melody are quantized to conform to the constraints of the duration categorization and significance of rhythmic events (see further section 7.2.2.3)

The second step in the analysis involves a preliminary analysis of central phrase structure. This might seem backwards regarding the general model of melody cognition which states that primary grouping is input to global grouping. However, the general model also states that primary structure can be revised by input from global structure. In the case of non-metrical melodies the assumption regarding non-metrical structural levels (*Assumption no 9*) designates the central phrase level as the main structural level of the melody. This implies that while primary grouping does not primarily influence this level, the identification of a central phrase level can strongly influence the primary grouping at figure level. Hence, it seems practical to perform this analysis before the grouping analysis, even if it means omitting a stage in the assumed cognitive model. The identification of a central standard duration category is,

however, essential for the preliminary analysis of central phrase structure and thus primary grouping is partly involved in the pre-analysis of central phrase level.

The analysis of central phrase level is based on mainly structural discontinuity at global level and can, in short, be said to replicate the analysis of global rhythmic breaks, onset-offset changes and global changes of pitch register for metrical structure (see chapter 5, section 5.2.4.6 *Analysis of Global Discontinuities*) In addition, the conditions of typical phrase duration, typical number of subcategories and good continuation in terms of phrase duration are applied in this analysis. (*Assumption no 10-12*).

The next step in the analysis regards the analysis of the primary grouping level – figure analysis. This involves analysis of local metricity and quasi-metricity within the preliminary phrases at central phrase level, and interpretation of grouping indications by means of local discontinuity and accent structure.

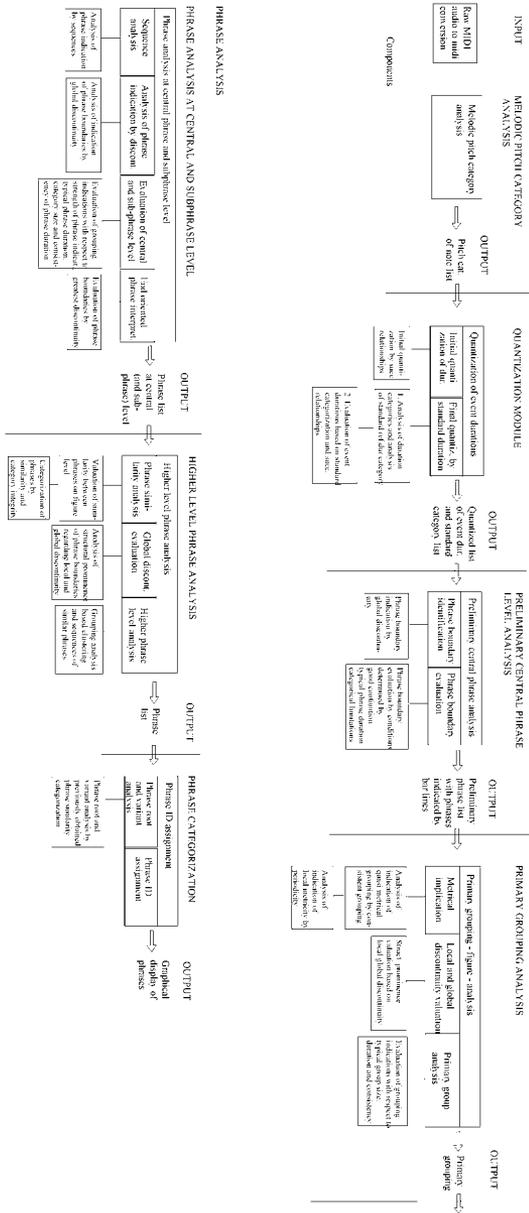
The primary grouping level is fundamentally start-oriented, grouping events before the most articulated event with previous events, i.e. regarding “up-beats” as being part of the preceding group. This is important for the subsequent phrase analysis since it allows for variation with regards to formulation of the figure level, but keeping the integrity of the structure based on most articulated events. This follows the concept of “*Gerüsttöne*”, the non-metrical local structure being structured in articulated and non-articulated events, the former of structural significance. The cognitive reality of this concept is indicated by several ethnomusicologists and also from notation practice and musical terminology (see above section 7.1.2.1).

In the next stage, phrase analysis at central level is once more performed, this time based on grouping of figures (c.f. Chapter 6, section 6.1.5.1 *Pitch Structure*). In this part of the analysis, indications of sequence structure are involved as well as grouping indications by local and global discontinuity, governed by the above restrictions of typical phrase duration, category content and good continuation regarding phrase length (see *Assumption no 8*). This analysis may lead to the identification of both central phrase and sub-phrase levels.

Possible higher phrase levels are identified in the next step of the analysis. This identification is based on both an analysis of melodic similarity between phrases and an analysis of relative strength of global discontinuity indications. Similarity is assumed to influence higher level grouping in basically the same way as for metrical structure - by sequencing of phrases and by clustering of similar phrases into higher levels - but to a lesser degree. Hierarchical categorization of global discontinuity involves the following (1) change of global register change; (2) change of melodic direction (3) hierarchical rating of global rhythmic breaks (interonset and offset-onset change); and (4) change of phrase duration in combination with analysis of local discontinuity. The relative structural significance of grouping indication by similarity and discontinuity is in contrast to metrical structure evaluated simultaneously, reflecting the assumed lesser impact of grouping by similarity.

In the last step of the analysis the obtained phrase structure is named according to the similarity rating of phrases. The outline of the model can be viewed in table 1.

Table 43. Outline of the model of structural analysis of non-metrical melody



Outline of the MODUS implementation of non-metrical structural analysis

## 7.2.2 Non-metrical quantization

### 7.2.2.1 From musical sound to midi

In the following we will examine the different stages in the analysis described briefly above more thoroughly by following the analysis of a typical non-metrical melody. I have chosen a melody from the Swedish style of melodic *berding calls*, traditionally called *kulning*, *kaukning* etc. (see e.g. Moberg 1959, Johansson 1986). The specific example is maybe the most famous among recordings of this repertoire, a *kulning* performed by Karin Edvardsson Johansson (born 1909) from Transtrand, Dalarna, recorded 1948.

Many herding calls may not be designated melodies in the sense of the concept used in this study, i.e. forming a meaningful whole in itself, but this particular *kulning*, improvised from a set of motifs that are common in calls from this specific area, has an elaborate form which makes it reasonable to consider it a melody.

It is also useful in the current context because melodies of this style are typically regarded as non-metrical. The input to the model is a raw audio-to-midi conversion made from the sound of this *kulning*. I have used the software Melodyne 1.5.2<sup>126</sup> to convert it from audio to midi. Though this conversion is essentially automatic, it is not entirely so, since field recordings such as the present example have disturbing noise, reverberation, extra-musical sounds etc. This makes it necessary to edit the conversion with regards to tone identification and discrimination between tone and rest. In the present example, the editions concern pitch discrimination at semitone level<sup>127</sup>, the discrimination between sound and rest and pitch separation of the three shortest notes of the sample, which were not recognized by the algorithm.



Figure 7-6. Excerpt from *Kulning* by Karin Edvardsson Johansson (b. 1909). Transcription by audio-to-midi conversion with pitches assigned at semitone level by the conversion software and with

<sup>126</sup> © Celemony Software GmbH 2000-2002

<sup>127</sup> Melodyne does not export midi files in 24 note per octave format that can be read by Finale

*editions regarding rest and tone discrimination and of three notes of short duration, displayed in Finale 2002 software.*

Because of the format of the implementation of the current model, the midi file created by Melodyne has to be imported to the software Finale<sup>128</sup>, in this case version 2002. A slight quantization is required, since the implementation of the current model does not allow for shorter note values than 1/32. The influence of this quantization can however be minimized by using a high tempo value in the conversion for perceptually significant durations.

When imported into Finale the transcription of Karin Edwards<sup>128</sup>son Johansson's *kulning* looks like in fig. 6 above.

This is then exported as a MIDI file which becomes the input to the current model.

### 7.2.2.2 Analysis of melodic pitch categories (MPC)

The first stage of the analysis performed by the program is to apply the analysis of melodic pitch categories (MPC analysis) to the input file. This part of the analysis is described in chapter 2 and will not be commented upon here. This analysis is essential for the subsequent stages of the analysis of melodic structure, since it determines which pitches belong to the same MPC and which belong to different MPC. Since the MPC level is the structurally determinative, level of pitch change in melodies, analysis of similarity and change regarding melodic contour is fundamentally related to this structuring.

In the figure 7 below the result of the MPC analysis is displayed. The differentiation of pitch categories and variants of the same category is as can be seen in this figure considerably different from the default categorization of pitches in the Finale software.



Figure 7-7. MPC analysis performed on the *Kulning* by Karin Edwards<sup>128</sup>son

---

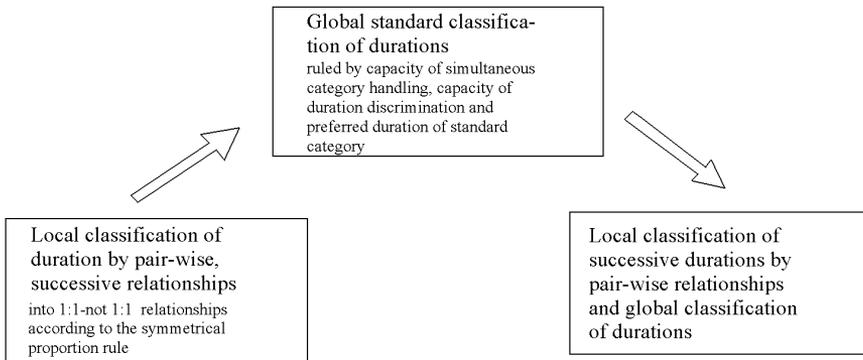
<sup>128</sup> © 1987-2001 by Coda Music Technology, a division of Net4Music Inc.

The cognitive relevance of the MPC analysis is supported by different transcriptions of both the same and variations of the same kulning, such as the one referred in Moberg (1959) (see figure 1 above).

### 7.2.2.3 Categorical quantization of event durations

The categorical quantization of event durations-starts differs from metrical quantization in that it does not assume the existence of a metrical structure at beat or division level with tempo fluctuations. Thus, it is essentially static, reducing the number of different durations by pair-wise successive relationships between event durations based on the assumption of category capacity within the perceptual present and the assumed limitations of categorical symmetrical relationships. (*Assumption no 3-6*).

This implies that the pair-wise relationships between event durations have to be examined by means of a global categorization of durations and that such a categorization must be obtained by the successive evaluation of pair-wise relationships between event durations. This cognitive process must thus involve three stages: (1) Local categorical perception of pair-wise duration relationships between events; (2) the formation of a global classification of durations based on (1); and (3) the classification of event durations based on both the pair-wise relationships and the standard classification of durations.



*Figure 7-8. Schematic model displaying the assumed cognitive process of categorization of temporal relationships between events in non-metrical structure*

The first step of the analysis must thus be the analysis of pair-wise relationships between events in the two classes equal duration – not equal duration. This is obtained by applying the symmetrical proportion rule to pair-wise durations, which states that two events are equal if they differ with less than 1/3 of the longest event based on the geometrical mean between no division and duple division.

By these means, the individual durations of the melody events are grouped into categories. In the current example, melody this initial category list will be:

Chapter 7

Table 44. Initial categorization of durations, expressed as midi note durations with corresponding durations displayed in common notation

Midi duration values	
(128)	
(256)	
(384 512)	
(768 896 1024)	
(1152 1536 1664)	
(1536 1664 1792 2048)	
(2560 3072 3328 3456)	
(4096 5120 5888)	

From this initial categorization, a standard duration categorization is formed by classification of the duration of initial duration categories according to several conditions. Initially, the hierarchical relationships, category integrity and discrimination limit between durations is evaluated. For instance, the two shortest note durations are in this example shorter than 150 ms, which is considered the ornament limit. They differ with 62.5 ms which, according to the current model, is not considered categorically significant in a non-metrical context with duration categories above 600 ms and are, therefore, merged into the same category.<sup>129</sup> Hierarchical category integrity implies that a higher category should be possible to express by means of a lower category, according to the symmetrical proportion rule allowing for duple and triple division of a category<sup>130</sup> Further, the categorization is evaluated based on the assumed limitations of simultaneous category capacity, by merging close and extreme categories when the number of duration categories exceed seven.

This leads, in the current example, to a revised categorization of event durations which expressed in midi click values looks like this:

(128 256)

<sup>129</sup> While JND, just noticeable differences, in isochronous sequences have been demonstrated to be as low as 6 ms for tone interonset intervals shorter than around 240 ms (Friberg 1995), the duration discrimination ability seem to be considerably worse for longer durations and complex musical situations (Clarke 1989, Friberg 1995), where the JND values have been reported to be considerably higher, between 5-20%. This indicates that the significance of duration difference is contextual, which is why the relatively high in-significance value is chosen here. The general assumption that duration differences below 100 ms are insignificant is also supported by research on subjective rhythmization, durational discrimination and studies of cortical processing of musical elements (see London 2002:535).

<sup>130</sup> Quadruple division is regarded insignificant in non-metrical contexts

(384 512)  
(768 896 1024)  
(1152 1536 1664 1792 2048)  
(2560 3072 3328 3456)  
(4096 5120 5888)

This categorization functions as input to the next stage of the analysis which concerns the identification of a central standard category to which higher and lower categories can be related. Essentially, this analysis is based on the assumption that the preferred duration of a central duration category is around 600 ms, which is indicated by studies of e.g. pulse salience (Parncutt 1994) and durational discrimination ability (Fraisse 1967, London 2002). Even if preferred pulse duration is obviously directly applicable to non-metrical context, it is reasonable to use pulse duration as a reference given the floating boundary between non-metrical and metrical structure.

Apart from preferred absolute duration, the position of the category within the category spectrum (*Assumption no 6*) is assumed to influence the conception of referential standard-duration. The categorical proportion rule, in combination with the limits of the categorical present, leads to the assumption that a referential standard category should ideally be divided into a maximum of three sub-categories and combined into a maximum of three super-categories. Further, neighbouring sub-categories should ideally be possible to relate to as divisions and combinations according to the categorical proportion rule and higher categories should ideally be able to express in terms of lower categories. Moreover, the frequency of occurrence of durations is assumed to influence the conception of durations in such a way that exceptional durations are assumed to have less categorical impact than frequent durations.

These conditions are implemented in the analysis of standard duration categorization and, in the current example this leads to the following list of standard categories denoted as midi click values with corresponding note values in common notation. The note value 1024 (quarter note) is appointed referential standard duration, and this category is enclosed in the figure below. Note that variations of standard categories are accepted when they can be expressed as categorical proportions of the next to the neighbouring lower category with a maximum of three subcategories within a standard category.

Chapter 7

Table 45. Standard duration categorization with corresponding midi click values and note values. The referential standard duration category is enclosed, the standard duration being 1024 or quarter note.

Midi duration values

(256)	
(512)	
(768 1024)	
(1536 1664 1792 2048)	
(3072)	
(4096 6144)	

This procedure thus creates a reduction of duration categories according to the assumed capacity of duration categorization within non-metrical structures. The above categorization could verbally be interpreted as very short durations, short durations, standard length durations, long durations, very long durations and extremely long durations.

The actual durations of the melody are then categorized by reference to this list of standard durations and to the successive relationships between events, stating that events which differ less than one-third of the longest duration of a pair are equalized while those of greater difference are quantized as belonging to different categories. Applied on the current *kulning* example, it creates the following notation:



Figure 7-9. *Kulning* by Karin Edvardsson after categorical quantization of event durations by standard categories and successive temporal relationships

#### 7.2.2.4 Preliminary phrase analysis

The object of the preliminary phrase analysis is chiefly to identify major discontinuities with implication for analysis of figure level grouping. The cognitive basis for this is that major discontinuities set limits for the figure level grouping, implying absolute limits for e.g. local metricity. The cognition of figure level grouping is thus assumed to be influenced by major discontinuities in such a way that the initial cognition of figures might be revised.

The preliminary phrase analysis begins with identification of possible phrase boundaries. These are identified by three different means: (1) global duration discontinuity, i.e. largest durational distance; (2) onset-offset-onset (IOI, tone-rest-tone) discontinuity and (3) registral discontinuity, all of these encompassing a psychological and categorical present. This means that a phrase boundary must be a local discontinuity maximum, within nine contiguous and rhythmically significant events, and encompass a time span which is closer to the tentative value of the perceptual present (or typical phrase duration) within the implementation, 5.625'', than to half of this value.<sup>131</sup>

In addition a phrase must be consistent with previous phrase duration, not be too close to a more marked discontinuity which would lead to the formation of a non-consistent phrase structure or create phrase lengths which do not comply with the above assumed restrictions to the psychological and categorical present. Moreover, discontinuities that involve durational difference have to be significant in relation to the assumed significant value of standard

---

<sup>131</sup> see Chapter 5, section 5.2.3.1 for a discussion of this value. See also London (2002:538)

Chapter 7

durations and, for onset-offset-onset discontinuity, the duration of the offset-to-onset duration must be over the limits of articulative rest, i.e. above ornament duration.

These phrase indications obtained by the initial global discontinuity analysis in the current example are marked in the figure 10 below.

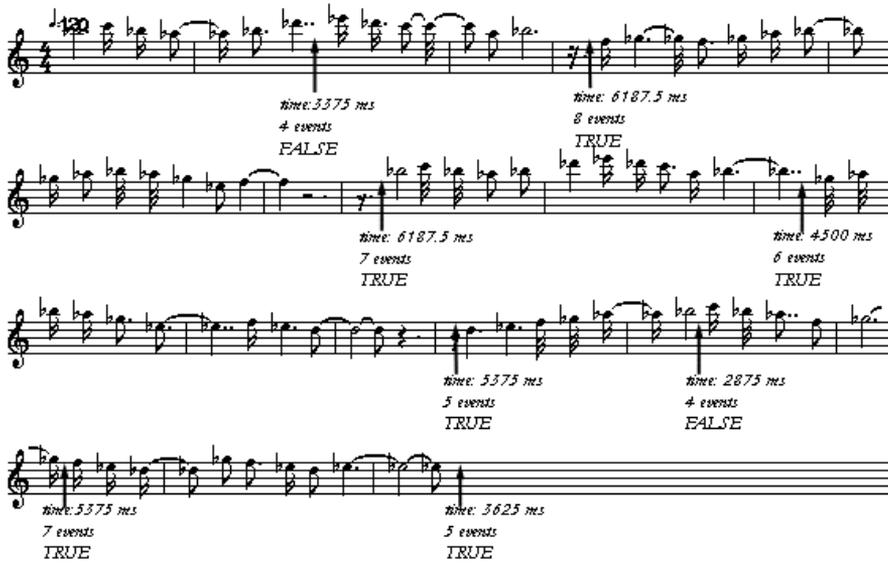


Figure 7-10. Global discontinuity indications obtained by the initial analysis of preliminary phrase structure. Duration of phrase segments (ms) and number of significant events within indicated phrase is given below phrase indication. True/False denotes if phrase indication is acknowledged by the analysis or not. (N.B. preliminary phrase analysis is performed on unquantized durations)

The phrase indications marked as false if they do not satisfy the conditions of phrase duration, consistency of phrase duration, number of rhythmically significant events or the conditions concerning proximity to next phrase start indication.

After this evaluation the phrase list is displayed in the notation by means of measures. The time signatures do not imply anything except the length of the preliminary phrases in terms of the common denominator value.



Figure 7-11. Preliminary central phrase level analysis of *Kulning* by Karin Edwardsson (included quantized note durations)

### 7.2.2.5 Primary grouping analysis at figure level

The figure analysis is the most problematic of the different stages of the non-metrical structural analysis, since cognition of primary groups is generally ambiguous. The assumed central level of cognition of non-metrical melodies is the phrase level, which is generally, as in the current example, indicated by global discontinuities, which by definition are more contextually evident. The figure level is hence a sub-grouping level, and the influence of different contextual and absolute factors such as metricity and absolute factors as limits of durational discrimination creates frequent perceptual conflicts where the structure becomes ambiguous.

The figure grouping analysis concerns the primary grouping level, and is performed mainly in a start-oriented grouping preference, that is, interpretation of group start from most gravely accented tones. (Start on rest is not considered applicable to non-metrical structure). This is for three main reasons: (1) The figure level analysis must be compatible with beat group analysis, since local true metricity or quasi-metricity may occur within a non-metrical context; (2) the gravest accented events are considered the primary skeleton tones of the non-metrical context; (3) To identify similarity between groups, insignificant variation such as ornamentation and sub-figuration have to be disregarded, which requires consistency between non-ornamented and ornamented figures.

The figure grouping analysis relies on group indication by analysis of local true metricity, quasi-metrical indication and local discontinuity indication. In this particular example, neither the analysis of local metricity, nor the quasi-metrical analysis leads to any grouping indications, hence the figure level analysis is based here entirely on local discontinuity analysis. The quasi-metrical analysis will be exemplified in section 7.3.

The analysis of group start indications by local discontinuity is in compliance with other parts of the analysis governed by a set of conditions applied to each event in its local context, primarily involving the three preceding and the three following events to the event in question. The fundamental principles are exactly the same as for metrical primary group analysis by discontinuity, which means that offset-discontinuity rules durational discontinuity,

Chapter 7

which in turn rules pitch discontinuity. The main difference regards the proportional influence of the different dimensions, assigning greater difference between the influence of rhythm and pitch than regarding metrical analysis.

This procedure will be exemplified in detail for the current example melody below:

a:1a x:1 x:3 b:8 h:4 b:8 h:4 x:4 b:8 h:4 b:8 a:4 a:1b b:8 j:4c x:1 x:3 x:6 h:4 x:6 x:1 x:3 b:8lm h:4 b:8 a:4  
 8 0 0 8 0 8 0 0 8 0 8 - 1 2 8 8 0 0 8 0 8 0 0 8 0 8 - 1

a:1a x:1 x:3 b:8 h:4 b:8 x:1 x:3 x:6 h:4 x:6 a:1a x:3 x:3 x:3 x:4 b:8 h:4 x:4 b:8 a:4  
 8 0 0 8 0 8 0 0 8 0 8 8 8 0 0 0 0 8 0 0 8 - 1

a:1a c: x:1 x:3 x:4 b:8 x:1 x:3 b:8 h:4 b:8 a:1c x:3 b:8 h:4 b:8 h:4 x:4 ö:2  
 8 8 0 0 0 8 0 0 8 0 8 0 0 8 0 8 0 8 0 8

Marker designations

- a:1a - first note  $\geq$  next  $\Rightarrow$  8
- a:1b - first note  $\leq$  next  $\Rightarrow$  2
- a:1c - ornament grouped with previous note  $\Rightarrow$  8
- a:4 - rest start, generally  $\Rightarrow$  0
- b:8 - local duration maximum, generally  $\Rightarrow$  8
- c: - longer than standard-duration, generally  $\Rightarrow$  8
- h:4 - local duration minimum, generally  $\Rightarrow$  0
- j:4c - shorter than previous, longer than next  $\Rightarrow$  8
- x:1 - ornament duration grouped with previous  $\Rightarrow$  0
- x:3 - ornament duration preceded by ornament  $\Rightarrow$  0
- x:4 - longer note on next, grouped with previous ornament  $\Rightarrow$  0
- x:6 - first after ornament, local duration maximum  $\Rightarrow$  8
- ö:2 - marked last notestart  $\Rightarrow$  8

Figure 7-12. Marker values for the individual events in *Kulning* by Karin Edvardsson. The discontinuity indications obtained by the computer implementation are displayed below the corresponding events with the connected marker value in the next row. The meaning of the different marker designations are briefly described in the table below the notation.

The local discontinuity valuations in this example are mainly related to one rule, grouping by durational maximum. This states that a local duration maximum groups the following notes up to its duration, i.e. grouping by durational accent. The only situation when shorter events are not grouped together with previous durational maximum is the beginning of measure 4, in which the number of shorter durations is too long to confine to the limits of the categorical present and consequently large enough to implicate a categorical present. The start is further supported by durational distance and within the group durational proximity.

As can be seen in the example above even longer rests after a shorter note are grouped with the preceding note, following the grouping by accent rule.

This leads to a grouping of events that can be viewed in the figure 13 below.



Figure 7-13. Figure level grouping in *Kulning* by Karin Edvardsson, designated by books between assigned groups at the top of the systems.

When comparing this figure level analysis with the figure level grouping assignment displayed by beamed notes in the close parallel *Kulning* by Karin Edvardsson presented by Moberg (see figure 1, section 7.1.1) one can see that they generally coincide in parallel situations. A start-oriented grouping at figure level is thus often used in notations of non-metrical melodies even though there is no beat conception (a feature of the structure, which is underlined by Moberg).

I have above presented the final figure grouping indications obtained by the computer model. There are, however, some situations in which there is conflict between different grouping principles, which may lead to a different grouping conception. Two of these will be exemplified below.

The first occurs in the very beginning of the melody. Consider the following excerpt of the opening phrase:

START-ORIENTED

grouping b 

grouping a 



grouping a 

grouping b 

corresponding END-ORIENTED grouping

Figure 7-14. Alternative grouping interpretations of the opening phrase of *Kulning* by Karin Edvardsson. Above the system the two different grouping start-oriented interpretations are displayed, with corresponding end-oriented interpretations below the system.

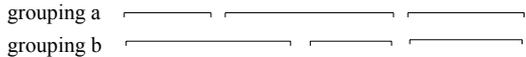
Grouping **a**, which is preferred by the implemented model, interprets the half note duration (note no. 6) as the chief local duration maximum of the opening section besides the first dotted half note duration. Thus, the intermediate two notes of rhythmical significance correspond to this duration, forming a separate group since they are preceded by two short durations which gives a durational accent to the first of the two dotted quarter notes. Grouping **b** reflects the rule that a stepwise ornamentation links the surrounding longer notes together, by durational proximity to the following and rhythmical insignificance in relation to the previous, i.e. as an extension by ornamentation of the previous. Thus, the four first notes form a complete group, while the second of the two dotted quarter notes becomes the beginning of the next group. In this case the start- and end-oriented interpretations concur, while the **a** grouping creates phase shifted start- and end-oriented interpretations.

The structural cue that makes the grouping **a** more prominent in the evaluation performed by current implementation is the relative and absolute duration of the second half note and the successive ornament, which gives a structural accent to this note that is interpreted as a start by the implementation. This is also the logical solution from a strictly non-metrical perspective.

I can myself actually conceive this passage in either way and really both interpretations influence my experience of the phrase, creating an essentially ambiguous conception where the first assumed **b** grouping becomes overridden by the **a** grouping. The solution provided by the computer model is, however, supported by Moberg's interpretation (see section 7.1.1 fig. 1).

The other example concerns the aforementioned start of phrase no. 4. Here at least two different start-oriented interpretations are also possible.

START-ORIENTED

grouping a 

grouping b 



grouping a 

grouping b 

corresponding END-ORIENTED grouping

Figure 7-15. Different grouping implication by parallel structural cues in phrase no. 4.

Grouping **a**, which once again represents the computer model solution, is supported by the aforementioned durational distance between the long ending note of the previous phrase and the assigned group start. It is also defined by the durational accent assigned to the dotted eighth note which becomes the final note of the figure. The figure is recognized since it contains enough number of elements; the ornament level durations become significant since the figure involves significant pitch change. The grouping **b** is determined by pitch accent,

through the recognition of the change of melodic direction at the third 16<sup>th</sup> note (symbolized by a “roof” image above the system). This change of melodic direction is also linked to the ending note of the previous phrase since it has identical pitch as this note. The two preceding 16<sup>th</sup> notes can thus be regarded an ornamentation extension of the ending note of the previous phrase giving structural accent to the top b<sup>2</sup> note.

What makes the computer model choose the **a** grouping, which in this case has concurring start- and end-oriented interpretation, is the absolute duration of the 16<sup>th</sup> notes which are below the ornament limit and the lack of duration differentiation at the point of change of melodic direction. The fact that the 16<sup>th</sup> notes are very short (from 78 ms – the first note – to 125 ms for the last of the notes.) enforces grouping by durational proximity.

But still, I can also, in this case, conceive both groupings and with generally longer durations, I would probably have conceived the **b** grouping as more salient. The choice made by the computer model is, however, logical in the sense that durational proximity will have stronger influence the shorter the durational distance. These examples of the inherent ambiguity typical of low-level grouping structure in non-metrical melodies, contrast to start-oriented grouping at figure level in metrical melodies which is uniformed by the accentuation implications given by periodicity. The predictions given by the model will thus probably be of lower significance when tested in relation to listener evaluations.

### 7.2.2.6 Phrase analysis at central phrase level

The phrase analysis is essentially performed in two steps, - the analysis of central and sub-phrase level and the subsequent analysis of higher phrase levels. The first part is based on phrase start indications determined by sequencing, i.e. the identification of adjacent repetitions of similar pitch and duration structures<sup>132</sup>, and global structural breaks based on the same structural features which determined the preliminary phrase analysis.

In the present example, one sequence of structural significance is identified by the sequence analysis. It identifies the similarity between the first and third phrase determined by the preliminary phrase analysis.



*Figure 7-16. Sequence identified by the sequence analysis in Kulning by Karin Edvardsson. The upper system represents first sequence half, the lower system second sequence half. The parts of the two sequence halves that are considered similar in the analysis are enclosed.*

<sup>132</sup> Sequence is here and elsewhere in this study used in the special significance of a structure determined by similarity between two adjacent series of sub-structural units. (c.f. chapter 6, *Metrical Phrase and Section Analysis*, section 6.1.4)

The two sections that are considered similar in this analysis, are also considered similar by Moberg in his analysis of a parallel *kulning* improvisation by Karin Edvardsson (see figure 1.) Even if this similarity would be revealed also by an event-by-event algorithm the benefit of the identification of a sub-structural level is evident here. As is obvious from the figure, the two similar parts differ substantially regarding event durations. And even if the repetition would have involved different ornamentation, the sequence analysis would be able to reveal the similarity as long as the figures concur. (An example of more obscure similarity will be provided in section 7.3.2, *folk hymn*)

There are, however, certain restrictions concerning the level and amount of similar events to determine a sequence with structural implication in non-metrical melodies in this analysis. The similarity has to extend over a psychological and categorical present or dominate two contiguous categorical presents. Furthermore it must involve higher-level similarity, with regard to both pitch and duration changes corresponding to NIL 4 and NIL 5 – type similarity within *Metrical phrase and section analysis* (see chapter 6, section 6.2.3.3).

The structural implications by sequence analysis are weighed against implications by global structural breaks by offset-onset, durational distance and pitch register change. The chief difference in relation to the previously described preliminary phrase analysis is that current analysis is based on difference between figures and not at event level. In other respects, it replicates this procedure. In this particular example, the final analysis of central phrase level comes up with exactly the same phrase boundaries.

If, however, the analysis comes up with phrase indications which imply phrases of different length, these are separated hierarchically by the categorical proportion rule (*Assumption* no. 5), and the phrase level which most consistently confirm typical phrase conditions, determined by the assumed standard duration of the perceptual present and the size of the categorical present, is appointed central phrase level.

### 7.2.2.7 Phrase analysis at higher phrase levels

The next step in the phrase analysis concerns higher levels, i.e. super-phrase and section levels. This is, on the one hand, based on similarity based classification of the central phrase levels segments which constitutes the foundation of a more elaborate sequence analysis and clustering of phrases by similarity; On the other hand, it involves an hierarchical analysis of the structural significance of global and local structural discontinuities. Here an analysis of global melodic direction (see chapter 5, section 5.2.4.6 *Analysis of Global Discontinuities*), which is not used in previous phrase analysis at central phrase level, is also added. The reason it is not involved in previous phrase analysis is that recognition of global melodic direction require more than one categorical present to be established. The different global phrase indications are then evaluated by quantification of structural impact.

The analysis of similarity between phrases was in the computer implementation performed by the same algorithm which is used in the metrical phrase and section analysis for segment root assignment (see chapter 6, section 6.2.4.5 *Sequence Analysis*). The similarity between beat groups, which is in the metrical phrase analysis, applied here concerns similarity between figures at corresponding positions. This implies that more general similarity, including similarity between non-adjacent primary groups, is considered in this analysis. The motive for this is that the phrases at central phrase level are considered the central structural units, which, due to their structural integrity, permit more general similarity to be perceived

between these units. Thus higher levels are not supposed to be of the same structural significance.

In the example melody, this analysis of more general similarity reveals a structural unity between the phrases in the *kulning*, which was but outlined in the previous sequence analysis.



Figure 7-17. Phrase root and variant similarity of central phrases as identified by the phrase similarity algorithm. The enclosed sections marks the unit which is regarded common to all phrases with root A.

As can be seen in the above illustration, the similarity identification involves e.g. similar contour, and similarity at root level does not require more higher level similarity between more than three figures. Equal variant, however, does require pitch set similarity.

Based on this classification of similar phrases, a sequence and clustering-by-similarity analysis is performed, which from the phrase root list (A A A B A C) identifies the super-phrase level (A A) (A B) (A C). The fundament of this analysis is that higher levels are constructed by combining lower levels into groups of two or three elements, governed basically by recurrent similarity (i.e. sequencing) and clustering of similar structures. In this particular example, the implied clustering into three contiguous A phrases and three other phrases (A A A) (B A C) is overridden by the alternating similarity (A B) (A C), implying a structuring in groups of two substructures.

The grouping obtained by sequencing and clustering of similar phrases is then evaluated in relation to local and global discontinuity analysis. This is done by adding the global register change value (0-3), the sequence indication value (0-3), an index value obtained by comparing lengths of preceding interonset durations (0-2), an index based on offset-onset occurrence (0-1), global change of direction (0-3) and a phrase length index based on the assumption of prominence of durational accent (0-2) to the original local discontinuity value of the phrase boundary (0-8). In the total valuation the sequence indication described above has a limited impact of 3/22. However, given that e.g. the basic local discontinuity value generally is equal for different segment boundary indications, and preceding interonset duration, global register change etc. usually also is close to equal, the influence of the sequence indication is in reality higher. This relatively small weight given to grouping by similar content – in comparison with the metrical phrase and section analysis - reflects the assumption that global discontinuity is

relatively more important in the conception of non-metrical melodies than in the conception of metrical melodic structure.

In the current example, the phrase grouping obtained by sequence analysis is enhanced by the analysis of global structural discontinuity. The total phrase-start-indication values is for the six phrases A1-18, A2-10, A3-16, B1-12, A4-14 and C-8. The valuation is influenced by the change of global melodic direction between A2-A3 and B1-A4 and the longer end durations at these break points. This results in a phrase analysis at a higher level which is included in the two output phrase analyses that are provided below. Internally, a start-oriented phrase analysis is basically performed, since more gravely accentuated groups are considered to be of greater structural significance. The preferred end output is, however, here by default end-oriented at phrase level, since non-metrical structure at higher levels is assumed to be conceived essentially end-oriented, determined by structural discontinuity.



Figure 7-18. Final phrase analysis of Kulning by Karin Edvardsson; Output from analysis by computer model. a) start-oriented interpretation (Upbeats are included in preceding phrases). The end-oriented interpretation is default output.



Figure 7-19. Final phrase analysis of Kulning by Karin Edvardsson; Output from analysis by computer model. b) end-oriented interpretation (Upbeats are included in preceding phrases). The end-oriented interpretation is default output.

## 7.3 Analysis of quasi-metrical melodic structure

### 7.3.1 General problems

To draw the borderline between metrical structure and non-metrical structure is, as has been mentioned, not a trivial matter. Because of the existence of mixed metrical and non-metrical structure, non-metrical sections within an essentially metrical melody make the tempo quantization module in the implementation of the current model fail; This implies that the structural analysis has to be performed within the framework of the non-metrical analysis. Further, it is very common within e.g. different traditions of European folk singing, that a melody with a metrical structure is performed and conceived in a non-metrical manner, obscuring metrical implications. (c.f. discussion in section 7.1.)

The question becomes then to what extent structurally determined quasi-periodicity is relevant for the conception of the melody? In a model which aims at identifying a cognitively relevant structural interpretation, the indications have to be evaluated by a weighted analysis based on the strengths of different indications. One would also want the model to ideally provide different structural interpretations, when the structure is ambiguous. Alternative interpretations are not yet provided in the current model, therefore the analysis must choose one interpretation on the basis of evaluation of the structural indications.

We will view some of the problems connected with quasi-metrical structure in two melodies, one of which has a more evident quasi-metrical structure than the other. Personally I conceive both these melodies from the performances used, as being performed in an essentially non-metrical manner, but it is no problem to attend to implied periodicity in either of the two performances of the melodies.

**Falska klaffare**  
*sjungen av Lisa Boudré, Vall, Ovansjö*

*Fritt, utan puls*

1. Fal - ska klaf - fa - re stå mig e - mot, -  
 de bju - da - till med sin ar - ga list och hot  
 De bju - da - till med al - lor lönn,  
 till - att om - gö - ra - vårt för - bund

*Figure 7-20. First strophe of a Swedish lyrical folksong sung by Lisa Boudré, Ovansjö, Gästrikland (b.1866), recorded 1935. My notation from recording (Rosenberg 1994:84)*

**220. P S A L M**  
*Gamla Psalmb. nr 381*

\* Närmast *ess*.

Figure 7-21. Folk variant of hymn originally sung by Pers Karin Andersdotter (b. 1835).  
Notation from live performance by Nils Andersson (Andersson 1922:130-131)

Regarding the first of these melodies, I have used the original recording from 1935 as input to the analysis. In the case of the second melody I have used a relatively new recording of the hymn melody, sung by the Swedish folk singer Susanne Rosenberg (Rosenberg 1991). She had originally learned the tune from the above transcription, but had been singing the tune for several years before the recording.<sup>133</sup> One interesting question regarding this example is if the quasi-metrical structure indicated in the original notation was in any way reflected in the performance.

The two melodies are chosen as examples for one main reason. The rhythmical structuring is quite different in the two performances, representing two types of rhythmical structuring in quasi-metrical melodies.

## 7.3.2 Analysis of two melodies with quasi-metrical structure

### 7.3.2.1 From Audio-to-midi-conversion to preliminary phrase analysis

The audio-to-midi-conversion was in these examples also performed by using the software Melodyne and imported into Finale 2002, with minimum quantization. Also, in these cases, the audio-to-midi-conversion had to be edited especially regarding some short articulations preceding notes of longer durations, which were to some extent identified as pitch transitions by the software, but due to variations in duration and pitch were inconsistently assigned as significant pitch shifts. The conversions also had to be slightly edited regarding the discrimination between rest and tone, the identification of short ornamented notes and equalization of notes into semi-tone steps.

Then the standard quantization and preliminary phrase analysis were performed on the two melodies respectively. We will not go into the details of this analysis here, since it is already described in section 7.2.2.4, and since the application on the two example melodies do not involve any other features of the model than what has already been described.

<sup>133</sup> It is taken from a commercial recording, and does not relate in any way to the current study

a) Unquantized input from audio-to-midi conversion



b) Resulting notation from quantization and preliminary phrase analysis



Figure 7-22. “Falska Klaffare” (first strophe) sung by Lisa Boudré. Audio-to-midi-conversion (a) and below (b) the result of the quantization and the preliminary phrase analysis

Regarding the song “Falska Klaffare”, sung by Lisa Boudré, the quantization ends up with a central standard category centred in the span of a quarter note up to a dotted quarter note. The phrase boundaries determined by the preliminary phrase analysis are given by extreme event durations and in some cases, also in combination with subsequent rests.

As can be seen in figure 23 below, the folk hymn variant “Hemlig stod jag” sung by Susanne Rosenberg, has a more elaborate structure:

Chapter 7

a) Unquantized input from audio-to-midi-conversion

a) Output result from quantization and preliminary phrase analysis

The image displays a musical score for the first strophe of the song "Hemlig stod jag" by Susanne Rosenberg. The score is presented in six staves of music. The first staff begins with a tempo marking of ♩ = 120. The music is written in a single melodic line on a treble clef staff. The notation includes various rhythmic values, including 32nd notes, and rests. The score is divided into measures by vertical bar lines. The key signature is one sharp (F#), and the time signature is 3/8. The music consists of six lines of notation, with the final line ending in a double bar line.

Figure 7-23. "Hemlig stod jag" (first strophe) sung by Susanne Rosenberg. Audio-to-midi-conversion (a) and below (b) the result of the quantization and the preliminary phrase analysis

The same standard quantization applies for this melody as for the other example; in this case more notes fall beneath the limit of notes of rhythmical structural significance, and are classified as ornament notes. They are all equalized into 32<sup>nd</sup> notes.

The preliminary phrase analysis make use of the same structural features as in the other melody, with phrase boundaries determined mainly by extreme durations and subsequent rests. In one case, a sub-phrase determined by a significant duration but followed by an articulative rest, which was acknowledged in the audio-to-midi conversion, was disregarded as a phrase boundary by inconsistency regarding phrase length and the category size requirements (see last phrase).

In both these melodies, the result obtained by the preliminary phrase analysis was consistent with the central phrase structure as indicated by the lyrics of the song. Each phrase concurred with the clauses of the lyrical lines as can be seen in the figures 24 and 25 below, where the lyrics are added to the output of the preliminary phrase analyses.

Chapter 7

♩ 120

Fal-ska klaff - a - re stú meig e - mo... De bjón - da\_\_ nill með sín arg-  
 en líst á þot De bjón\_\_-da\_\_ nill með al\_\_for lömna ni\_\_lítt o\_\_  
 m gö - ra\_\_ vör, för-boend

Figure 7-24. Preliminary phrase analysis of “Falska klaffare” sung by Lisa Boudré with the lyrics added below.

♩ 120

Hem-lig s sto... od ja... e... e... mo... or- go... Up-  
 þá e... is lö... m... is... sig stad Með me... che... græt... o...  
 úð... vör... de... Hó... rde... jag þot... - ror...  
 bað... dæ... de... o... om... e... om... ni... sig... me... De... ha...  
 me... ha... de... grí... þit... Ó... úð... þot... - döt... me... d... sta...  
 arkt... ba... ná

Figure 7-25. Preliminary phrase analysis of “Hemlig stod jag” sung by Susanne Rosenberg with the lyrics added below.

### 7.3.2.2 Primary grouping analysis at figure level

Quasi-metrical structure provides a great challenge for the primary grouping analysis at figure level. The two melody examples used here are evident examples of the problems connected with the identification of a figure level grouping.

To judge from the words of the songs as a reference for primary grouping, the rhythmic interpretation in the two songs represent two converse grouping principles in relation to durational accent. When comparing the rhythmic structuring in e.g. the first phrase in “*Falska klaffare*” with the second phrase in “*Hemlig stod jag*” by a simplified rhythmical notation in two duration categories - short and long - with reference to stressed syllables by word accent, the difference with regards to durational accent becomes apparent:

Figure 7-26 consists of two musical staves, A and B, each with a treble clef. Staff A shows a melody with notes grouped into five circles. Below the notes are the syllables: Fal - ska klaf - fa - re stá mig e - mot. The syllable 'stá' is underlined. Staff B shows a melody with notes grouped into four circles. Below the notes are the syllables: E - en mor - gon. The syllable 'E' is underlined.

Figure 7-26. Simplified rhythmic notation of the first phrase of “Falska Klaffare” (A) and the second phrase in “Hemlig stod jag” (B) with implied primary grouping by word accent marked by enclosure. Gravely stressed syllables are underlined.

The same durational pattern, when reduced into two categories, will thus be conceived in converse ways with relation to grouping by word accent. In relation to the durational pattern, the *A* interpretation is essentially start-oriented, grouping from the most accentuated event by durational accent; The *B* interpretation is essentially end-oriented, grouping by temporal proximity.

The question becomes whether this rhythmical grouping, as determined by word accent, can be traced by the analysis of the rhythmical structure of the two examples?

According to the rule of prominence of start- and end-oriented grouping at primary level (see chapter 5, section 5.3.4.2), the prominence for start-oriented grouping over end-oriented grouping relates to primarily seven different factors:

- level of strict periodicity (categorical integrity)
- number of iterations of the pattern (SL/LS)
- duration of the LS/SL pattern
- temporal categorical integrity of the LS/SL pattern
- durational contrast
- pitch distance
- phase and completeness of the iteration

Start-oriented interpretation is assumed to be more prominent if any of the following conditions apply: (1) the more strictly periodical a series of alternating long and short sounds (the LS/SL pattern) are (relating to assumed stronger influence of periodicity); (2) the greater the number of iterations of the pattern (relating to recurrence as factor which favours

periodicity perception); (3) the shorter the duration of the LS/SL pattern (relating to the assumed possibility to conceive the two events as a singular event – as an interonset), (4) the greater the temporal categorical integrity (c.f. strict periodicity), i.e. the less possible it is to interpret long durations as short durations and vice versa (which supports perception of reoccurrence of a pattern) (5) the less the durational contrast (in relation to the rule of categorical proportional conception and the decreasing influence of temporal proximity as a grouping factor); (6) The less the pitch distance within the SL pair is in relation to the LS distance (relating to the grouping by pitch proximity); and (7) finally if the phase and period completion is congruent with the start-oriented interpretation, implicating that a (LSLSLS) alternation will be more likely to perceive as a (LS) (LS) (LS) pattern than a (L) (SL) (SL) (S) pattern.

The reverse is assumed to apply for prominence for end-oriented grouping.

When examining the two phrases in relation to the factors noted above in the context of the less rhythmically reduced structure it is clear that they do not differ substantially in respect to the duration of the LS/SL pattern. Both sequences also have a pattern which, with respect to start phase (LS), implies a start-oriented interpretation, but, with respect to phase end, would favour an end-oriented pattern (SL). In general, this would imply a start-oriented interpretation for both phrases since phase implied by sequence start is more influential than phase implied by sequence end.

A 

Fal - ska klaff - a - re sta mig e - mo\_\_t

B 

e \_ \_ \_ \_ \_ e \_ \_ \_ \_ \_ mo \_ \_ \_ \_ \_ go \_ \_ \_ \_ \_

Figure 7-27. Two phrases with LS-iteration from “Falska Klaffare” (A) and “Hemlig stod jag” (B).

The two phrases are, however, more obviously different in other respects:

(1) Phrase **A** has a better fit with perfect periodicity than **B** for a LS periodicity; (2) **A** has a greater number of categorically discrete LS iterations (four iterations) than **B** (three iterations), which is assumed to be of significance regarding small number of iterations; (3) The LS groups are more categorically discrete in **A**, while they are not categorically discrete in **B** since e.g. the total duration of the fourth and fifth event is shorter than the third event; (4) The durational contrast within the pattern is measured as an average relatively similar in the two samples; the variation in durational contrast is, however, higher in **B**; (5) Pitch repetition within LS-pairs in the beginning of the phrase and distance between LS-pairs strongly supports LS grouping in **A**, while the pitch distance is equal within and between groups in the **B** phrase

Taken together, this supports start-oriented grouping in phrase **A** according to a LS-pattern. Conversely end-oriented grouping is supported in **B**, forming an interpretation of the SL alternation into a (L) (SL) (SL) (SL) pattern.

Thus, one can conclude that the word *accentuation* is actually mirrored in the rhythmical interpretation (see further below).

A further important implication of this in the case of “*Hemlig stod jag*” is that the end- and start-oriented interpretations concur at primary level, while in the case of “*Falska klaffare*” they are phase shifted at the primary level.

This analysis is performed by the quasi-metrical pattern recognition algorithm within the analysis of primary grouping at figure level. The phrases are first tested for perfect metricity starting from every element of the phrase and subsequently for quasi-metrical grouping. If none of these metrical tests result in a consistent grouping which lasts for at least three implied “beats”, the local and global discontinuity analysis evaluates the possible grouping from a strictly non-metrical perspective. (see previous section 5.2.2.5 *Primary grouping analysis at figure level*).

In the case of “*Hemlig stod jag*”, quasi-metrical iterative patterns are identified to be structurally significant in phrases 2, 3, 4, 6 and 7, the last of which a (SSL)(SSL)(SSL) pattern is recognized.

In “*Falska Klaffare*” as quasi-metrical grouping by iterative patterns only regarded significant in the first phrase.

The image shows a musical score for the song "Falska klaffare". It consists of three staves of music in treble clef with a 3/4 time signature. The tempo is marked as ♩ = 120. The lyrics are written below the notes. Vertical book-like symbols are placed at the top of the staff to indicate figure boundaries. The lyrics are: "Fäl-ska klaff-a-re stå mig e-moa De björ-da till med sin ar-ga av list å bot De björ-da till med al-lor löv-er ni ill att o-m gö-ra värt för-bered".

Figure 7-28. Primary grouping analysis performed on “Falska klaffare” sung by Lisa Bondré. Figure boundaries indicated by books at top of the system. Lyrics added to the computer model output.

Chapter 7

The image shows a musical score for the song "Hemlig stod jag" by Susanne Rosenberg. The score is written in treble clef with a tempo marking of 120. The lyrics are written below the notes, and figure boundaries are indicated by brackets above the notes. The lyrics are: "Hem-lig stod jag e-nne mo-ot-ter-son Upp-på e-nne lö-ven-ns sig stad Med my-cket grat o- och vä-der- Hjä-rtat jag hö- -ret bad dö-der o- om e-ns sorg-er ma-ns De-ns hö- om ha- de giv-its O- ch det- me- d sta- arkt ba- na".

Figure 7-29. Primary grouping analysis at figure level performed on “Hemlig stod jag” sung by Susanne Rosenberg. Figure boundaries indicated by hooks at the top of the system. Lyrics added to the computer model output.

In the figures 28 and 29 above, I have added the lyrics to the computer model output of the figure analysis. From this it can be seen that the great majority of figures concur with the structuring indicated by the lyrics, by indicated figure boundaries at word accents. This further supports the notion that the lyrical structure is reflected in the musical interpretation in these two examples.

In one important respect, however, the model fails. It regards several short pre-articulations, especially at phrase endings and phrase, as beginnings belonging to the previous group when the performers insert a short *acciaccatura* before the final syllable. These are not always identified as pitches by the audio-to-midi-conversion software, since they are generally

short and often do not start at a stable pitch. According to the generally start-oriented model of group analysis, these are regarded non-accentuated in relation to the succeeding tone, which is accentuated by durational accent.

The beginnings of the word make, however, these notes accented by grave dynamical accent hence making the succeeding note of longer duration un-accentuated with respect to dynamical accent. This feature of articulation is frequent in folk singing in Scandinavia (Rosenberg 1986) and can be regarded rather as an extended articulation of a tone than a separate pitch transition.

This stresses one important limitation to the current model. Since dynamical accent is not input information to the current model, it will fail in cases where dynamical accent is structurally more significant than durational and pitch accent.

Still, regarding the highly dynamically complex performances which are analysed here, the level to which the grouping indication by word accent is reflected in the analysis by the computer model can be considered substantial.

### 7.3.2.3 Phrase analysis

The phrase analysis at central phrase level does not differ in any important respect in the two current examples from the preliminary phrase analysis and will thus be left uncommented.<sup>134</sup>

The higher-level phrase analysis of “*Falska Klaffare*” sung by Lisa Boudré benefits mainly from structural contrast at global level, resulting in two super-phrases.



Figure 7-30. Final phrase analysis of “*Falska Klaffare*” sung by Lisa Boudré. Output from computer model.

<sup>134</sup> There is, however, a sub-phrase level indication in the phrase analysis of “*Falska Klaffare*”, which separates the phrases into two sub-phrases of about equal length. This is interesting from a metrical point of view since it reflects a possible phrasing at measure level. It is, however, disregarded by the phrase analysis since it does not consistently fulfill the categorical requirements of a phrase. It would be possible in a future development of the model to include such an quasi-metrical indication of sub-structure in the output of the model.

Chapter 7

This analysis reflects the lyrics of the song regarding the last super-phrase, but is not generally supported by the lyrics regarding the first super-phrase. The implication of this super-phrase is, however, not strong from the analysis of global discontinuity, but is implied by the structural unity of the two last central phrases.

Moreover, the phrase identification does not reveal the structural similarity between the phrases A1, B1 and D1 at central level, since the level of similarity is too low to be regarded significant for identical root designation. This might have created a phrase structure of (A1 A2) (B1 A3). If this similarity is cognitively significant to a degree for to be regarded by the current model needs to be evaluated by listener tests, which have not yet been performed.

The phrase analysis of “*Hemlig stod jag*” on the other hand, reveals an interesting conflict between grouping by phrase similarity and hierarchical relationships implied by global structural changes.

The image shows a musical score for the song "Hemlig stod jag" in 3/8 time, starting at a tempo of 120. The score is divided into several systems, each with a treble clef and a key signature of one sharp (F#). Above the notes, horizontal lines with brackets indicate the boundaries of phrases identified by a computer model. These phrases are labeled as follows: A1 (two occurrences), A2, A3, A4, B1, C1, D1, and E1. The annotations show how the model groups similar phrases across different parts of the song, such as A1 and A2, or B1 and D1, despite their different positions and melodic contexts.

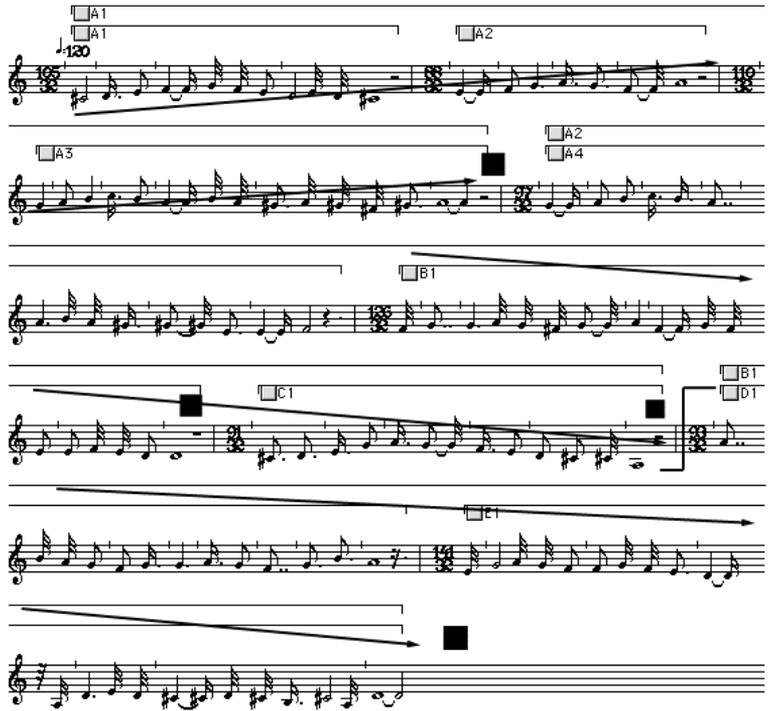
Figure 7-31. Phrase analysis of “*Hemlig stod jag*” sung by Susanne Rosenberg. Output from computer model.

The similarity analysis of phrases at central phrase level identifies similarity at root level for the first four phrases, which all have similar melodic contour in their beginnings when the ornamentation is disregarded.

The four subsequent phrases are all different at root level according to the similarity analysis.

A construction of a higher phrase level by clustering of similar phrases based on root similarity alone would imply a higher-level structure of four plus four phrases, this leads to a possible further symmetrical division in two plus two phrases (A A A A) (B C D E).

However, the analysis of global melodic change, register change and global duration changes tells a different story. This implicates a boundary after the third A phrase and between the C and D phrases. In the first case this is determined chiefly by change of global melodic direction, global change of duration and phrase impact by phrase duration. In the second case, the boundary is implied by change of melodic direction in combination with global register change and durational difference.



*Figure 7-32. Global structural changes considered significant in the analysis of higher phrase level of “Hemlig stod jag”. Lines with arrows designate change of global pitch register between phrases designated as melodic direction. Black boxes refers to hierarchically significant rhythmical breaks*

regarding phrase duration and ending event duration. The crooked line before the last super-phrase designates global shift of pitch register.

When this phrase analysis is compared to the phrase indications given in the original notation of the song, which was the basis for the interpretation that has been object for the analysis here, it reveals a striking similarity.

**220. P S A L M**  
*Gamla Psalmb. nr 381*

\* Närmast *ess.*

Figure 7-33. The original notation of “Hemlig stod jag”, from which the performer Susanne Rosenberg originally learned the song.

The central phrase level in the computer model analysis is here generally indicated as a sub-phrase level, to judge from the notation where phrases are marked by double bar lines and sub-phrases are indicated by fermata signs. The super-phrases generally concur and it would actually be possible to reduce the computer model analysis, based on the performance, to the original notation by more “fierce” quantization of figures into equal duration.

## 7.4 Evaluation of non-metrical analysis

### 7.4.1 Means of evaluation

The evaluation of the model of non-metrical analysis is, as has been said, not as extensive as the evaluation of the metrical phrase and section analysis and the metrical analysis.

The most important deficiency is that it has only been tested on melodies from the repertoire of Swedish folk music. The validity of the model with regards to other musical styles require further evaluation. Further, no formal listener tests have been performed to test the model. The reference material has only been notations, with indications of phrase and figure structure inferred by researchers and transcribers. Yet another means of evaluation have been to assume that lyrical structure reflects melody cognition, using the structure of the lyrics as reference for the analysis of melody structure.

Thus, a comparative study has been made involving twenty melodies from the Swedish herding calls and herding tune repertoires, identical to or parallel to the melodies used by Moberg (1959) in his study of melodic structure in this repertoire. An additional twenty from the Swedish folk song repertoire, mainly folk hymns and lyrical songs, which I have classified as having a typically non- or quasi-metrical structure was also used. Here the reference, was

mainly the lyrical structure. The result of this test of the model showed a significant result, but the implications of the result have to be considered in the light of the very limited and stylistically homogenous material used.

*Table 46. Results of test of non-metrical analysis performed on Swedish vocal folk music material*

Category	<i>Tot no of figures</i>	Ratio of correctly identified figures	Ratio of correctly identified phrases (central phrase level)
Herding calls (20 melodies)	482	0.62	0.81
Folk songs (20 melodies)	707	0.73	0.85
<b>TOTAL:</b>	1189	0.69	0.83

### 7.4.2 Discussion

As mentioned above, a more general test of the model of non-metrical structural analysis has to be performed to evaluate the validity of the model for other styles than those currently tested. One reason for the very limited study is the extensive amount of work involved in the editing of audio-to-midi-conversion, which could be considerably facilitated by further development of the current implementation of the model, as to allow for a more extreme note-values in the imported files. The use of another software for audio-to-midi-conversion that would allow for microtonal intonation would also be helpful in an extended study.

In spite of these difficulties the results indicate that a rule-based model can predict the cognition of non-metrical melody better than chance also in a non-metrical context. The relatively low ratio of correctly identified figures in relation to the study by Moberg, can e.g. partly be explained by the relatively high degree of quantization used in his transcriptions. They are also, to my mind, highly inconsistent regarding the means of identification of phrase length and similarity between phrases.

For example, a comparison of the phrase analysis provided by Moberg, on the parallel *Kulning* by Karin Edvardsson, with the analysis presented here, shows that the similarity between phrases, based on similar melodic contour, is not considered by Moberg.

Figure 7-34. Phrase analysis (end-oriented) of Kulning by Karin Edvardsson by the current model

Figure 7-35 Phrase analysis of a parallel Kulning variation by Karin Edvardsson by Moberg (1959:47). Phrases are indicated by reference to parallel version and ending note cadence, placed at the beginning of the assigned phrase.

In above example, Moberg, designates the phrases solely based on ending note similarity. Thus, the first and the third phrases are regarded as belonging to the same category, while the second and fourth phrases in Moberg's analysis are not regarded as similar. This implies further that the last phrase in Moberg's analysis, which he seemingly, according to the two phrase slurs, defines as consisting of two separate parts, is inconsistent in level designation in relation to the previous phrases. Hence, he will not be able to reveal the super-structure consisting of three super-phrases by his analysis, and consequently also neglects the variation structure at this level. Since he does not consider hierarchical phrase levels, the figure level interpretation becomes inconsistent: The figure interpretation at the end of the first phrases, in which the to short notes before the ending note are beamed, are different the corresponding melodic and rhythmic structure in the last phrase, in which the corresponding notes are separated.

Over a period of ten years, I have performed informal listening tests in which more than forty music students have participated. The task has been to segment this particular melody. In the resulting segmentations, the super-structural level is most frequently assigned the main phrase level. The phrase analysis and similarity between phrases are also generally confirmed by these tests to be of structural significance.

This implies that the application of the current model can reveal cognitively relevant structures which are not entirely trivial and may contribute to the understanding of melodic structure in non-metrical contexts.

The questions of validity of the results, which can be raised from making a comparison with the segmentations provided by a single researcher, emphasizes the need for further study of the cognition on non-metrical melodies.

The current method of analysis does, however, produce promising results.

# Appendix:

## List of Assumptions

---

### Assumptions in chapter 4

- (1) Pitches which follows, in immediate succession belong to different melodic pitch categories (with the exception of chromaticism noted below)
- (2) Neighboring pitches which do not follow in immediate succession in a melody can (under certain circumstances, see below) be experienced as pitch variants of the same category
- (3) Chromaticism within heptatonic, pentatonic or other MPC settings with less than 12 notes per octave are herein regarded as an exception of the general rule that different pitches in immediate succession belong to different MPCs. It is assumed that conception of chromaticism require categorical differentiation in relation to the MPC level, determined by the restriction of melodic motion within MPCs to immediate succession.
- (4) Chromaticism are assumed to be conceived as a temporary suspension of the MPC structure, comparable to glissando.
- (5) I assume that most MPC sets in melodies in different cultures can be delimited by the octave range.
- (6) I assume that there are certain typical maximum numbers of MPCs per octave, most frequently three, five, seven and twelve. I also assume that the maximum number is constant between octaves.
- (7) I assume that there are typical limits to the variation of intonation in relation to a reference pitch within MPCs in different musical styles. For analytical reasons, I have chosen a tentative limit of 1,5 diatonic steps (minor third).
- (8) I assume that a division of a pitch range into to near-to-equal intervals is preferred, within the intonation limit given in (7)
- (9) When the pitch set in the melody conform to a division of an octave into a perfect fifth and a perfect fourth in a heptatonic MPC set, the number of MPCs within perfect fifth is assumed to be limited to 5 (three intermediate MPC) and correspondingly four per perfect fourth.
- (10) When this initial assumption is in conflict with the above rules the MPC set is recalculated according to the assumption of a equidistant MPC set (including the possibility of dodecaphonic MPC set)

- (11) Categorization of MPC is a result of melodic motion with the possibility of different implications. The method of analysis has therefore to take more than one pitch shift (transition) into account
- (12) Implications of pitch set structure are symmetrical in relation to interval direction
- (13) The conception of MPC set emerges gradually and can be established through recurrent stimuli
- (14) The tendency to form a new MPC set based on divergent melodic information rather than to experience this as a deviation from an existing MPC set is inversely dependent on how well the existing MPC set is established, i.e. the number of pitch transitions involved.
- (15) Since the pitch changes are the fundamental means of creating a MPC set, pitch repetitions, returns and chromatic passages are omitted from the initial analysis.
- (16) Melodic pitch categorization can be interpreted differently among individuals and in different musical styles. Hence, a method of analysis of MPCs can only be a prediction of MPC conception (with the exception of formally defined MPC concepts).
- (17) This method of analysis of MPC aims to give a weighted analysis based on both context dependency and pitch identity
- (18) The absence of cultural and stylistic input sets limits for the validity of the predictions given by the method of analysis. The validity is dependent on the amount of information given by the melody and presupposed within a musical style/culture respectively.

## Assumptions in chapter 5

- (1) Meter is used here in the sense of perceived and/or structural periodicity in music, to which musical events are conceived as related. Meter is thus, whenever conceptually present, a primary level of the temporal organization of music
- (2) Conception of meter and gestural rhythm in music can be simultaneous and interrelated, but meter and gestural rhythm can also exist independently of one another, such as in music without conceived meter or music without conceived gestural rhythm.
- (3) Conception of metrical structure, the degree to which a perceived periodicity is conceived as a temporal grid, can be regarded as relative. There can be extra-metrical events within a generally metrical context.
- (4) The ability to conceive meter in music is assumed to be innate and spontaneous, closely related to the ability to relate to periodicity as a means of coordinating actions and can be regarded as a streaming phenomenon.

- (5) Meter is here used in the sense of conceived periodicity, including quasi-periodicity, which implies that events that are not phenomenally isochronal can be conceived metrically. It also implies that a quasi-periodical temporal grid may be conceived as regulative, involving expectations of temporal attention according to a significant scheme of period fluctuations.
- (6) Metrical conception is in concordance with the general model of melody conception assumed to be evolving gradually through the listening to a piece of music
- (7) People are assumed to be sensitive to physical periodicity (regularity) in sound patterns (timing) and to the consistency of periodicity.
- (8) There are individual differences in the tendency to relate metrically to music. Such differences can possibly even be culturally or stylistically discrete.
- (9) A periodical change in any perceptually recognizable dimension of musical sound can be conceived metrically
- (10) Pulse is the periodic flow of beats in music, the fundamental structure of conceived periodicity in music, whenever present it constitutes the basic temporal mapping of music
- (11) Beats are perceived or structural periodic markings of time (i.e. pulse periods) or points of temporal expectation in music, and can be conceived as the basic units to which temporal events relate (including complex metrical events)
- (12) If a perceived periodicity is conceived as pulse relates to tempo (pulse frequency). Periodicity is perceived as pulse within a certain tempo-range.
- (13) It is herein assumed that, for a pulse to serve as the basic temporal mapping, the periodicity must be perceivable within the basic scope of attention limited to the perceptual present.
- (14) If a perceived periodicity is conceived as pulse is related to interonset structure, which duration categories that are present. A pulse is ideally a central duration category and conforms to the limitations of the categorical present.
- (15) Meter is assumed potentially be conceived as complex in a hierarchical manner, as more than one level of periodicity in simultaneous coordination. This can be expressed as conception of superimposed pulses (polymetricity) or the simultaneous experience of time and pulse. One of these levels are often conceived as the central pulse level or tactus.
- (16) The ability to conceive metrical conflict between pulses of different tempo in a complex metrical conception is assumed.
- (17) Metrical complexity at pulse level can include conception of superimposition of pulses as well as composite beat groups / beat asymmetry. Pulse should however be possible to conceive as the basic periodical measurement of the music, while it has to conform to the conditions given by the perceptual present.

- (18) Pulse at atomic level is structurally defined mainly by periodical interonset intervals, while higher metrical levels such as composite pulse and time are dependent to larger extent on periodicity in grouping of events, which involve other musical dimensions such as pitch. The strength of durational accent is assumed to be inversely relative to the number of events within the period.
- (19) Time (meter at measure level) is defined as regulative periodicity determined by periodical grouping of beats
- (20) Meter is inferred from grouping by the perception of group starts as impulse points, as structural accents. Grouping is inferred from meter by the perception of temporal positions determined by the metrical grid as impulse points, implicating group start.
- (21) Meter at measure level is in analogy with the general concept of meter assumed to have punctuate dimension, implying impulse points. Thus measures, in terms of pulse groups, will be counted from the beat conceived to have the most prominent accentuation.
- (22) It is assumed that people possess the ability to experience structural similarities between accent types (phenomenal/structural/metrical) with human motion schemata and the behaviour of objects of different weight on a subconscious level. The categorization in grave and acute accents is thus assumed to be conceptually relevant.
- (23) It is assumed that grave accents are more metrically prominent than acute accents. Thus, metrical groups are more likely to be counted from beats with grave accent, than from beats with acute accents.
- (24) Metrical grouping concerns primarily low-level, primary grouping of elements within the limits of the perceptual and categorical present
- (25) Shorter groups are preferred to longer. This applies for the total duration as well as the number of consisting elements.
- (26) Primary grouping in two or three elements is preferred. Conception of larger groups as composed of subgroups of 2 or 3 elements is preferred.
- (27) Grouping by similarity, continuity and proximity – 'sameness': Similar and close events tend to be grouped together.
- (28) Grouping by discontinuity, dissimilarity and distance – 'difference'. Change / difference indicates group boundaries
- (29) 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
- (30) Grouping by symmetry. Symmetrical grouping is more prominent than asymmetrical grouping and can be implicative in the case of grouping by hierarchical symmetry.
- (31) Grouping by perceptual prominence / foreground, i.e. articulation, impulse and gravity: Grouping between most articulated events is prominent.

- (32) Grouping by integrity/'prägnanz' / contrast:  
Discrete grouping is prominent.
- (33) More unambiguous and precise structural indications of group boundaries are superordinate to structural features which give more ambiguous or vague indications of group boundaries
- (34) Indications of group boundaries which involve more elements are superordinate to structural features which involve less elements.
- (35) Combinations of different indications are superordinate to single indications

## Assumptions in chapter 6

- (1) The substructure of a melody is syntactic; substructures relate to each other to form a meaningful whole.
- (2) Melody is thought to be conceived in local segments, where segment size is determined basically by the limitations of the perceptual present. Thus, there is always at least one substructural level to a melody.
- (3) Since the conception of the melody takes place in time and can be revised throughout the experiencing process, the method of analysis must to consider the whole melody.
- (4) For the same reason as in (2), i.e. global structure is revealed in retrospect, the global substructures determines the local substructures, which implies that the identification of more global substructures, e.g. sections, will override more local sub-structures, e.g. phrases.
- (5) The identification of global structures is determined by the analysis of local structure, which implies that local structure must be input to the analysis of global structure.
- (6) Segment boundaries are conceptually determined either by the start or the end of the segment, start- or end oriented. They can be considered as complementary conceptions; the first focuses on impulse quality, the latter on gestalt quality and may be combined in forming a conception of segment structure.
- (7) Segmentation by similarity is generally start-oriented, since the similarity is recognized by the recurrence of a pattern.
- (8) Segmentation by similarity is generally structurally significant in metrical melodies since similarity between segments is a powerful factor for determining syntactic relationships between segments.
- (9) Similarity has greater significance as a structurally determining factor on a global level, while discontinuity is more influential on a local level.
- (10) Sequence in the sense of segmentation by contiguous repetition is a stronger structural factor than melodic parallelism in general.
- (11) Similarity between segments is easier to recognize when segment boundaries are prominent, thus allowing more general similarity to be recognized

- (12) In a metrical context, melodic segments of structural significance will be congruent to the conceived central pulse level/tactus. Thus, structurally determinant groupings will primarily begin on beats.
- (13) The existence of a metrical grid makes it possible to conceive time spans of melodic segments, which makes duration of segments a possible structural factor. Symmetry and periodicity/good continuation thus become more forceful as determinants of grouping structure and makes structural hierarchy easier to conceive.
- (14) As a structurally determining factor the strength of segment duration is inversely related to the length of duration in terms of metrical units. The longer the segment, the less determining its duration relationship to other segments will be.
- (15) And as a consequence of (12), there is a limit to the degree which symmetry between successive segments can be perceived as a structurally determining factor. This is tentatively in the proposed model set to 48 beats<sup>135</sup> based on the limitations of category conception
- (16) There are two different basic aspects of pitch structure when regarding similarity between segments, pitch set similarity and pitch change similarity, the former being based on absolute local pitch memory and the latter being based on relative pitch change.
- (17) While pitch set similarity is assumed to be more specific pitch change is assumed to be of greater structural significance, being a stronger determinant of melodic identity.
- (18) The main structurally significant quality of pitch in melody is assumed to be pitch height, which implies that change between melodic pitch categories are perceived in terms of pitch height, patterns of ups and downs.
- (19) Pitch change in melody is assumed to be perceived categorically. The level of melodic pitch categories (MPC) is hence assumed to be the primary structural level of pitch in melodies. Two segments which are identical with regard to melodic pitch categories will therefore be regarded structurally identical.
- (20) Melodic motion, in terms of pitch-change between melodic pitch categories (MPC), is cognitively grouped into two basic classes, steps and leaps. The former type refers to motion between neighboring MPC categories. A step is assumed to be of the maximum size of a major third in this model, based on MPC systems in different traditional musical styles.
- (21) Leaps are categorized as small or large, a difference which is considered to be structurally significant. This categorization is fundamentally contextual, and also has general limits. Leaps exceeding a perfect fifth are assumed to be conceived as large in this model. However, the influence of this categorization is context dependent.

---

<sup>135</sup> In concordance with the beat scope model, which sets the maximum periodicity perspective to eight successive basic psychological/categorical presents of 6 beats.

- (22) Pitch change is regarded as possible to conceive between groups of tones, and between articulated tones.
- (23) Pitch change between the initial notes of beat-groups (primary groups determined by beat period), is, along with successive tone-to-tone pitch change, the most structurally significant level of pitch change in metrically conceived melodies.
- (24) Global pitch changes regard the changes between groups of beats and are categorized into global change of register and global change of melodic direction, typically defined within the limits of the categorical present ( $7 \pm 2$  beats)
- (25) The more precise structural quality of pitch becomes, the more salient it will be in the determination of melodic structure. However, it need not reflect a difference of structural significance.
- (26) Patterns of event duration are the most specific level of rhythmic structure<sup>136</sup>
- (27) Patterns of duration changes are generally more structurally significant than patterns of absolute duration
- (28) Patterns of interonset/duration are basically determined by successive relationships between pairs of events, the categorization into long and short durations. On a more specific level, the successive temporal relationships are conceived in terms of small integer ratios, the number of categories of duration relationship being limited to the capacity of the categorical present.
- (29) It is assumed that the categorical relationship between interonsets/ duration of events and beat length (at central pulse/tactus level) can be structurally significant.
- (30) On the basis of (26) it is assumed that rhythmic structure in metrical melodies can be conceived as patterns of beat categories based on interonset qualities, as patterns of divided, undivided and tied beats.
- (31) At a more specific level, this rhythmic categorization of beats also considers the relationship within divided beats as being structurally significant, involving a categorization into the two basic categories short and long durations within the beat.
- (32) Global rhythmical changes, based on rhythmical qualities of groups of beats, can be structurally significant
- (33) More specific rhythmic structural means rule takes precedence over more general means, given that they concur with the primary metrical structure
- (34) Prominence of rhythmic structural means relates to the metrical impact of the change, i.e. the number of beats involved
- (35) Rhythmic similarity is more structurally significant on a local level (within the perceptual present), while pitch similarity is more structurally significant on a global level (the more events involved)

---

<sup>136</sup> This concerns interonsets as well as offset-onsets

- (36) Melodic structure in melodies with a metrical structure is typically hierarchical. It is assumed that it is possible to conceive multiple hierarchical levels of melodic segments even in a monophonic melody when metrical structure exists.
- (37) Grouping by similarity, continuity and proximity – 'sameness': Similar and close events tend to be grouped together
- (38) Grouping by discontinuity, dissimilarity and distance – 'difference'. Change / difference indicates group boundaries
- (39) Grouping by good continuation / constancy. Group size, group start etc., which is coherent with previous grouping is prominent and can be implicative
- (40) Grouping by symmetry. Symmetrical grouping is more prominent than asymmetrical grouping and can be implicative
- (41) Grouping by perceptual prominence / foreground, i.e. articulation, impulse and gravity: Grouping between most articulated events is prominent
- (42) Grouping by integrity/'prägnanz' / contrast: Discrete grouping is prominent. Groups with a discernible, salient, unique inner structure or groups which are discernible/salient/discrete in the context are prominent.

## Assumptions in chapter 7

- (1) There is no clear boundary between metrical and non-metrical rhythmic structure from a phenomenal point of view. The same structure may be conceived metrically and non-metrically by different people. There can also be a mixture between metrical and non-metrical sections within the same melody.
- (2) There is a categorical difference between metrical and non-metrical rhythmic structure, regarding the conception of meter being a regulative factor.
- (3) Temporal relationships between musical events in a non-metrical rhythmic structure are assumed to be perceived primarily locally, based on relative and categorical perception of pair-wise relationships between durations, but influenced by global categorization of standard durations (see 6).
- (4) The categorization of durations in non-metrical contexts is assumed to be limited to the perceptual present, typically extending up to 5-6'', and the categorical present, typically allowing for  $7 \pm 2$  simultaneous categories.
- (5) The basic categorization of durations is assumed to be governed by the categorical proportion rule, which states that conception of proportions between atomic units in time ranges is limited to equal, duple, triple and quadruple division.
- (6) The global categorization of durations is also assumed to be influenced by the absolute duration of events and event groups, favoring the conception of a referential standard duration category, which is ideally concurrent with salient pulse duration and ideally in the centre of the category spectrum.

- (7) Complex rhythmic relationships based on the relationship to a meter, such as syncopation, and complex rhythmic relationships which require metrical subdivision are not assumed to be applicable in a non-metrical rhythmical conception.
- (8) The lack of a general temporal grid makes the conception of structure in non-metrical contexts highly dependent on structural contrast. The conception of grouping structure in non-metrical contexts with low structural contrast is thus assumed to be highly ambiguous.
- (9) The lack of a general temporal grid, influences the strengths of the factors determining melodic structure. The primary grouping principle of similarity, specifically regarding grouping by sequencing, is assumed to be generally subordinate to discontinuity. The secondary grouping principles of good continuation regarding periodicity, successive and hierarchical symmetry are assumed to be less influential than in metrical contexts. Thus are these grouping principles have been dethroned to ternary grouping principles with regards to non-metrical structure on higher levels, i.e. not being of implicative strength.
- (10) The grouping levels in non-metrical melody may be divided into primary grouping level, sub-phrase levels, central phrase level and super-phrase/section levels.
- (11) Of these, the central phrase level seems to be the most conceptually salient level. One might rather designate non-metrical rhythmical structure as phrase-oriented structure.
- (12) The primary grouping level, here designated figure level, is assumed to be corresponding to grouping at beat level in metrical music, while the central phrase level is assumed to be basically confined to the limits of the psychological and categorical present. This makes the central phrase level comparable with e.g. 2-3 measure phrases in metrical music made up from typically  $7 \pm 2$  primary groups, while phrases at measure level in metrical contexts may rather be compared to sub-phrase level within non-metrical structure
- (13) It is here assumed that grouping at primary level gives rise to a sub-structuring in more structurally significant events defined primarily by articulation *Gerüststöne*, and structurally less significant events which can be considered as embellishment of the structure made up by the more articulated events.
- (14) Conception of higher phrase levels and sections are assumed to be relevant in non-metrical contexts but to a lesser degree than in metrical contexts, since the means of establishing temporal hierarchical relationships are restricted by the lack of a general metrical grid.
- (15) The same structural principles are assumed to apply for the conception of figure level groups as for primary groups at beat level in metrical contexts, with the exception that the influence of secondary grouping principles such as

periodicity and symmetry is less prominent. (see chapter 5, section 5.2.4.3 and chapter 6, section 6.1.7)

- (16) As a consequence of the assumption (1) that there is no clear structural boundary between non-metrical and metrical rhythmic structure, conception of periodicity is assumed to be of possible significance in an essentially non-metrical conception, without being a regulative factor with structural implication. This implies that in a melody that is essentially conceived as phrase-grouped, local periodicity can influence grouping at lower levels.

# References

---

## Literature

- Ahlbäck, Sven (1986). *Jembergslåtar*. Gävle: Länsmuseet i Gävleborgs län.
- Ahlbäck, S (1989) *Metrisk analys – en metod att beskriva skillnad mellan låttyper*, Trondheim, Norsk Folkmusikklags skrifter nr 4
- Ahlbäck S & Emtell S, (1993) A Computerized Method of Analysis of Tonality in Monophonic Melodies. *Annual Conference of the Society for Music Perception and Cognition : Abstracts* (SMPC) Philadelphia
- Ahlbäck, Sven (1993), *Tonspråket i äldre svensk folkmusik*, Stockholm: Udda Toner
- Ahlbäck, Sven (1983/1985/1996), *Låtpuls – karaktäristiska egenskaper för låttyper*, Stockholm: Udda Toner
- Ahlbäck, Sven (1994), A computer-aided method of Analysis of Phrase Structure in monophonic melodies, in I Deliège, (Ed), *Proceedings of the Third International Conference for Music perception and Cognition*. Liege, Belgium: ICMPC.
- Ahlbäck, Sven (1997) A computer-aided method of melodic segmentation in monophonic melodies, in A. Gabrielsson, (Ed), *Proceedings of the Third Triennial ESCOM Conference*. Uppsala, Sweden: ESCOM
- Ahlbäck, Sven (2001) From where do we count? Asymmetrical beat in polska melodies. In: Polish-Scandinavian Symposium on the Polska. Publikationer från Svenskt Visarkiv (in press)
- Andersson, N & Andersson, O (1922-1940). *Svenska låtar*, Stockholm: Gidlunds. *Facsimile 1972*
- Baily, J., (1985). Musical structure and human movement. In P. Howell, I. Cross, and R West (Eds.), *Musical Structure and Cognition*, London (237-258)
- Bartelette, J. C., & Dowling, W. J., (1980). The recognition of transposed melodies: A key-distance effect in developmental perspective, *Journal of Experimental Psychology: Human Perception and Performance*, 6, 501-515
- Bartók, B. (1920). Ungarische Bauernmusik, *Musikblätter des Anbruch*, No. 11-12, (422-424)

- Bertrand, D (1994). A developmental approach to preference rules for grouping: principles of organization in the music domain. *Escom newsletter no. 5*, April 1994
- Bengtsson, I., & Gabrielsson, A., (1977). Rhythm research in Uppsala. In *Music, Room, Acoustics, Publications issued by the Royal Swedish Academy of Music* No. 17, Stockholm
- Bengtsson, I., (1979). Rytym, *Sohlmans Musiklexikon*, volume 5:248-250, Stockholm: Sohlmans förlag.
- Bengtsson, I., & Gabrielsson, A., (1980). Methods for analyzing performance of musical rhythm. *Scandinavian Journal of Psychology*
- Bengtsson, I., & Gabrielsson, A., (1983). Analysis and Synthesis of Musical Rhythm. In J. Sundberg (Ed.), *Studies of Music Performance*, Publications issued by the Royal Swedish Academy of Music No. 39, Stockholm
- Björnerberg, A., (1996) Om tonal analys av nutida populärmusik. *Dansk årsbog for musikforskning XXIV* 1996
- Blacking, J., (1973/1976), *How Musical is Man?* Seattle: University of Washington Press
- Blacking, J., (1987) *A Commonsense View of All Music': Reflections on Percy Grainger's Contribution to Ethnomusicology and Music Education*, Cambridge
- Blom, J-P. & Kvifte, T. (1986), On the Problem of Inferential Ambivalence in Musical Meter. *Ethnomusicology* vol 30
- Blom, J-P. (1989). *Rytme og metrikke i svensk folkemusikk – en kommentar til Sven Ahlbäck's typebeskrivelse*, Trondheim, Norsk Folkemusikkklags skrifter nr 4
- Blom, J-P (2001). Making the music dance. On dance connotations in Norwegian fiddling. In: *Crossing Boundaries: Conference on North Atlantic Fiddle Traditions*. University of Aberdeen
- Bod (2001), Memory-based models of melodic analysis: Challenging the gestalt principles. *Journal of New Music Research* 30(3)
- Brown, H., (1988), The Interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception* 7, 219-250
- Burns (1999) In D. Deutsch (Ed.), *The Psychology of Music* (pp. 215-258). New York: Academic Press.
- Cambouropoulos, E. (1998) "Towards a general computational theory of musical structure" Unpublished doctoral dissertation, University of Edinburgh, Scotland.

*Appendix*

- Cambouropoulos, E., (2003) Pitch Spelling: A Computational Model. *Music Perception*, 20 (411-429)
- Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L., (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, 113, (394-412)
- Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music*, Stockholm: Royal Swedish Academy of Music. (pp. 19-34)
- Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research*, 51, 2-9
- Clarke, E. F., & Krumhansl, C. L., (1990). Perceiving musical time. *Music Perception*, 7, 213-252.
- Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 473-500). New York: Academic Press.
- Cook, N., (1987). *A Guide to Musical Analysis*, London: J.M.Dent & Sons Ltd
- Cook, N., (1990). *Music, Imagination, and Culture*. Oxford: Oxford University Press.
- Cooke, D., (1959). *The Language of music*. London
- Cooper, G. & Meyer, L.B. (1960) *The Rhythmic Structure of Music*. Chicago: University of Chicago Press
- Cross, I., (1995) Review of E. Narmour's: 'The Analysis and Cognition of Melodic Complexity', *Music Perception*, 12(4):486-509
- Cuddy, L. L., Cohen, A. J., & Miller, J., (1979). Melody recognition: The experimental application of rules. *Canadian Journal of Psychology*, 33, (148-57)
- Deliege, I., (1987). Grouping Conditions in Listening to Music: An Approach to Lerdahl and Jackendoff's Grouping Preference Rules. *Music Perception*, 4 (325-360).
- Desain & Honing (1992) *Music, Mind and Machine*. Thesis Publishers, Amsterdam.
- Deutsch, D. (1972a). Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology*, 93, 156-62
- Deutsch, D. (1972b). Octave generalization and tune recognition. *Perception & Psychophysics*, 11, 411-12

- Dictionary of Musical Quotations*, (1994) compiled by D. Watson, Hertfordshire: Wordsworth Ed Ltd
- Dowling, W.J. & Fujitani, D.S. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 49, 524-531
- Dowling, W. J., (1973). The perception of interleaved melodies. *Cognitive Psychology*, 5, 322-37
- Dowling, W. J., (1978). Recognition of Melodic Transformations: Inversion, Retrograde, and Retrograde-Inversion. *Perception and Psychophysics* 12 (417-421)
- Dowling, W. J. & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, 1, 30-49
- Dowling, W. J., & Harwood, D. L. (1986). *Music Cognition*. Orlando: Academic Press.
- Dragoi, S. V., (1959). Musical Folklore Research in Rumania and Béla Bartók's Contribution to it, in Z. Kodály (Ed), *Studia Memoriae Belae Bartók Sacra*, Budapest: Publishing house of the Hungarian Academy of Sciences
- Edworthy, J. (1985). Interval and Contour in Melody Processing. *Music Perception*, 2:3 (375-388)
- Emtell, S. (1993). *Computer Aided Music Analysis regarding tonality in monophonic melodies*, Stockholm: Royal Institute of Technology Unpublished Masters Thesis
- Fraisse, P. (1956). *Les structures rythmiques*. Editions Universitaires de Louvain
- Fraisse, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*. Vol. 8 (pp. 203-254). New York: Academic Press.
- Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). New York: Academic Press.
- Fraisse, P. (1987). A historical approach to rhythm as perception. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 7-18). Stockholm: Royal Swedish Academy of Music.
- Frances, R., (1958). *La perception de la musique* [The perception of music (W. J. Dowling, trans.) Hillsdale, NJ Earlbaum], Paris: J. Vrin
- Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, 98, (2524-2531)
- Friberg, A., (1995) *A Quantitative Rule System for Musical Performance*, Stockholm: Royal Institute of Technology

*Appendix*

- Friberg, A & Frydén & Sundberg, J. *The Quantitative Rule System for Musical Expression*
- Gabrielsson, A. (1973b). Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology*, 14, (244-260)
- Gabrielsson, A., Bengtsson, I., Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology* 24
- Gabrielsson, A. (1984). Studies of Musical Rhythm. *Comunicazioni Scientifiche di Psicologia Generale*. (249-266) Rom: Bulzoni Editore
- Gabrielsson, A. (1987) Once Again: The Theme from Mozart's Piano Sonata in A Major (K.331), In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 7-18). Stockholm: Royal Swedish Academy of Music.
- Gabrielsson, A (1988), Timing in music performance and its relations to music experience. In J Sloboda (Ed.) *Generative processes in music*. Oxford: Clarendon Press
- Gabrielsson, A. (1993). The complexities of rhythm. In T. J. Tighe & W. J. Dowling, (Eds.), *Music: The understanding of melody and rhythm* (pp. 93-120). Hillsdale, NJ: Lawrence Erlbaum.
- Garner, W. R. (1974) *The processing of Information and Structure*. London: Lawrence Erlbaum Associates,.
- Grove's Dictionary of Music and Musicians, The new 6a ed 1980 Macmillan, London*  
Articels: Bar, Beat, Metre, Rhythm, Pitch,
- Gurvin, O. (1959), *Norwegian Folk Music, series I*. Oslo: Oslo University Press
- Hood, M., (1971). *The Ethnomusicologist*. New York: MacGraw-Hill
- von Hornbostel , E. M., (1948). The Music of the Fuegians, in *Ethos* (61-102)
- Hugardt, A. *Individuality and Development in Children's Spontaneous Tempo and Synchronization*. Göteborg: Publications from the Department of Musicology, Göteborg University, No 66
- Höthker, K., Thom, B., Spevak, C., (2002) *Melodic Segmentation: Evaluating the Performance of Algorithms and Musicians*. Technical Report 2002-3. Faculty of Computer Science, University of Karlsruhe
- Höthker et al 2002a (see Spevak et al 2002)
- Hötkher et al 2002b (see Thom et al 2002)
- Höthker et al 2002c (see Höthker et al 2002)

- Jersild, M & Ramsten, M. (1987). Grundpuls och lågt röstläge – två parametrar I folkligt sångsätt. *Sumlen*, Stockholm
- Jersild, M. & Åkesson, I. (2000). *Folklig koralsång, en musikeknologisk undersökning av bakgrunden, bruket och musiken*. Södertälje: Gidlunds.
- Johnson, A. (1980). Sången i skogen, skogen i sången. Om fåbodamas musik. In J Ling et al (Eds) *Folkmusikboken*, Stockholm: Prisma
- Johnson, A. (1986). *Sången I skogen, studier kring den svenska fåbodmusiken*, Institutionen för Musikvetenskap, Uppsala Universitet, Uppsala
- Jones, A.M. (1959). *Studies in African Music*. London: Oxford University Press
- Kessler, E. J., Hansen, C., and Shepard, R. N., (1984). Tonal schemata and the perception of music in Bali and in the West, *Music Perception*, 2, (131-165)
- Kolinski (1947). The Determinants of Tonal Construction in Tribal Music. *The Musical Quarterly*. New York: G. Schirmer
- Kolinski (1961). Classifications of Tonal Structures. *Studies in ethnomusicology* New York : Oak Publications
- Krumhansl, C.L. (1990) *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press
- Krumhansl, C. L. (1995). Effects of musical context on similarity and expectancy. *Systematische Musikwissenschaft (Systematic musicology)*. 3, (211-250)
- Krumhansl, C. L. (1996). A Perceptual Analysis Of Mozart's Piano Sonata K. 282: Segmentation, Tension, and Musical Ideas. *Music Perception*, 13, (401-432)
- Krumhansl, C. L. (1997). Effects of perceptual organization and musical form on melodic expectancies. In M Leman, *Music, gestalt, and computing: studies in cognitive and systematic musicology*. Berlin: Springer Verlag
- Krumhansl, C.L., & Shepard, R.N. (1979) Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 579-94.
- Krumhansl, C. L., Louhivuori, J. Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectancy in Finnish folk hymns: convergence of behavioral, statistical, and computational approaches. *Music Perception*, 17, (151-197)

*Appendix*

- Krumhansl, C. L. Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (1998, 1999 and 2000). Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks, *Cognition* 76 (13-58)
- Kvifte, T. *Om flertydighet I opplevelse av metrum.* (manuscript) Oslo 1983
- Lande, V., (1983). *Slåttar og spelemenn i Bygland.* Oslo: Norsk Folkmusikksamling
- Large, E.W. & Kohlen, J.F. (1994). Resonance and the perception of musical meter. *Connection Science*, 6, 177-208.
- Lee, C. S. (1985). The rhythmic interpretation of simple musical sequences: towards a perceptual model. In P. Howell, I. Cross and R. West (Eds.), *Musical Structure and Cognition*, London
- Lee, C. S. (1991) The Perception of Metrical Structure: Experimental Evidence and a Model. In Representing Musical Structure, P. Howell et al. (Eds). (pp. 59-127). Academic Press, London.
- Ledang, O. K., (1966). *Song Syngemåte Stemmekarakter*, Oslo. Universitetsforlaget
- Lerdahl, F. & Jackendoff, R., (1983). *A generative Theory of Tonal Music*, Cambridge, Mass. The MIT Press
- Levy, M. (1983/1989). *The World of the Gorrlaus Slåtts*, Acta Ethnomusicologica Danica No. 6, Köbenhavn: Köbenhavns Universitet
- London (2002). Cognitive Constraints on Metric Systems: Some Observations and Hypotheses. *Music Perception* 19:529-550
- Longuet-Higgins, H. C. and Lee, C. S. (1984). The Rhythmic Interpretation of Monophonic Music. *Music Perception*, 1:424-441.
- Longuet-Higgins, H. C. and Lee, C. S. (1982). The Perception of Musical Rhythms. *Perception*, 11:115-128.
- Longuet-Higgins, H. C. (1976/1987) The Perception of Melodies. In Mental Processes Studies in cognitive science. The MIT Press, Cambridge (Ma).
- Merriam, A. P. (1963). The purpose of ethnomusicology, an anthropological view. I: *Ethnomusicology* 4, (206-213)
- Merriam, A. P. (1964). *The Anthropology of Music*. Evanston: Northwestern University Press.
- Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press

- Meyer, L. B. (1973). *Explaining Music*. Berkeley: University of California Press.
- Michon (1978). The making of the present: a tutorial review. In J. Requin (Ed.), *Attention and performance VII* (89-111). Hillsdale, NJ: Erlbaum
- 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. (1959). Om Vallåtar II Musikaliska strukturproblem. Särtryck ur *Svensk Tidskrift för Musikforskning*, årgång 41 , Stockholm: Isaac Marcus Boktryckeri AB.
- Monohan, C. B. & Hirsh, I. J. (1990). Studies in auditory timing: 2. Rhythm patterns. *Perception and Psychophysics*, 47, 227-242.
- Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realisation Model*. Chicago: The University or Chicago Press
- Narmour, E. (1992). *The analysis and cognition of melodic complexity: The implication-realization model*. Chicago: University of Chicago Press.
- Narmour, E. (1999) Hierarchical Expectation and Musical Style. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 441-472). New York: Academic Press.
- Nattiez, J. J. (1972). Is a Descriptive Semiotics of Music Possible?, *Language Sciences* 23, Bloomington.
- Nattiez, J. J. (1990). *Music and Discourse: Towards a Semiology of Music*. Princeton University Press, Princeton.
- Nettl, B. (1956). *Music in Primitive Culture*. Cambridge, Harvard University Press
- Nettl, B (1964). *Theory and Method in Ethnomusicology*. London, Macmillan Ltd
- Nketia, J. H. K. (1974) *The Music of Africa*. New York: Norton
- Palmer, C. & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 15, (331-346).
- Parncutt, R. (1987). The Perception of Pulse in Musical Rhythm. In A. Gabriellson (Ed.), *Action and perception in rhythm and music* (pp. 127-138). Stockholm: Royal Swedish Academy of Music.
- Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach*, Heidelberg: Springer Verlag

*Appendix*

- Parncutt, R. (1994) A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception*, II (4): 409-464.
- Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 3-18.
- Povel, D. J. and Essens, P. (1985) Perception of Temporal Patterns. *Music Perception*, 2:411-440.
- Powers, H. S. (1980). Language models and musical analysis. *Ethnomusicology*, January 1980
- Rosch, E. (1978) Principles of Categorization. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum
- Ramsten, M., (1982). *Einar Övergaards folkmusiksamling*. Stockholm: Svensk visarkiv
- Rosenberg, S. (1986) *Lisa Boudrés sångliga och melodiska gestaltning i tre visor. En analys av folkligt sångsätt*. Uppsats, Musikvetenskapliga institutionen, Stockholms Universitet
- Rosenberg, S (1994). *Visor i Gästrikland*, Gävle: Länsmuseet i Gävle,
- 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
- Sachs, C. (1953). *Rhythm and tempo : a study in music history*. New York
- Sachs, C. (1962). *The wellsprings of music*, edited by Jaap Kunst, The Hague : Nijhoff
- Schellenberg, E. G. (1996). Expectancy in melody: tests of the implication-realization model. *Cognition*, 58 (75-125)
- Schellenberg , E. G. (1997). Simplifying the implication-realization model of melodic and harmonic processes. *Music Perception*, 14 (295-318)
- Scabolci, B. (1956). *Bausteine zu einer Geschichte der Melodie*. Budapest 1959.
- Seeger, C. (1977). On the moods of a music logic. *Studies in Musicology 1935-1975*. Berkeley, CA
- Singer, A. (1974). The Metrical Structure of Macedonian Dance. *Ethnomusicology* 18.3: 379-404.
- Sloboda, J. A. (1985) *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Oxford University Press

- Spevak, C., Thom, B., Höthker, K. (2002) Evaluating Melodic Segmentation. In: Christina Anagnostopoulou, Miguel Ferrand and Alan Smaill, Ed. *Music and Artificial Intelligence. Proceedings of the Second International Conference ICMAI 2002*, LNAI 2445, (168-182) Springer-Verlang
- Sundberg, J. (1980). *Röstlära*, Stockholm: Proprius
- Temperley, D. (2001). *The Cognition of Basic Musical Structures*. Cambridge, MA, London, England: The MIT Press
- Temperley & Bartlette (2002). Parallelism as a Factor in Metrical Analysis. *Music Perception*, 20 (117-151)
- Tenney, J. & Polansky L. (1980). Temporal Gestalt perception In Music. *Journal of Music Theory*, 24:2 (205-241).
- The Fiddle Tunes Volumes*. Norwegian collection of folk music at the University of Oslo. (see also Gurvin 1959)
- The dens, H-H. (2001) *Untersuch den ganzen Mann – so wie er vor Dir steht: Der Spielmann Salve Austenå*. Oslo 2001
- Thom, B., Spevak, C., Hötkher, K. (2002) Melodic segmentation: evaluating the performance of algorithms and musical experts. In Nordahl (Ed): *Proceedings of the 2002 International Computer Music Conference* (65-72). Göteborg, Sweden
- Thompson, W. F., Cuddy, L. L. & Plaus, C. (1997). Expectancies generated by melodic intervals: evaluation of principles of melodic implication in a melody-completion task. *Perception & Psychophysics*, 59, (1069-1076)
- Thompson, W. F., & Stainton, M. (1998). Expectancy in Bohemian folk song melodies: evaluation of implicative principles for implicative and closural intervals. *Music Perception*, 15, (231-252)
- Todd, N. P. (1994b). Metre, grouping and the uncertainty principle: A unified theory of rhythm perception. *Proceedings of the Third International Conference for Music perception and Cognition* (pp. 395-396). Liege, Belgium: ICMPC.
- Trehub, S. E., (1987). Infants' perception of musical patterns. *Perception & Psychophysics*, 41, 635-41
- Tversky, A. (1977). Features of Similarity. *Psychological Review*, 84(4):327-352.
- Tversky, A. and Gati, I. (1978). Studies of Similarity. In . E. Rosch and B.B. Lloyd (Eds.) *Cognition and Categorisation*, London: Lawrence Erlbaum Associates.

## *Appendix*

- Valkare, G. (1987). *En model för överföring av interval mellan kulturer*, unpublished manuscript
- Wallin, N. L., (1982). *Den Musikaliska Hjärnan, en kritisk essä om musik och perception i biologisk belysning*, Skrifter från musikvetenskapliga institutionen, Göteborg: 6, Stockholm: Kungliga Musikaliska Akademien
- Wertheimer, M. (1923). Laws of Organization in Perceptual Forms. In W. D. Ellis, (Ed.) *A Source Book of Gestalt Psychology*, London, Routledge & Kegan Paul, 1938
- Wiora, W., (1941). Systematik der musikalischen Erscheinungen des Umsingens. *Jahrbuch für Volksliedforschung*, Berlin
- Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology*, 14, (1-66)
- Övergaard, E. (1935) (manuscript, ULMA 11642)

## **Sources of Music Examples**

- Abbogi* south Indian raga transcribed from sa-re-gam –and notated by K. Shivakumar 1992 (private/ authors manuscript)
- A Foggy Day* by Ira Gershwin and George Gershwin © Gershwin Publishing Corp., 1937
- Alfanska Racenica* , traditional Bulgarian tune (manuscript)
- '*Blue Rondo a la Turke*' by Dave Brubeck (my transcription)
- Boléro*, Maurice Ravel 1928 © Durand
- Bridal March* played by Per Danielsson (Svenska Låtar, Jämtland, Andersson no 21)
- Cetvorno* Traditional Balkan tune, Bulgaria (private/ authors manuscript)
- Fascinatin' Rhythm* by George Gershwin, (Gershwin 1924) private manuscript
- Folk song from Rumania*, sung by Susana Crăciunescu, Hunedoara, Rumania, transcribed by M. Rabinovici (Dragoi 1959:19)
- Folk song from Rumania* sung by P. Ispas, Hunedoara, Rumania, transcribed by E. Comisel (From Dragoi 1959:27)
- Fugue XV* of "Das wohltemperiertes Klavier" (BWV 860) by J.S. Bach (1722).
- Habanera*, from Carmen (opera in 4 acts), G Bizet

- Halling* after Ole Evju, Eggedal (from “The tune collection”, no 160 k)
- Halling* from Inger Aalls collection
- Hemlig stad jag* (Svenska Låtar Dalarna 1 Andersson 1922:130-131)
- Indiana Jones theme* melody, abstracted from a composition by John Williams for the motion picture “Raiders of the lost Ark” (Paramount pictures 1981).
- Jesu Bleibet meine Freude* the *obligato* melody From Cantata no 147, Herz und Mund und Tat und Leben BWV 147), J S Bach (1723)
- Norafjells* played by Andres Rysstad (Levy 1983)
- Partita in B minor for Violin solo*, the *Courante* from (BWV 1002), J.S. Bach (1720)
- Piano Sonata in A major*, (K 331, 300i) W.A. Mozart (1781-1783)
- Polska* no 83, after Bleckå Anders Olsson, Orsa (N. Andersson 1921).
- Polska* no 4, Per Danielsson, (Svenska Låtar, Jämtland, Andersson)
- Polska no. 10* played by Per Danielsson, (Svenska Låtar, Jämtland, Andersson 1926:10).
- Polska no. 11, second section* of played by Per Danielsson, CSvenska Låtar, Jämtland, Andersson 1926:11)
- Polska no. 13* played by Per Danielsson. (Svenska Låtar, Jämtland, Andersson 1926:12).
- Polska no 13* played by Per Danielsson (Svenska Låtar, Jämtland, Andersson V:13)
- Polska no. 46, second section*, variation, played by Bengt Bixo. (Svenska Låtar, Jämtland, Andersson 1926:28)
- Polska* played by Bengt Bixo and Per Danielsson. (Svenska Låtar, Jämtland, Andersson)
- Raag Suddha Danyasi*, transcribed from sa-re-gam –and notated by K. Shivakumar 1992 (private manuscript)
- Skjallmöjsslajé* gangar ‘ (Lande 1983 p 208)
- Symphony No.9* (Op.125) L. v Beethoven (1822-24)
- Symphony in G minor, No. 40* (K 550), W.A. Mozart (1788)

## *Appendix*

*Sonata G-minor Violin Solo*, Presto movement, (BWV 1001), J S Bach (1720)

“*Stormy Weather*” (Koehler/Arlen). “Vocal Real Book” version

*Sordölen*, performed by Eivind O. Hamre, Norway. (Lande 1983:116)

*Waltz* played by Per Danielsson, Mörsil, Jämtland, Sweden (From Svenska Låtar, V:18)

-----

*Seven Swedish polska melodies* from the repertoires of Gössa Anders Andersson and Spak Olof Svensson

*Ten Norwegian Halling/Gangar and Rull melodies*  
transcriptions made by e.g. Arne Björndahl, Morten Levy and myself.

*Thirteen popular Balkan melodies* notations from different origin (Bulgaria, Macedonia and Serbia)  
(provided by P. Landin 2003)

*Five South Indian ragas*, notated by K. Shivakumar

## **Recordings**

Andersson, Gössa Anders, *Hambreuspolskan* Polska SVL 104, Musica Sveciae/ Folk music in Sweden. Caprice CAP

Austenå, *Salve Kjempe-Jo*, (Austenå 1991)

Boudré, Lisa (1935) “*Falska klaffare*”, Landsmålsarkivet I Uppsala ULMA 155:3

Edvardsson Johansson, Karin (19XX) “*Kulning/berding call*” (# 3a) from “Lockrop & Vallåtar/ Ancient Swedish Pastoral Music”. Musica Sveciae/ Folk music in Sweden. Caprice CAP 21483

Fitzgerald, Ella “*Fascinating rhythm*” by G. Gerschwin, (Verve records 1962).

Landin, Petter, *Cetvorno*. Bulgarian tune. (private recording)

Rosenberg, Susanne (1991) “*hemlig stod jag*”, from “hemlig stod jag”, Rotvälta. Siljum 1991 BGS-CD 9121

Rysstad, Andres K. *Norafjells* (rec. 1973)

## **Software and Technical Equipment**

*Macintosh Common Lisp* version 2.0-4.1, © Apple Computer 1989-1994,  
© Digitool Inc. 1995-1997

*Finale*, version 2.6.3, 3.7, 2001 and 2002 © 1987-2002 by Coda Music Technology, a division  
of Net4Music Inc.

*Melodyne* 1.5.2 © Celemony Software GmbH 2000-2002

*Logic Platinum version 4.8.3* © 2002 Emagic Soft- und Hardware GmbH

*Proteus 1* Synthesizer Module

*Powerbook G4, 1Ghz, Apple computer*