JYVÄSKYLÄ STUDIES IN HUMANITIES 9

Tuomas Eerola

The Dynamics of Musical Expectancy

Cross-Cultural and Statistical Approaches to Melodic Expectations

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa (S212)
joulukuun 5. päivänä 2003 kello 12.

Academic dissertation to be publicly discussed, by permission of
the the Faculty of Humanities of the University of Jyväskylä,
in the Auditorium S212, on December 5, 2003 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2003

The Dynamics of Musical Expectancy

Cross-Cultural and Statistical Approaches to Melodic Expectations

JYVÄSKYLÄ STUDIES IN HUMANITIES 9

Tuomas Eerola

The Dynamics of Musical Expectancy

Cross-Cultural and Statistical Approaches to Melodic Expectations

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2003

Editors
Matti Vainio
Department of Music, University of Jyväskylä
Pekka Olsbo, Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

Cover Picture: Tuomas Eerola

URN:ISBN 9513916553
ISBN 951-39-1655-3 (PDF)

ISBN 951-39-1602-2 (nid.)
ISSN 1459-4331

Jyväskylä University Printing House, Jyväskylä 2003

ABSTRACT

Eerola, Tuomas
The dynamics of musical expectancy: cross-cultural and statistical approaches to melodic expectations
Jyväskylä: University of Jyväskylä, 2003, 84 p.
(Jyväskvlä Studies in Humanities
ISSN 1459-4331 ; 9)
ISBN 951-39-1655-3
Diss.

Melodic expectancy is a field of research that provides ways to understand the processes and knowledge people use in structuring, interpreting, remembering and performing music. Expectancy encapsulates the temporal nature of music, in which pitch and temporal patterns create dynamically changing expectations for ensuing events. Central issues included evaluation of the expectation models and their dependence on data-driven and schema-driven processes, assessment of melodic complexity and similarity and the development of a dynamic approach to measuring and modelling expectancy.

Cross-cultural comparisons were conducted to explore the role of data-driven and schema-driven knowledge in the formation of melodic expectations by using probe-tone experiments and predictability ratings. The musical excerpts consisted of Finnish spiritual folk hymns, North Sami yoiks, European folk melodies and African folk songs. Participant groups varied in their familiarity with the musical styles in question and consisted of Sami, Finnish, Central European and South African musicians. The contribution of data-driven musical properties towards melodic similarity and continuous predictability ratings were also investigated.

The results support the view that music draws on common psychological principles of expectation, but that cultural background and stylistic knowledge shape the contribution of these principles. A range of data-driven principles of expectations explained the responses of all groups, although the responses of non-experts more closely reflected event-frequency models and Narmour's implication-realization model (1990) whereas the experts' responses exhibited reliance on schema-driven and style-specific models. Moreover, frequency-based melodic features accounted for a proportion of listeners' similarity ratings and a dynamic model combining data-driven and schema-driven melodic information was able to account for listeners' continuous ratings of melodic predictability.

Keywords: expectancy, music, cognition, cross-cultural, similarity, complexity, psychology, modelling

Author´s address Tuomas Eerola
Department of Music
University of Jyväskylä
P. O. Box 35 (M)
FIN 40014 University of Jyväskylä

Supervisors Professor Petri Toiviainen
Department of Music
University of Jyväskylä
P. O. Box 35 (M)
FIN 40014 University of Jyväskylä
FINLAND

Professor Carol Krumhansl
Department of Psychology
Cornell University
Ithaca, NY 14850
USA

Professor Jukka Louhivuori
Department of Music
University of Jyväskylä
P. O. Box 35 (M)
FIN 40014 University of Jyväskylä
FINLAND

Reviewers Professor Ian Cross
Faculty of Music
University of Cambridge
11 West Road
Cambridge CB3 9DP
ENGLAND

Docent Mari Tervaniemi
Cognitive Brain Research Unit
Department of Psychology
P. O. Box 9
FIN 00014 University of Helsinki
FINLAND

Opponent Professor Ian Cross
Faculty of Music
University of Cambridge
11 West Road
Cambridge CB3 9DP
ENGLAND

PREFACE

I am not the first person to wonder how "sheep's guts should hale souls out of men's bodies" (Benedick in William Shakespeare's *Much Ado about Nothing*). I started my undergraduate studies in Jyväskylä partly seeking an answer to this mystery and was quickly drawn into cognitive musicology, one of the three specialization topics available within the musicology programme at the Department of Music. At the beginning of my graduate studies, I recognised that the issues covered by music psychology and music cognition would provide answers to questions that are relevant from the common listener's perspective, challenging from the musician's point of view, and attainable as scientific activity.

During my years as a postgraduate student I was lucky to benefit from several outstanding advisors. Professor Carol Krumhansl introduced me to the world of empirical endeavours and showed that working in this field is great fun as well. Professor Petri Toiviainen's brilliant, analytic and insightful comments encouraged, supported and guided my scientific adventures. In addition to their words, they both also set an example of broadening one's musical skills whilst pursuing an academic career. Professor Jukka Louhivuori introduced me to the wonders of the scientific world. He was essential in providing financial and mental support for my graduate studies and also prompted me in the direction of cross-cultural issues. I am grateful to the reviewers of my work, Professor Ian Cross and Docent Mari Tervaniemi, for their valuable comments on this thesis. I also wish to thank Professor Matti Vainio for having faith in me in my current position as an assistant in musicology. He also acted as an editor for the publication series in which this study was agreed to be published (*Jyväskylä Studies in Humanities*).

Most of my work has been carried out at the Department of Music, first as a postgraduate student (1997), then as a research assistant (1998) when I was fortunate to work with Professor Carol Krumhansl, then as a graduate student (1999-2001) at the *Pythagoras Graduate School of Sound and Music Research* and finally as an assistant in musicology (2002-2003). I want to thank all the staff members at the Department of Music for being such terrific colleagues. To single out some of them is easier said than done but Topi Järvinen gave good comments and argued for the use of proper tools for each task. Pekka Toivanen gave me healthy reminders of the necessity to incorporate cultural context and musical practice into the research of musical cognition. Kirsti Hämäläinen provided collegial support and merriment to the daily work at the Department.

There have been many others who have contributed to this work by providing good atmosphere and support. I thank Adrian North, Lorraine Sheridan and Mike Beauvois from the Leicester University, England, for their friendship, sharp observations and promoting the importance of having a life outside academia during my visit to Leicester in 1998-1999 (funded by the

Finnish Academy). I am especially grateful to Adrian for showing me how to say the maximal amount with a minimum of energy.

My visit to Cornell University in 2001 allowed me to broaden my knowledge of music cognition considerably, strengthened my conviction of empirical and rigorous methodology and helped me to gain new perspectives on academic life in the real world. Joel Snyder gave numerous helpful comments regarding the real-time project and not only introduced *Matlab* and *MAX* to me, but guided me patiently through my first attempts using these powerful tools. Another Cornell graduate student, Erin Hannon, provided energetic remarks on my research ideas.

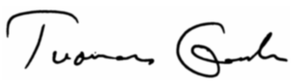
Edward Lebaka from Pretoria, South Africa, had a crucial role in facilitating the work in South Africa as well as providing a wealth of information on native African traditions. Annukka Hirvasvuopio-Laiti from Tampere University acted as an expert in Sami traditions and she graciously and skilfully guided our expedition to Kautokeino, Norway. Tommi Himberg helped me to carry out experiments in South Africa and proved to be a reliable collaborator with an excellent sense of humour.

My colleagues at the Pythagoras Graduate School of Sound and Music Research (funded by the Finnish Ministry of Education), especially those at Jyväskylä, Jan-Markus Holm and Francis Kiernan, helped me to laugh at various nerve-wracking things that one undergoes during ones graduate school career. Also, the teachers of the graduate school posed many constructive questions and remarks without whom this thesis would have ended up being more restricted in its scope and style.

Furthermore, my gratitude goes to Donald Adamson, David Underwood and Micah Bregman, who proof-read parts of the text.

My biggest gratitude is without a doubt to my wonderful wife, Päivi-Sisko. She put up with my long hours at work, travels around the globe, and the anxiety of paper deadlines. She also acted as a manager for the most of our trips abroad, charmed the locals while I stayed at the laboratory after office hours, and took up the duties of editor-in-chief during the final stages of writing.

Jyväskylä, 28 October, 2003



CONTENTS

LIST OF PUBLICATIONS

List of publications (reprinted after the introductory part), which are included in this thesis:

I Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception, 17*, 151-196.

II Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks, *Cognition, (76)1*, 13-58.

III Eerola, T., & Louhivuori, J. (submitted). Expectancy in North Sami yoiks revisited: The role of data-driven and schema-driven models in the formation of melodic expectations.

IV Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (submitted). Perceived complexity of Western and African folk melodies by Western and African listeners.

V Eerola, T., Järvinen, T., Louhivuori, J., & Toiviainen, P. (2001). Statistical features and perceived similarity of folk melodies. *Music Perception*, *18*, 275-296.

VI Eerola, T., Toiviainen, P., & Krumhansl, C. L. (2002). Real-time prediction of melodies: continuous predictability judgments and dynamic models. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, J. Renwick (Eds.) *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 473-476). Adelaide: Causal Productions.

SUMMARY OF PUBLICATIONS

I The first publication examined melodic expectations in terms of schematic and data-driven knowledge. The former consisted of typical patterns in Finnish hymns and schematic Western tone transitions and the latter were frequency-based models and the implication-realization model developed by Narmour (1990). A probe-tone experiment containing excerpts from Finnish spiritual folk hymns was conducted for listeners either familiar (expert) or unfamiliar (non-experts) with hymns. The results indicated that a range of underlying psychological, data-driven principles of expectations explain the responses of both groups although non-experts' responses reflected more simple frequency-based models whereas expert listeners' ratings could be explained using schema-driven models. Statistical style analysis and simulation of melodic expectations with neural network models were also used in the analysis.

The author was responsible for a part of the statistical analysis, data collection and analysis.

II The second publication studied melodic expectations using North Sami yoiks and three groups of listeners (experts, semi-experts, and non-experts). The results portrayed a continuum between the schema-driven and data-driven models in relation to familiarity with yoiks. The results generally supported the predictions from the implication-realization model of Narmour (1990) although the relative importance of the principles varied across styles, and even listeners unfamiliar with the music showed sensitivity to this variation. Moreover, event-frequency models and schema-driven models (Western generic and yoik) explained different portions of expectations, depending on the listeners' backgrounds. In sum, these results support the idea that music draws on common psychological principles of expectation, but that cultural background and stylistic knowledge shape the contribution of the principles.

The author was partly responsible for the statistical analysis of yoiks, data collection, and data analysis.

III The third publication reanalyzed the material used in publication two: North Sami yoiks. Empirical results from a non-Western group of listeners (South African traditional healers) indicated that the data-driven models are robust in explaining the expectations regardless of the cultural background of the listener. Event-frequency models were more successful than an auditory memory model and Narmour's implication-realization model (1990), although support was obtained for certain principles in the latter model, supporting its culture-transcending qualities.

The author carried out the major part of the work, including the data collection and analysis, the modelling and the writing.

IV The fourth publication examined the schema-driven and data-driven processes in music perception by comparing melodic complexity ratings given by Western and African listeners to Western and African folk songs.

The results illustrated no large overall differences between the groups although certain rhythmic aspects of the melodies reflected culturally determined differences between the groups.

The author was responsible for the dominant part of the work, including experimental design, data collection, analysis, the majority of the melodic complexity models and the writing.

V The fifth publication studied the contribution of data-driven, frequency-based musical properties in accounting for the melodic similarity in folk songs. Similarity of frequency-based musical properties accounted for a moderate proportion of listeners' similarity ratings and a slightly better predictive rate was achieved using descriptive variables such as rhythmic and melodic predictability.

The author was responsible for the major part of the work, including experimental design, data collection and analysis, the similarity models and the writing.

VI The sixth publication developed a new methodology and novel models for melodic expectations. Listeners' continuous predictability ratings were collected in two experiments, one containing folk melodies and the other artificial sequences. Listeners' continuous ratings were better accounted for by dynamic models that combined data-driven and schema-driven, frequency-based melodic information than rule-based models due to their better context sensitivity.

The author carried out the central part of the work, including the data collection and analysis, dynamic models and writing.

1 INTRODUCTION

Expectancy has a central role in all forms of human behaviour, including perception, planned motor movements, speech production, and speech comprehension, that continuously shape our reactions to events in the real world. For example, a stroll through the centre of town involves countless expectations concerning the behaviour of other pedestrians and, more importantly, predictions about the movements of the motor vehicles. Similarly in human relationships, expectancies consisting of general social norms and cues provided by social interaction mount up and provide the conditions in which we fall in and out of love.

While many kinds of anticipatory events are frequently experienced in the ordinary situations of life, listening to music is especially beset with patterns that play with listeners' expectations, forming conflicts, tension, even jarring surprises along with their resolution in many pleasurable ways. The role of expectancy as a fundamental musical process has therefore been acknowledged and has received much attention in the psychology of music. (e.g., Meyer, 1956; Carlsen, 1981; Jones & Boltz, 1989; Narmour, 1990; Krumhansl, 1995b). These studies have examined the ways the expectations are formed in Western music, especially with respect to the melodic aspects of music. The pervasiveness of expectations has also generated theoretical discussion about the origins of expectations in music and whether they are mostly learned from ones musical culture or derive from certain universal principles that transcend musical cultures. Take for example listening to speech. Finnish speech can be predictable, full of expected regularities, to a listener with previous experience of the language, making it possible to parse the auditory stream into relevant chunks, phonemes, words, and phrases, that make up the language. Without a long-term immersion in the Finnish language, expectations about auditory events occurring in Finnish speech may concern only such details as monotonic repetition of strange syllables and comparatively small deviations in the gross contour of speech. Despite not understanding the content of the speech, a listener without knowledge of the Finnish may still pick up the tone of voice, whether, for example, the speaker is calm, agitated or bored.

The topic of musical expectancy has a number of theoretical implications for music cognition and applications derived from music cognition. The main incentives to investigating expectancy in music can be summed up as follows:

1) Expectancy is a way to explore the processes and the knowledge listeners use to structure, interpret, remember and perform music.
2) Expectancy encapsulates the temporal nature of music, in which pitch and temporal patterns create dynamically changing expectations for ensuing events.
3) Expectancy is connected with aesthetic and emotional responses to music.
4) Expectancy features strongly in music theory, including Schenker's goal-directed harmonic and melodic progressions (1935/1979), Piston's notion of harmonic motion (1978), and the prolongational reduction studied by Lerdahl & Jackendoff (1983).

The main topic of the thesis is the formation of melodic expectations and how these expectations are mediated through different processes and types of knowledge. This topic is studied by contrasting the performance of listeners with different kinds of musical knowledge in various tasks connected with the phenomena of melodic expectancy, complexity and similarity. The theoretical considerations raised by the topic concern first the role of listeners' automatic expectations in music and how musical knowledge reflects these, as well as opening up the research in a less Western-biased direction. In addition, other approaches to expectancy formation, namely the perception of similarities in melodies and the dynamic, temporal nature of melodic expectations, are considered.

The first chapter outlines the approach taken in this thesis. The second chapter portrays the processes and scope of expectations and makes operational definitions of them with respect to music and its features. The third chapter contains selected examples of melodic expectations. In the fourth chapter, various empirical methodologies used to explore expectations are reviewed. The fifth chapter presents the aims of the individual studies and the sixth chapter summarizes the results of the studies in turn. In the final chapter, conclusions are drawn from the empirical findings.

1.1 Music cognition

The paradigm this thesis adopts for the study of expectancy is music cognition. Music cognition employs an experimental approach and computer modelling, the former borrowed from cognitive psychology and the latter from the field of artificial intelligence. Music cognition has its roots in the naturalist tradition (Helmholtz, Stumpf), which was transformed in the 1960's and 70's by a movement called cognitive revolution. This paradigm focused on mental processes – skills, perception, knowledge and motivation – and how they are mentally encoded, stored and retrieved (Gardner, 1987). For cognitive science music offers a virtually ubiquitous domain of human processing, that has attained a structural complexity that matches that of language, and which can be learned without explicit training at a very early age. In music cognition the listener's highly sophisticated perceptual and cognitive mechanisms for understanding, appreciating, and participating in musical activity are put under scrutiny (Leman, 1997; Laske, 1988). The paradigm has produced influential models of music perception and cognition (Lerdahl & Jackendoff, 1983; Krumhansl, 1990; Bregman, 1990; Sloboda, 1985). In this paradigm, notational, acoustic and perceptual perspectives on musical behaviour are taken into account although it is acknowledged that the definition of music requires a broader context, where the interaction between listener, performer, sound and the context creates musical experiences (Carterette & Kendall, 1999).

The field of artificial intelligence brought into being the notion of a computational theory of mind, which had a profound impact on the cognitive research paradigm. Although this notion is often taken to represent any cognitive system that can be implemented as a program on a computer, this does not mean that the mind literally works like a computer, consisting of hardware, software, processor and memory although the two latter terms are frequently encountered in cognitive literature. The fundamental significance of the computational theory of mind lies in its ability to explain complex human actions (knowing, thinking, listening, etc.) as the product of "mindless physical process" (Pinker, 2002, p. 33). On a more practical level, description of a theory as a formal, computerised model "provides a standard of rigour and completeness to which theoretical explanations should aspire (which is not to say that a program in itself is a theory)." (Boden, 1990, p. 108). This notion is still subscribed to today in cognitive research, including research in music cognition (Handel, 1989; Cook, 1999; Leman, 1997), although the output of a computer simulation is considered to be valuable only if the computer model is itself based on an empirically approved theory of underlying mechanisms (Leman, 1995).

What can be represented by a computational model of music is another significant topic in music cognition. It is vital to carefully define the details of mental representation in music cognition, since they are essential components of a sensory system whose input is entirely transitory. Information about acoustic events must be collected across time, attended to dynamically, and processed with respect to previous experience while new information is being received. In the early days of music cognition, a symbolic approach – where separate symbols such as pitch or chord are used to represent music – was taken in the majority of studies. Whilst still at the heart of many music cognition models, symbolic representation has since been challenged by subsymbolic representation, where perceptual processes are represented by signal processing models and neural networks (Leman, 1993; Camurri & Leman, 1997; Griffith & Todd, 1988; Bharucha, 1991; Toiviainen, 1996).

Representation leads to another hallmark of music cognition, namely gathering empirical data from experiments. The experimental method, especially the use of controlled experiments, is a powerful methodological tool applied when certain aspects of musical behaviour are investigated in a controlled setting by varying certain elements of interest while holding many extraneous features stable. Although it is possible to infer causal relationships in a strictly controlled experiment, the very constraints the experimental method imposes upon the research topic pose challenges as to how well the results apply to the real world. However, music cognition seeks out convergent evidence drawing upon different types of experiments and samples and it also gathers evidence across disciplines and cultures to understand which factors contribute essentially to the observed facts. For the present thesis, converging evidence was sought across cultures, an approach known as *cross-cultural music cognition*. This approach is relevant in studying the sources of expectations in music.

Developmental psychology is another approach that may be used to uncover the sources of expectations in music. This approach has already yielded results describing several human predispositions for processing music and illustrated how culture-specific schematic processes in music develop (for a review, see Trehub & Trainor, 1993; Trehub, Schellenberg & Hill, 1997; Trehub, 2000). However, the cross-cultural approach will be presented in more detail next, as it is the approach taken in several studies in this thesis.

1.2 Cross-cultural music cognition

In order to study the role of possible innate and learned factors in processing music, it is necessary either to investigate listeners who have little acculturation (infants) or else to compare the responses of listeners from culturally separate musical backgrounds. In theory, a comparison between those who are heavily encultured within a musical tradition (i.e., musicians) and those who are less so (non-musicians) falls within the same approach. For example, it can be assumed that musicians have more fine-tuned musical schemata than non-musicians but the differences in innate, automatic processes are likely to be rather small (Cuddy & Lunney, 1995; Krumhansl 1995). Then again, cross-cultural music cognition experiments usually involve participants from distinct and separate cultural backgrounds.

There have been many attempts, on philosophical, theoretical and practical levels, to investigate the universality of innate predispositions and the learned particularities of individual cultures and environments. In psychology, these approaches have taken two strands, cross-cultural psychology and cultural psychology. The proponents of cultural psychology (for example Cole, 1996; Shweder, 1990), hold that it is not enough to study culture as an independent variable, which is usually the case in cross-cultural psychology. However, the contributions of cultural psychology have been so far mostly philosophical and theoretical, and have charted the ways of incorporating cultural contexts into research. Cross-cultural psychology, on the other hand, has opted for a more empirical approach, and dealt with cultural contexts in a different way. In this approach, the cultural background of the participants has been the independent variable, and the aim has been to explore what exactly in the cultural background causes the variation in the dependent variable, and not just to be content with the result that it is something called "culture". Regarding the aesthetic responses of people from different cultures, cross-cultural psychology has provided evidence for both substantial differences and substantial similarities (Russel, Deregowski & Kinnear, 1997).

Cross-cultural studies of music have echoed findings on aesthetic responses in general. Cross-cultural studies of affective reactions to music have shown some examples of culturally-specific emotional responses. A study by Gregory and Varney (1996) portrayed how European and Indian listeners disagreed in describing the mood of Indian classical and Western classical music, which suggested that stylistic information largely determined the musical moods perceived by the listeners. Meyer, Palmer and Mazo (1998) and Balkwill and Thompson (1999) found influences of both schematic knowledge and culture-transcendent factors in their cross-cultural studies of Russian laments and Hindustani ragas, respectively. Other studies using the cross-cultural approach in music have focused, not only on melodic expectations and the structural organisation of tones, as will be described later, but also on beat finding (Stobart & Cross, 2000; Magill & Pressing, 1997; Toiviainen & Eerola, 2003), and on interval discrimination (Sampat, 1978).

Cross-cultural studies of music also focus upon the level of attention in cross-cultural music cognition, which underlines the importance of perceptual processes for deciphering acoustical cues and forming melodic and rhythmic expectations in music. Hardly anyone would attempt to claim that there is a cultural difference in how a single neuron operates within a listener's auditory cortex nor would it be wise to try to explain the diverse musical behaviour exhibited across cultures using a single, unified theory.

2 EXPECTATIONS

> ″A mind is fundamentally an anticipator, an expectation-generator″
> (Daniel Dennett, 1996, p. 57)

The philosopher Daniel Dennett's succinct appraisal of the function of the mind highlights the crucial role that expectations play for an organism. Expectations are hypotheses about the configurations underlying the real world. The machinery behind the expectations is an evolutionarily adaptive design for anticipating future events, which has helped us to react successfully to temporally organised events. Anticipating events even moderately accurately has had major biological advantages and the development of this system has been possible because of the regularities in our ancestors' environment. The importance of anticipatory mechanisms for humans is brought out by the fact that 8-9-month-old babies already have the capacity to generate expectations of future actions (Munroe, Munroe & Whiting, 1981). Because of the ubiquity of expectancy in human life, expectations are an important focus of psychological research, from describing the perceptual processes in language and vision to comprehending how social behaviour is constructed (e.g., Mandler, 1975; Olson, Roese & Zanna, 1996; Chun & Jiang, 1999).

With regard to terminology, the term *expectancy* refers to the general notion of anticipating something. According to Oxford English Dictionary (OED), expectancy refers to "the quality or state of being expectant" (OED, 1989). It is wider and more generic in scope than the other term used extensively in this work, namely *expectation*, which refers to ″the action of mentally looking for some one to come, forecasting something to happen, or anticipating something to be received; anticipation" (OED, 1989). *Expectation* is therefore more concrete and specific than *expectancy*. In the literature on musical or melodic expectancy, however, the use of these terms is not consistent. For example, in the 44 studies that employ the word expectation or expectancy in their titles, cited in this work (see references), 30 studies refer to expectancy and 14 to expectation without a common logic in the use of the terms (e.g., "expectation in music", Schmuckler, 1989 and "expectancy in melody", Schellenberg, 1996, but "modeling melodic expectation", Larson, 1993 and "measuring melodic expectancies", Adachi & Carlsen, 1995). In this work, the term *expectancy* is reserved for a generic notion of what is to come, and the term *expectation* for specific notions of melodic continuations listeners have in mind when listening to music.

2.1 Expectation processes

A useful way to divide various factors that affect expectations is to refer to two types of process. The first, *data-driven process*, arises from the senses, while the second, *schema-driven process*, is a more complex cognitive process as it is affected by knowledge. It is also possible to distinguish certain processes that come between the two. These automatic processing tendencies are called *processing predispositions***.** These three terms are explained in the following section.

2.1.1 Data-driven and schema-driven processes

In a short-term time frame, perceptual information coming directly from the senses supplies cues to behaviour. This is called a *data-driven* (also known as bottom-up, sensory priming, adaptive) *process* in cognitive science. It is automatic and consists of basic principles of perceptual organisation. For example, the grouping of auditory events is achieved by pitch proximity and temporal contiguity. If these data-driven dynamics are beneficial for an organism and occur frequently enough, the knowledge they represent may be transferred into the long-term memory and may be used on future occasions. This process is called *schema-driven process*, which comprises learned behaviour, involving detection of familiar events, which then guide the organisation of relevant details. Of course the actual process of perception is a dynamic one, influenced by the interaction of both processes and this is often referred to as the perceptual cycle (Neisser, 1976). In this cycle, data-driven processes are guided by what we expect to perceive (schema-driven knowledge), which may in turn, be modified by data-driven knowledge.

There are various types of memory structure involved in both processes. *Sensory stores* hold information coming from the senses briefly, and operate in different ways for different senses. In the auditory domain this memory is called *echoic memory* and it is estimated to have a storage capacity of about 1 second (Darwin, Turvey & Crowder, 1972; Treisman & Rostron, 1972; also referred to as *auditory primal sketch*, Todd, 1994; Näätänen & Winkler, 1999). Auditory events can also be processed in the *short-term memory* (or *working memory*, Baddeley & Hitch, 1974, 1986), which is somewhat more flexible and has a larger capacity than the echoic memory. Its capacity has been described in terms of a time span of 2 to 8 seconds (Baddeley, 1986; Hitch, Halliday, Dodd & Littler, 1989) but more commonly it is specified by the number of items it can hold (corresponding to Miller's famous 7±2 items rule [1956]). In addition to acting as a working memory, the sensory stores have the effect of priming – facilitating – the processing of contextually related events. Priming effects have been documented largely in the domain of language (Cutler, Mehler, Norris & Segui, 1987; Forster & Davis, 1984; Dupoux & Mehler, 1990; Tulving & Schacter, 1990).

However, the capacity of the short-term memory is also dependent on the encoding of the items in it, which may be influenced by the *long-term memory*, itself virtually unlimited in its capacity and theoretically permanent in duration. This memory employs, among other things, abstract cognitive structures called *schemata* that encode generic knowledge, as summarized by Rumelhart (1980, p. 33):

> "Schemata are the building blocks of cognition. They are the fundamental elements upon which all information processing depends. Schemata are employed in the process of interpreting sensory data (both linguistic and non-linguistic), in retrieving information from memory, in organising actions, in determining goals and subgoals, in allocating resources, and, generally, in guiding the flow of processing in the system."

In 1932, Bartlett demonstrated how the schemata of typical stories influenced strongly the recall of a particular story and often caused it to be distorted. Schemata do not only work retrospectively but also influence expectancy (Mandler & Johnson, 1977). Reader's expectations of what is to come at the sentence level are governed by schematic knowledge (Taraban & McClelland, 1988). Schemata also influence perceptual processes by decreasing the need to evaluate all aspects of a visual scene (Schank & Abelson, 1977; Rumelhart, 1975) and organise new information in terms of past experience in order to facilitate anticipation (Neisser, 1976; Mandler, 1984). Schemata also have an obvious connection to *prototypes*, which are exemplary or idealized representatives of a class (Rosch & Lloyd, 1978), and important in structuring the environment into distinct classes. At this point, it is also necessary to separate *episodic memories* from schematic knowledge. Episodic memories represent the specific instances of events (Rumelhart, 1980), which will also affect expectations with respect to particular events. In music, expectations based on these instances are called *veridical expectations*, as explained later (Bharucha, 1987).

Whilst data-driven and schema-driven processes account for most of the processes involved in expectations, it is also possible to distinguish certain processes in perception which do not strictly correspond to these processes but which nevertheless have a significant effect on perceptual processing. These may be called *processing predispositions* and they will be briefly discussed next.

2.1.2 Processing predispositions

Because the regularities in the real world are bound to particular contexts it is also important for an organism to confine expectations to particular contexts (Cosmides & Tooby, 2000). If a certain context is especially stable over an evolutionary time period, an organism may develop a *predisposition* towards learning that particular regularity. Such a predisposition may be called an *innate* mechanism. Hence in evolutionary psychology innate and learned knowledge structures are not opposed to each other. Take for example three human predispositions, face-recognition, facial expressions, and communication by means of a language. Infants less than 10 minutes old prefer face-like patterns to scrambled versions of the same patterns (Johnson & Morton, 1991), and even newborn babies display the same facial expressions (Camras, Holland & Patterson, 1993). Language has also been demonstrated to be more of an innate property of the human mind than a fundamentally learned ability (Pinker, 1994) although learning is of course necessary to bring this capability into full operation. It follows from this that certain processes involved in structuring auditory signals, speech and music may have found their way into *permanent human predispositions*. Neuropsychological findings (Peretz, 2001; Tervaniemi et al., 1999) support the idea that there are specialized neural structures for musical processing, and this would have to be so in order to sustain any claim as to permanent predispositions in this domain, as these specialized neural structures need an evolutionary time period (i.e., tens of thousands of years) for their permanent organisation.

Examples of innate mechanisms in the auditory domain are reflexes such as the orienting response, where an unexpected loud noise produces physiological and neurophysiological changes that aid in collecting vital information from the environment and prepare the individual to take defensive action (Lang, Simons & Balaban, 1997). Research on infants has outlined several processing tendencies for music which demonstrate that babies are able to detect subtle changes in melody (Trainor & Trehub, 1992) and rhythm (Trehub & Thorpe, 1989), which suggests a biological basis for several aspects of musical processing (summary in Trehub, 2000). Moreover, the skills of trained and untrained listeners show more similarity than difference (Bharucha & Stoeckig, 1986, 1987; Cuddy & Badertscher, 1987; Regnault, Bigand & Besson, 2001; Trainor, Desjardins & Rockel, 1999) and there are many striking similarities between the acoustic codes used to convey emotions in different cultures (Gabrielsson & Juslin, 2003). All these observations imply certain musical predispositions that may also be evident in musical expectations. However, it might also be that the heuristics captured by innate principles for processing music merely reflect the structural properties of the music and speech that are learned by the listeners instead of being innate or "hardwired". As put by Bharucha, "If schematic expectancies can be easily learned by mere exposure, there is no reason for them to have been wired innately" (1994, p. 231).

The question of human predispositions relates to the scientific debate about the role of *nature versus nurture*, which has received a great deal of attention in many areas of human information processing in recent years. On the one hand there are scientists who subscribe to the "blank slate" doctrine, according to which learning completely shapes the human mind, while on the other hand there are those who argue for biological determinism and the genetic construction of behaviour. Most scientists nowadays would argue against the dichotomy of nature versus nurture, and advocate instead the idea that both in some way or another interact to produce the constraints of human behaviour. For instance, the cognitive scientist Steven Pinker suggested that "behavior may vary across cultures, but the design of the mental programs that generate it need not vary" (Pinker, 2002, p. 40). Pinker's studies are mostly about language, which he says we should consider as an innate, specialized mechanism of the human mind that adapts to the particular language of the infant's cultural surroundings (Pinker 1994). This view, which in cross-cultural psychology is called *universalism*, is in sharp contrast with *absolutism* and *relativism*. The former regards human phenomena as essentially the same in all cultures while the latter holds the opposite view, and holds that cultures cannot even be compared meaningfully (Berry, Poortinga, Segall & Dasen, 1992). The position taken in this thesis is that of universalism, seeing the study of common mental mechanisms as a fascinating area of research into musical behaviour in view of the fact that large differences exist in notions and features of music across cultures.

The very same debate concerning nature versus nurture is apparent in the music research community. In the humanities, ethnomusicologists, music anthropologists and musicologists mainly approach music as a cultural activity, and therefore give more weight to the influence of culture and learning (Becker, 2000; Feld, 1982; Herndon & Mcleod, 1981; Davidson & Torff, 1992; Feld, 1984; Feld & Keil, 1994; Walker, 1996; Gourlay, 1984) although there are significant exceptions (Merriam, 1964; Blacking, 1995). In the field of music cognition, researchers deal more with perceptual and developmental issues relating to music and are therefore more open to the idea of culture-transcendent, universal mechanisms than researchers in the humanities (e.g., Carterette & Kendall, 1999; Trehub & Trainor, 1993; Papousek, 1996). Even so, there are those within the field of music cognition who are less inclined to talk about human universals in music (Walker, 2002; Cook, 1994).

Even if the background of musical expectancy and predispositions can be explained in terms of evolutionary psychology and biology, it does not necessarily follow that music has an important value for survival, although some researchers have argued in favour of this idea (Brown, Merker & Wallin, 2000; Brown, 2000). If minds anticipate all types of environmental stimuli, the same mechanisms apply to behaviours that have no direct survival value. The development of innate predispositions for organising expectations about the environment is not plausible when the environment contains a high degree of variance (e.g., a particular language or a particular musical style). In that case, a better evolutionary strategy is to form expectations through learning, as originally outlined by Baldwin (1896). Consequently, it is more probable that most musical expectations correspond to this category and are acquired through learning. This issue about innateness and learning will be reconsidered when different sources of melodic expectations in music are described in detail.

2.2 Expectations in music

One of the founders of ethnomusicology and comparative musicology, Erich M. von Hornbostel, regarded the study of music as a combination of ethnology, anthropology and psychology. Exactly a hundred years ago, he pinned down the relationship between the arousal of attention and expectation in music (Abraham & Hornbostel, 1903, as cited in Blum, 1991, p. 16-17):

> "Music arouses attention by various means. It attracts attention by a very strong, very high, very low tone, by constancy and shifting of tone quality. At present the ultimate reasons are unfathomable. Besides the aforementioned, secondary criteria also play a role in attracting attention […]. If we hear the sounds a b c d, we reproduce previous impressions connected with the continued alphabet, so if we expect anything, we expect the continuation of the alphabet; likewise in music."

Although we are not much wiser now as to the ultimate reasons why music arouses attention, we can agree with Hornbostel that music seems to be a game of expectations. Listeners effortlessly form expectations of the forthcoming events in music, even if they are not familiar with a particular musical piece. This occurs due to the regularity of musical structure, consisting of hierarchical and temporal relations that have been recognized, for example, in music-theoretical treatises (Piston, 1941/1978; Schenker, 1906/1954; Schoenberg, 1911/1978; Meyer, 1956, 1973; Narmour, 1990, 1991, 1992). Meyer laid the groundwork for the study of expectancy by stating that expectations consist both of features arising from musical stimuli and learned, culturally specific stylistic codes (1956, 1973). Both sources of knowledge are employed in creating dynamic changes of tension that may give a partial answer to Hornbostel's question of how Western music arouses attention. Changes in musical tension lead to gratifying releases, which are occasionally delayed, that may form an essential part of our emotional responses to music (Bissell 1921; Meyer, 1956). Violations of regularity trigger autonomic nervous system arousal that activates further cognitive activity in a search for meaning. The meaning, when determined, merges with the arousal of the experienced emotion. For example, Sloboda (1991) observed that the emotional peaks in music identified by listeners were associated with unexpected changes in music, such as a change of key or the entry of a new instrument. Also, locations of high musical tension display changes in the syntactic structures that also rely on the expected properties of music (Krumhansl, 1996). At a more simple level, expectations affect memory of melodies (Schmuckler, 1997), and detection of wrong notes in melodies (Janata, Birk, Tillmann & Bharucha, 2002).

In sum, there is an important connection between expectancy and the sustaining of attention in music. The way this comes about is achieved is reviewed next in connection with the processes involved in listening to music.

2.2.1 Expectation processes in music

Listening to music brings together numerous automatic processes involved in structuring auditory stimuli and the interpretation of these percepts in the light of learned, stylistic knowledge. Almost a half a century ago, Leonard Meyer distilled the essence of these two different processes:

> "The probabilities of style and form, the norms upon which expectancies rest, differ from culture to culture and style to style. What remains constant in the flux of music history is not any particular organization of the materials of sound. The patterns of style are fixed by neither God nor nature but are made, modified, and discarded by musicians. What remains constant is the nature of human responses and the principles of pattern perception, the ways in which the mind, operating within the framework of a learned style, selects and organizes the sense data presented to it" (Meyer, 1956, p. 73).

For example, the data-driven processes in music consist of such heuristics as determining the sound source according to frequency proximity and spectral similarity. As regards expectancy, listeners constantly adjust their expectations according to what has occurred before, using the data-driven processes at work in various types of short-term memory. In music, these priming effects are common. Harmonic context, for example, primes the processing of target chords (Bharucha & Stoeckig, 1986, 1987; Tekman & Bharucha, 1998). Moreover, findings on pre-attentive processing of auditory information using brain recordings[1] suggest that even intricate musical stimuli are processed automatically and pre-attentively in the auditory cortex (Tervaniemi, 1999, 2001). This processing is related to echoic sensory memory but also to syntactic processes on a larger level and it is also mediated by training, since musicians have been proven to exhibit more definite responses (Koelsch, Schroger & Tervaniemi, 1999; Regnault et al., 2001; Tervaniemi, 2001).

Schema-driven processing of sound is learned behaviour and involves the identification of familiar acoustic patterns from which processing mechanisms construct a hypothesis about a stimulus, which then guides the organisation of relevant auditory details (Bregman, 1990, p. 397). Consequently, this schematic information in music requires learning the common patterns of one's musical culture, and storing these in the long-term memory. The proper application of this knowledge is activated through data-driven processing. Hence, both processes play a vital role in the processing of musical signals (Bigand, Parncutt & Lerdahl, 1996; Regnault, Bigand & Besson, 2001) and the two may also form conflicting expectations.

Although musical expectations are probably predominantly learned, they are mostly unconscious and processed automatically. This calls for an additional expression to differentiate those occasions when music conforms to schemata and is familiar from those when expectations in music are explicit and consciously known. Bharucha (1987) calls this *veridical* expectancy. It represents the specific knowledge of a particular work that the listener is familiar with (episodic memory). For example, in thinking of the opening theme of the Beethoven's Fifth symphony one's expectation is veridical. In contrast, when one expects a subdominant chord to move to the dominant and then resolve to the tonic, one's expectation is of a more general type, that is, schematic. In schematic expectancy, expectations are influenced by other music with which the listener is familiar whereas in veridical expectancy, expectations are influenced by the knowledge of a particular piece of music.

Bharucha and his colleagues (1999) have shown that veridical listening is nevertheless heavily influenced by schematic knowledge (also Plamondon, 1995; Schmuckler, 1989). For example, a deceptive cadence may still evoke a surprise response in physiological terms, even when a listener is certain of its impending occurrence. This inconsistency is explained by the modularity of these processes. In vision and language research, the concept of modularity is used to account for the way learning is accomplished by mechanisms particular to the domain in question and way the modules function independently of other modules (Fodor, 1983). Similarly in music, Jackendoff (1987) has argued that the schematic and veridical processes are separate and disconnected from each other. Hence, our veridical knowledge of a musical work does not disrupt the formation of more general, schematic expectations. To take the earlier example, the schematic system is surprised when the deceptive cadence occurs even though the listener knew (veridical memory) that it was coming.

Finally, on top of our schematic knowledge of, for example, Western music we also have more precise schematic knowledge of the different styles of music that we are familiar with (e.g., Eurovision song-contest style, Early Mo-

---

[1] Electro-, magnetoencephalographic and positron emission tomographic investigations, summarized in more detail in chapter 4.

zart style, industrial techno, etc.). We can label this encapsulation of schematic knowledge of particular musical styles as *stylistic knowledge*, which enables us to maintain the different expectations for different musical styles separate from each other. At this point, it is useful to summarize the different levels of expectations mentioned above.

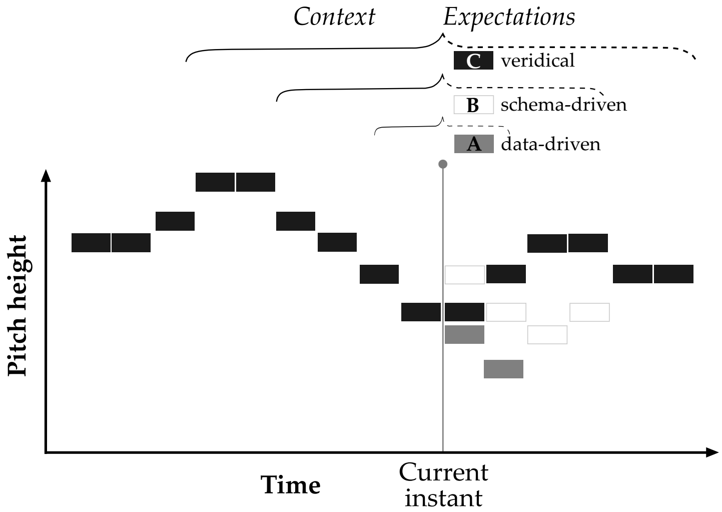


FIGURE 1 A schematic illustration of different levels of expectations in a familiar melody.

In Figure 1, a familiar musical context with a steady temporal structure is shown in graphic format for visual emphasis. The black boxes denote the musical context (sequence of notes) and the dotted line represents the current instant at which point three possible expected continuations are given (the black, white and grey streams of notes). The sources of expectations, as explained above, are divided into data-driven processes, schema-driven processes, and veridical knowledge. Data-driven expectations encompass the acoustic information obtained from the senses and certain rule-based models accounting for the organisation of the melodic continuations. These data-driven expectations are mostly based on a short-term context. In Figure 1 the continuations generated by data-driven principles illustrate a continuation of the descending scalewise movement (A), which follows the Gestalt principle of good continuation. This possible continuation is just an example of certain rule-based expectations and not the short-term sensory priming. Schema-driven expectations generated by the context (B) follow the principles of phrasing, typical melodic archetype and typical harmonic frame, which in this case drive the expected tones towards the tonic of the sequence, forming a typical cadence pattern. The schema-driven expectations typically need a larger context than the data-driven expectations. For example, to create an expectation of typical harmonic and tonal progression, a sense of key must be established and this is usually based on a larger context than just a few preceding notes. The veridical expectation, which represents the knowledge of how this famous little tune actually continues, will firmly steer the expectations towards the remembered, correct continuation (C). In this case, the actual continuation of the melody repeats the interval pattern and the melodic contour, aspects which make the actual continuation even more predictable and complete. The temporal context of the veridical expectations is typically more extensive than the one needed by the other two processes as the whole melody can be remembered once it has been recognized.

It should be noted, however, that although all these processes are different and each process may have different rules, principles, and models within it, most of the expectations are generated automatically and unconsciously by the listeners.

2.2.2 Musical features and expectations

There are many musical devices available for composers and performers to occasionally confound the expectations of the listeners while still keeping them sufficiently on the track of the music, which means that there are numerous musical features that listeners may be on the alert and generate expectations for. In Western music, these features are often divided into melodic, temporal, harmonic, structural and timbral categories, of which a detailed review has been given by Narmour (2000). Moreover, listeners possess different expectations for different kinds of musical styles that may have distinct schemata for pitch-, and temporal structures, as documented by several empirical studies (Krumhansl, 1991; Leman, 1995; Handel, 1989).

This work focuses on melodic expectations, a central aspect in Western music, and especially on pitch-related aspects of melodies although there are also other important musical features that contribute to Western listeners' melodic expectations, such as harmony and phrasing.

Ever since Rameau (1722) put forward his theory on the subject, harmony has received considerable attention in music theory (Riemann, 1895; Piston, 1978) and in music psychology. For example, Cuddy, Cohen, and Mewhort (1981) showed how harmonic structure influences the perception of tone sequences (also Povel & Jansen, 2002). Also, global and local harmonic expectations are known to influence listeners' responses (Bigand, Madurell & Pineau, 1999; Tillmann, Bigand & Pineau, 1998). The question of sensory priming is important in harmonic expectancy, as Western harmonic principles correlate with the psychoacoustic structures of sounds (Parncutt, 1989; Bigand, Parncutt & Lerdahl, 1996) and their influence on harmonic expectations has been observed (Bigand & Pineau, 1997).

In music theory, attention has been given to structural expectations, such as large-scale tonal progressions (Schenker, 1935/1979) or large-scale periodicities and "hypermeters" (Lerdahl & Jackendoff, 1983). However, the plausibility of these large-scale expectations has been called into question by empirical studies in which musical segments have been reordered without this having any detrimental effect on listeners' preferences or recognition (Gotlief & Konecni, 1985; Karno & Konecni, 1992; Cook, 1987). On the other hand repetition and forms such as rondo form in classical music or popular song forms with, for example, four- or eight-bar phrases organized in structures such as AABA or AABC, create more plausible structural expectations that are based on clear repetition of musical material and occur on a shorter time-scale than large-scale expectations.

Phrasing also guides expectations. Phrase endings and beginnings are usually marked by stylistic devices, based on harmonic or melodic conventions, which impart clues about phrase structure to listeners. Both temporal and pitch information are used to give these cues and even infants are sensitive to phrasing information (Krumhansl & Jusczyk, 1990; Jusczyk & Krumhansl, 1993). The role of phrase structure in forming expectations has been directly studied by Boltz (1993) who found that deviant pitch changes were more easily detected at phrase endings, which means that listeners expect certain regularities at phrase endings, and regular phrasing aids recognition and melody recall (Boltz, 1991; Chiappe & Schmuckler, 1997). Phrasing and segmentation, however, are difficult to infer and model from a musical sequence due to the different emphasis of local structures (rests, durations) and higher-level structures (e.g. recurring motives, harmony, melodic parallelism) as well as style-dependent norms. The current models of segmentation mostly deal with local structures (Tenney & Polansky, 1980; Cambouropoulos, 1997; Bod, 2002; Temperley & Sleator, 1999). Tempo also creates a sense of expectation and steady alterations in tempo (accelerando and ritardando) are tracked by adjusting the internal clock of the listeners (Vos, van Assen & Franek, 1997). Other types of temporal expectations will be addressed later.

Listeners also assume that instrumental timbres and instrumentation will be constant after they have been defined or that changes in, for example, the texture will occur when more instruments are brought in to the music. Hence, expectations concerning the timbre, instrumentation and texture exist. So too it may be with dynamics (volume) as it is with texture that gradual changes lead into further interpolation of the change by the listener. Expectations are also reflected in the expressive timing and expressive dynamics of performances, especially with respect to phrasing (Palmer, 1996). Finally, it may be that expectations are generated by lyrics and other extramusical phenomena or features of the situation such as visual information. When the percussionist picks up the cymbals we certainly expect them to meet shortly with a sonorous clash. This work, however, is restricted to investigating the basic way musical expectations operate in the melodic domain.

2.2.3 Representation and acquisition of musical expectations

The basic components of melody such as pitch and rhythm need to be represented in some way. In the cognitive science, the ways in which knowledge is represented are essential objects of study (Anderson, 1983; Fodor, 1983). The visual system, for example, makes use of multiple parallel representations of various features – location, colour, contrast, shape, lines, orientation, motion, etc. – present in the visual signal (Marr, 1982). In music cognition, a wealth of work has considered the nature of musical knowledge and its representation (Huron, 1992; Honing, 1993; De Poli, Piccialli & Roads, 1991; Clarke, 1988; Baily, 1985; Leman, 1999). Despite the lack of general agreement on the precise nature of these representations, *discrete* elements, such as tones, scales and chords are often used (McAdams, 1989) and the *continuous* aspects of music, such as waveforms, dynamics and tempo curves are used in the study of expression, timbre, and the performance attributes of music (Honing, 1993; Clarke, 1987) as well as sub-symbolic accounts of musical processing. The epistemological nature of discrete and continuous representations of music merits further comment. The discrete representation can be called atomistic, as it refers to small particles that cannot be divided further. However, a non-atomistic epistemology holds that musical knowledge is the result of a continuous processing of sensory information, often called non-symbolic or sub-symbolic representation (Leman, 1995). At the sub-symbolic level, neural mechanisms detect the invariant and discriminant features of the stimulus. For example, the emergence of tonal centres has been modelled using this approach, where first the low-level perceptual systems are modelled and the emergence of higher level phenomena is explained in terms of self-organizing neural networks (Leman & Carreras, 1997; other high level musical phenomena by Toiviainen, 1996; Bharucha & Todd, 1989). However, discrete, event-based models populate musicological theories as well as cognitively-oriented models. In this work the majority of the models used belong to the discrete, event-based category since the existing literature on music cognition has provided a wealth of theoretical and empirical accounts based on these concepts.

Even if the scope of representation is limited to discrete, event-based representations, there is still the question of what is represented and how many representations are needed to provide even a cursory account of musical processing. For example, absolute pitch-classes, intervals, temporal information, phrasing, accents and melodic contour can all provide clues about musical structure for a listener (e.g., Dowling, 1978) and these, as multiple parallel representations, have been employed in new ways of measuring melodic similarity (Toiviainen & Eerola, 2002). Thus, one way of representing musical knowledge is to express it in terms of probabilities, such as the probabilities of the occurrence of individual tones or specific chord progressions in baroque music. The existence of multiple representations and representational levels may generate conflicting expectations between the different representations along with asynchrony in expectations generated by different hierarchical levels (Sloboda, 2000). For example, the leading note in a cadence may be highly expected but its timing may be postponed to create a deviation from the expected temporal structure.

The acquisition of schema-based knowledge occurs by learning, also understood as a *cultural transmission* on a larger level. Cultural transmission refers to a process in which a "cultural group perpetuate[s] its behavioral features among subsequent generations through teaching and learning mechanisms" (Berry, Poortinga, Segall & Dasen, 1992, p. 17). In cultural transmission, two modes of transmission are distinguished, *enculturation* and *acculturation*. The first is a fundamental and automatic socialisation process that takes place within one's own culture. *Acculturation* is learned culture that results from contact with cultures other than one's own (Berry et al., 1992). The acquisition of musical schemata may be studied by means of developmental and cross-cultural studies. Developmental studies of acquired culture have demonstrated how children acquire sensitivity to tonal-hierarchical relations (Krumhansl & Keil, 1982), with young children first differentiating between diatonic and non-diatonic pitches and later between more and less stable diatonic pitches. The results of subsequent studies (Cuddy & Badertscher, 1987; Lamont & Cross, 1994) suggest that the acquisition of tonal organisation is affected by musical training rather than simply materialising with increasing age. Also infants' spontaneous production of melodies may be analysed in order to understand the development of musical schemata (e.g. Mitroudot, 2001). Developmental studies of melodic expectations are rare and these have mainly demonstrated that the dominant model of melodic expectations (Narmour, 1990, detailed later) has predictive accuracy even among listeners aged between 5 and 11 years old (Schellenberg, Adachi, Purdy & McKinnon, 2002; also Adachi & Carlsen, 1994, 1995). In the present thesis, the process of learning or development will not be studied as such but cross-cultural comparisons are used to evaluate the part played by different types of schema-driven and data-driven processes in melodic expectations.

3 MELODIC EXPECTATIONS

Melody, defined as "pitched sounds arranged in musical time in accordance with given cultural conventions and constraints" (Ringer, 2003, ¶ 1), is a universal human phenomenon that features prominently in most kinds of music around the world and can be traced back to prehistoric times. Melodic expectations address the question "what" and "when", that is, what tones tones are expected to occur and when in a given sequence of music. The first question concerns pitch and the second the temporal structure of the sequence.

During the last two decades, a number of empirical studies of music have focused on melodic expectations, mainly with regards to the first question, that of pitch (Bharucha, 1984, 1987; Dowling, 1990; Jones, 1982, 1990; Krumhansl, 1990; Lerdahl & Jackendoff, 1983; Schellenberg, 1996, 1997; Schmuckler, 1989; Aarden, 2002; von Hippel & Huron, 2000). Schema-driven processes in the formation of expectations have been studied more than data-driven processes. While the focus of these studies has mostly been on Western classical music using musically experienced listeners as participants, it is assumed that musically naïve listeners have acquired comparable representations of the regularities in Western music and that their processing of music is largely similar to that of musicians.

In the following, melodic expectations are presented in relation to the underlying processes, whether data-driven or schema-driven. These are, in turn, divided into pitch-related and temporal sources of expectations, although the distinction does not do justice to the fact that both sources interact in the perception of melody as well as in the formation of expectations (Boltz & Jones, 1986; Deutsch & Feroe, 1981; Monahan, Kendall & Carterette, 1987; Palmer & Krumhansl, 1987, 1990). Nevertheless, division into pitch and temporal dimensions is warranted in the light of the neuropsychological evidence, which has shown numerous cases of selective loss of either component (Peretz, 1990; Peretz & Kolinsky, 1993). It should also be noted that while it is possible to refer separately to schema- and data-driven processes, both are constantly involved in the perception of music and in some cases, it is difficult to draw a line between them. Tonality offers an example from the musical domain that combines schema-driven and data-driven processes (Parncutt & Bregman, 2000).

3.1 Data-driven expectations

Data-driven expectations are based on the sensory traces related to musical events accumulated in the sensory memory, which in turn influence listeners' anticipation of what is to come. We can separate these processes into (a) short-term auditory priming, (b) auditory stream segregation, (c) sensitivity to frequency of occurrences, and (d) rule-based heuristics of melodic continuations.

3.1.1 Pitch-related expectations

Short-term auditory priming. The role of short-term auditory priming in the formation of expectations has been relatively little studied (Huron & Parncutt, 1993). Leman (2000), for example, has used an auditory memory model to study melodic continuations to the classical chord and melodic cadences (Krumhansl & Kessler, 1982). The model is based on the workings of the auditory system. The auditory periphery and basilar membrane are first simulated and further processing occurs using the autocorrelation technique and leaky integrators, representing the two types of short-term memory, *sensory memory*, extremely brief in duration and episodic in nature and *echoic memory*, which is integrated over a longer period of time. These two memories are compared to obtain a prediction of the acoustic fitness of current pitch to the context.

Auditory stream segregation. One of the fundamental aspects of melody perception has been studied in the form of auditory stream segregation, a feature of auditory scene analysis, explained in detail by Bregman (1990). Auditory stream segregation is greatly influenced by many aspects of the so-called *Gestalt laws* (good continuation, similarity, proximity, etc.) and depends on automatic functioning of the auditory system making non-conscious guesses about the nature of the sources of sounds in the real world. In auditory stream segregation, a sequence is considered as a single auditory stream that has to be extracted from the environment. When the tones of a sequence are close to each other in time and in pitch, the sequence is perceived as one stream but when the pitch intervals grow larger and the tempo faster, the probability of perceiving two streams is greater. The separation of these streams is governed by the principles of *fission boundary* and *temporal coherence boundary* (Miller & Heise, 1950; van Noorden, 1975; Bregman & Campbell, 1971). Existing music can be demonstrated to follow these two rules closely; for example, the distribution of interval sizes in samples of folk music shows that small intervals predominate (Merriam, Whinery & Fred, 1956; Vos & Troost, 1989). Dowling (1967) demonstrated how pseudopolyphony is created in baroque solo works by consistently exceeding the fission boundary. Huron (2001) also showed how the majority of tone durations in musical compositions fall within the temporal coherence boundary. The association of single streams with close pitch proximity may be a result of the limitations in vocal production systems (von Hippel, 2000).

The principles of auditory stream segregation have implications for melodic expectations as they provide heuristics for predicting what kind of sequences are perceived as a coherent melody. For example, melodic leaps may form a possible second perceptual stream and the cohesiveness of the original stream is best maintained if the following interval returns to the boundaries established by the original stream.

Sensitivity to frequency of occurrence. Another data-driven explanation of melodic expectations draws on listeners' sensitivity to the frequency of occurrence of tones, intervals, and durations, which has also been shown to influence the formation of melodic expectations (Oram & Cuddy, 1995; Krumhansl, 2000; von Hippel, 2002; Cuddy, 1993). These event-frequency heuristics are often presumed to transcend musical cultures, but this view has been put in question by cross-cultural experiments using different styles of music. In these studies listeners from separate musical cultures usually evaluate the fitness of various possible continuations of the melodies (Carlsen, 1981; Unyk & Carlsen, 1987; Castellano, Bharucha & Krumhansl, 1984; Krumhansl, 1995b). Castellano et al. (1984) demonstrated how Indian listeners' ratings of Indian music could be predicted from their schematic knowledge of Indian music plus short-term statistical properties related to the excerpts, whereas American listeners' ratings of the same music reflected only the statistical properties of the excerpts. Similar findings have been obtained in studies concerning Balinese music (Kessler, Hansen & Shepard, 1984), Korean Court music (Nam, 1998), and serial 12-tone music (Krumhansl, Sandell & Sergeant, 1987). In the last study, listeners unfamiliar with serial music rated the fitness of continuation tones based on the tones that had already occurred in the contexts. On the other hand, the listeners familiar with 12-tone music rated the tones actually missing from the contexts as more fitting, consistently with the pitch structure of serial music.

As well as by simple tone profiles, expectations are also easily generated by higher-order statistical patterns in music. For example, tone- and interval-transitions (first-order probabilities) influence both adult and infant listeners, as demonstrated by Saffran and her colleagues using both language-like sequences and tone-sequences (Aslin, Saffran & Newport, 1998; Saffran, Johnson, Aslin & Newport, 1999). Frequency-based heuristics are also common components of prediction in other domains than music and language (Kelly & Martin, 1994; Hasher & Zacks, 1984). Besides tones, similar priming of certain patterns also occurs with chord sequences (Bigand, Madurell, Tillmann & Pineau, 1999; Bigand & Pineau, 1997). Moreover, recurrent higher-order pitch patterns, melodic motifs, will cause listeners to expect them (Jones, 1987; Deliège, 1996; Zbikowski, 1999). Narmour (2000) has provided the most full-blown account of how repetition and variation is achieved by means of cognitive rule-mapping, which involves pattern repetition. It is true that exact repetition is commonplace in music (ostinato, groundbass, canon, etc.) but repetition must be defined in terms of similarity, which is much more difficult to frame in cognitive terms than as it has been used in the vocabulary of music theory (Hewlett & Selfridge-Field, 1998). One way to measure melodic similarity is to use the statistical properties of melodies (tone, interval, duration profiles and their higher-order variants) as the element of similarity (detailed in Study V).

Narmour's implication-realization model. The first description of tone-to-tone expectations was presented by Meyer (1956, 1973). Later, Narmour (1990,

1992) described in more detail how these expectations could work in music theory. Narmour's *implication-realization* model (1990) describes tone-to-tone expectations for melodic continuations and Narmour holds that parts of these processes are innate, automatic, and require no learning. The predictions of the model are expressed in terms of a small number of principles relating to interval size and direction.

The central terms in the model are the *implicative interval* and the *realized interval*. When there are three successive tones (i.e., two successive intervals), the implicative interval is the first and the realized interval is the one that follows. The implicative interval *implies* that some tones are more likely to follow (realized) than others. These three-tone patterns have been formalized into five principles*: registral direction, intervallic difference, registral return, proximity*, and *closure*. These principles contain simple rules for the size and direction of the realized interval in relation to the implied interval. A quantification of these principles is illustrated in Figure 2. For example, the principle of *registral direction* states that a small implicative interval (up to 5 semitones) implies a continuation in pitch direction, whereas a large interval implies a change in direction. The principle of *intervallic difference* states that a small implicative interval implies a similar-sized realized interval whereas a large implicative interval implies a relatively smaller realized interval. The principle of *registral return* states that the second tone of the realized interval is proximate in pitch to the first tone of the implicative interval. The principle of *proximity* states that there is a greater expectation of small realized intervals than large ones. Indeed, melodies all over the world show a tendency to favour small intervals (Merriam, Whinery & Fred, 1956; Dowling, 1967; both mentioned in Huron 2001, which also demonstrates the prevalence of small intervals using samples from various music cultures). This is in line with Narmour's claim and the principles of auditory stream analysis, detailed earlier. The principle of *closure* describes a change in pitch direction and whether the realized interval is smaller than the implicative interval. *Closure* is also dependent on other factors such as duration and the metrical position of tones as well as harmony. Finally, the principle of *consonance* was added by Krumhansl (1995b), in which the consonance (fusing or blending of the two tones in an interval) of the realized interval is calculated according to an averaged set of empirical and theoretical values of consonance, as summarized in Krumhansl (1990, p. 57).

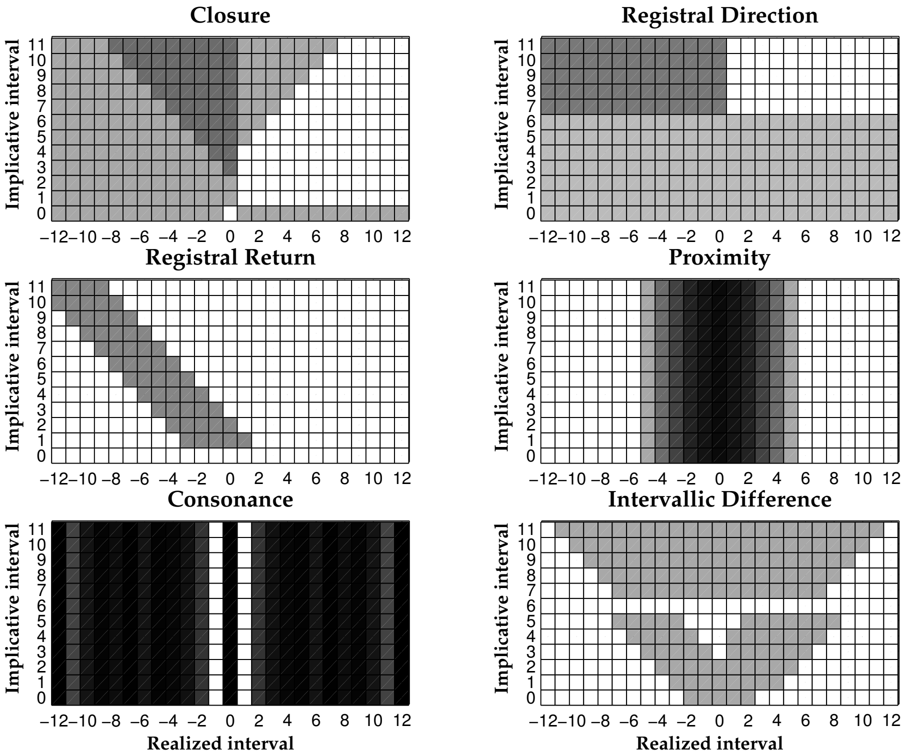


FIGURE 2 Quantification of the Narmour's implication-realization model (Krumhansl, 1995b). The vertical axis corresponds to the implicative interval, ranging from 0 to 11 semitones. The horizontal axis corresponds to the realized interval, ranging from 12 semitones in the opposite direction from the implicative interval (-1 to -12) to 12 semitones in the same direction (as the implicative interval, +1 to +12). The shaded grids indicate the combinations of implied and realized intervals that fulfil the principle. The darker the shading the better, the more expected, the realization of the implied interval.

Figure 2 displays all six principles of Narmour's model using grid representation, where implicative interval size is shown on the Y-axis and the realized interval size and direction is shown on the X-axis. The full implication-realization model is more comprehensive, including processes and contextual effects, especially relating to metrical and harmonic factors, and has been further extended by Narmour (1999, 2000). However, much empirical work has been conducted using these principles of melodic continuation, and as a result the number of principles and their implementation have since been modified and improved by several authors (Krumhansl, 1995a, b; Schellenberg, 1996, 1997; von Hippel & Huron, 2000; von Hippel, 2000; Thompson & Stainton, 1998). Take for example one of these recent revisions of Narmour's model by von Hippel (2000). This revision is related to restrictions of melodic range and brings in the principles of *tessitura* and *mobility*. The former predicts that forthcoming tones will be close to median pitch, while the latter uses autocorrelation between successive pitches to evaluate whether the tone is predictable in relation to the previous intervals and mean pitch. Larson (1993, 1994, 1997, 2002) argues that melodic expectations are governed by physical parameters or analogies such as *gravity*, *magnetism* and *inertia*. However, similar principles have been included in revisions of Narmour's model that lend themselves more easily to empirical testing.

The claim that the principles of Narmour's model reflect, innate, 'hard-wired' and automatic heuristics, may be too bold, as the principles could very well be products of non-musical acoustic regularities, such as speech intonation. For example, the expectation of small melodic intervals may be associated with the physical constraints of the sound source (e.g., vocal cords). However, some support for the universality of the principles of Narmour's model has been obtained in studies where the principles of the model predict the responses of listeners in the same way regardless of familiarity with the musical style (Krumhansl, 1995b; Schellenberg, 1996, 1997; Cuddy & Lunney, 1995). For example, Krumhansl (1995b) used excerpts from British folk songs, atonal songs, and Chinese folk songs to which different possible continuations were rated by listeners according to how well they agreed with their expectations. In the cases of the Western musical excerpts, judgments made by musicians and non-musicians were compared. In the case of the Chinese excerpts, judgments made by native Chinese and American listeners were compared. The results supported the principles of the model, and were consistent with its claims concerning universality, since the differences between the groups were relatively small. Hence, the predictive power of the principles does not appear to depend on extensive training in music (Thompson, Cuddy & Plaus, 1997).

Nonetheless, style-dependent factors affect how much variance in the expectancy ratings by listeners there is left for the implication-realization model to explain. For instance, with atonal melodies (Krumhansl, 1995b; Schellenberg, 1996; Roh & Yi, 2000), the predictive power of Narmour's principles has proved weaker than with tonal melodies (Chinese or British folk songs). Moreover, in comparisons between children and adults it has been noticed that the model more successfully predicted the adults' expectations (Schellenberg et al., 2002), which suggests that the model reflects learned knowledge rather than innate.

3.1.2 Temporal expectations

In addition to the question of *what* future events might occur, it is important to know *when* the events might occur. In a concise summary of the universal processes in temporal organisation of auditory sequences, Drake (1998) distinguishes two basic psychological processes: "the segmentation of an ongoing sequence into groups of events on the basis of their physical characteristics and the extraction of an underlying temporal regularity or pulse" (p. 13). In the first process sequences are grouped in terms of changes in the parameters. For example, a temporal gap in the sequence creates a boundary that terminates the perceptual unit and signals a new one. This grouping process is assumed to be universal, and present even in early infancy (Krumhansl & Jusczyk, 1990). The second basic process involves the extraction of temporal regularities from the sequence. Listeners commonly expect musical events to occur at regular intervals, sometimes known as the referent period, of which the most salient is approximately at 600-ms interonset interval (Drake & Baruch, 1995; Fraisse, 1982; Parncutt, 1994; van Noorden & Moelants, 1999; Clarke, 1999). This process is also assumed to be universal and it appears early in infants (Baruch & Drake, 1997). This summary is in accord with a review by Carterette and Kendall (1999, p. 780), who also list grouping strategies, the use of reference pulses and subdivision of time pulses as the most plausible cognitive musical universals that relate to the temporal structuring of music.

The music-theoretical notions of rhythm, beat and metre have received a great deal of attention in music theory and cognitive modelling. However, when the discussion is concerned with data-driven processes relating to rhythm, many music-theoretical terms are probably not valid as they are constructed especially for Western music culture and require the acquisition of its typical temporal patterns. However, the terms related to the two above-mentioned basic processes are also detailed in the theory of Lerdahl and Jackendoff (1983). They describe the organisation of events into groups and how the sense of metre is produced. The latter involves "series of perceived pulses marking subjectively equal units in the temporal continuum" (Large & Kolen, 1994, p. 182). Metre, therefore, refers to the measurement of the number of beats between regularly recurring accented events. The perceptual accentuation of the events depends on many different factors (see summary in Parncutt, 1994).

Western music theory also holds that metrical organisation is hierarchical (Lerdahl & Jackendoff, 1983; Cooper & Meyer, 1960). However, it seems unlikely that the Western notion of metrical hierarchy transcends cultural boundaries. As Magill and Pressing (1997) noted, the asymmetric timeline-ground model based on African understanding of rhythm corresponded better with the beat of a Ghanaian percussionist than the Western model. Also Stobart and Cross (1994, 2000) have questioned the universal applicability of the concepts of Western music theory in their studies of Bolivian Easter songs. Despite the implausibility of hierarchical metrical organisation as a general organizing principle of temporal structure in music, Temperley has argued that the preference rule system, drawn from the theory by Lerdahl and Jackendoff, is also applicable to African music (2000), popular music, and jazz (2001). These issues will be addressed in more detail in describing schematic temporal expectations.

The models of temporal aspects of music produced since Lerdahl and Jackendoff's work are often based on such basic cognitive processes as *dynamic attending* (Boltz, 1993; Jones & Pfordresher, 1997), *entrainment,* where the stimulus rhythms serve to synchronize a perceiver's internal rhythms, usually modelled as hierarchically nested oscillators (Jones & Boltz, 1989; Large & Kolen, 1999), and *sensory-guided action schemes* (Todd, Lee & O'Boyle, 2002). Pressing (1983) has theorized about the common cyclic structures of the pitch and rhythm in music using group theory and provided applications of the notion to jazz, West African, and Balkan music.

Furthermore, there is an interaction between pitch and rhythm. For example, memory recall of melodies is degraded if pitch and temporal patterns are out of phase (Boltz & Jones, 1986; Monahan, Kendall & Carterette, 1987). Harmonic information also affects the perception of metre (Dawe, Platt & Racine, 1994).

3.2 Schema-driven expectations

Knowledge of stylistic conventions permits listeners to recognize similarity between percept and memory and hence to correlate their learned, schema-driven expectations with current percepts. Stylistic information also helps to carve the musical input into distinct chunks (phrases, motifs, etc.), which is economical in the cognitive sense. Most music-theoretical writings concern the nature of stylistic information and how it influences our notions about pieces of music (Meyer, 1989; Narmour, 1992, 1999). The exact terms used to describe stylistic information vary, for example, from schematic clusterings (Gjerdingen, 1988), archetypes (Meyer, 1973), to style structures (Narmour, 1990, 1992). However, this knowledge of regularities in music, represented by schemata, is learned and is usually tacit, that is, "a matter of habits properly acquired (internalized) and appropriately brought into play" (Meyer, 1989, p. 10). Although a variety of these terms are used in music theory, the fundamentals of the terms stem from cognitive psychology, and especially from schema theory, reviewed earlier. The information represented by schematic knowledge in music has been modelled using various probabilistic models, derived from empirical experiments and style analyses but also by neural networks (Bharucha & Todd, 1989; Gjerdingen, 1992; Tillmann, Bharucha & Bigand, 2000). Schematic knowledge has been demonstrated to influence a wide variety of processes, including musical preferences (North & Hargreaves, 1997; 1999; Tekman & Hortaçsu, 2002), key-finding (Krumhansl, 1990), and remembering music (Krumhansl, 1979; Boltz, 1991).

3.2.1 Pitch-related expectations

A number of perceptual studies have outlined the role of schema-driven, pitch-related information in the perception of music. For example, DeWitt and Samuel (1990) showed the effects of schema-driven expectations in a task where the participants were required to distinguish between two types of items: those in which a target tone was removed from the recording and replaced with noise, and those in which noise was merely added to the signal. Participants were more accurate in detecting the target tone when the melody was either familiar or predictable. Similarly, Bey and McAdams (2002) examined the role of schema-based processes in the perceptual organisation of tone sequences by presenting the listeners with target melody interleaved with a distractor sequence and asking whether the probe melody was identical to or different from the target melody. If the listeners had heard the melody previously, their recognition was markedly better. Hence, expectancy may improve perceptual processing by directing the attention to particular points in time and pitch (also Dowling, Lung & Herrbold, 1987; Jones & Boltz, 1989).

Tonal hierarchy. Western tonal hierarchy acts as a strong schematic influence in remembering and producing music (e.g., Krumhansl, 1979; Boltz, 1991). For example, listeners tend to hear chromatically inflected melodies as departures from diatonic schemata (Bartlett & Dowling, 1988; Cohen, Thorpe & Trehub, 1987). Tonality and tonality induction has been represented by the tonal hierarchy theory and the concept of *key-profiles* that stand for the stability of the twelve pitch-classes relative to a given key, where the tonic is the most stable, followed by the other tones of the tonic triad, then the diatonic tones, and the nondiatonic tones being the least stable (Krumhansl, 1990). The tonal hierarchy defines the relationships between the tones, which act as cognitive reference points that have different probabilities, which in turn influence melodic expectations. For example, a sequence of G_3 E_4 E_4 D_4 D_4 creates a strong expectation of C_4 but hardly any expectation of, say, $A\sharp_3$ (the numeral specifies the octave; C_4 = middle C). This hierarchy of stability is learned through exposure to a particular style of music, since infants do not possess such schemata for tones. For example, infants can detect mistunings equally well in Western key modes and in Javanese pelog scales whereas older children and adults detect mistunings only in their native scale systems (Lynch & Eilers, 1991; Lynch, Eilers, Oller & Urbano, 1990). It has also been documented how the tonal hierarchy is acquired in different stages during childhood (Krumhansl & Keil, 1982; Speer & Meeks, 1985). Further evidence for the learning of tonal hierarchy by exposure is provided by the high correspondence between the distribution of tones in Western music and the tone profiles derived from the experimental data obtained by Krumhansl and Kessler in their probe-tone experiments (1982; noted in Krumhansl, 1990).

The schematic expectations created by the tonality have been studied in terms of neural networks (Bharucha & Todd, 1989; Leman, 1994; also Tillmann, Bharucha & Bigand, 2000; Toiviainen & Krumhansl, 2003). Moreover, psychoacoustically oriented models, which take into account pitch salience and sensory memory decay (Parncutt, 1989; Huron & Parncutt, 1993) and geometric models of tonality have also been proposed (Shepard, 1982; Chew, 2000; Longuet-Higgins & Steedman, 1971).

Western schematic expectations. Another schematic model of Western pitch-related expectations is represented by results from a study by Krumhansl (1995a), in which an interval was followed by a third tone, and listeners rated this tone in terms of how well it agreed with their expectations of what interval would follow. A wide variety of intervals (-11 to +11 semitones) and probe-tones (-12 to +12 semitones) were used. The pattern of results was similar to the results obtained in production tasks (Schmuckler, 1990; Thompson, Cuddy & Plaus, 1997). Pitch-related expectations are not only created between successive tones but they arise also at more abstract levels (Narmour, 1990, 1992). Krumhansl (1997) has demonstrated how melodic expectations are carried over short phrases.

Harmony. The temporal order of tones also influences the sense of tonality. After hearing the first two or three tones of a melody, listeners already assume that the key has been defined and that subsequent tones are attributed to the same chord and same tonal frame. A similar process occurs with harmonic sequences. This was first observed in a study by Sloboda and Parker (1985), where participants recalled a Russian folk song. The analysis of sung reproductions suggested that metrical and harmonic frames were largely intact although individual tones were often changed. Bharucha and Stoeckig (1987) explored these expectations by measuring reaction time in a priming paradigm, whereas another study involved the detection of "wrong notes" (Janata, Birk, Tillmann & Bharucha, 2002). Both studies produced results that followed the underlying tonal hierarchy and underlined the importance of harmonic frame for processing melodies. In a series of studies, Jansen and Povel (1999, 2000; Povel & Jansen, 2001; also Holleran, Jones & Butler, 1995) have teased out how the implied chord structure improves the melodic quality of simple melodic sequences. As tonality creates expectations of individual tones, it also creates expectations of implied chords, which also generate expectations of the other chord tones. For instance, a sequence of G_4 E_4 C_4 B_3 D_4 creates an expectation of G_4, completing the V chord in C major.

Melodic anchoring. The interplay of implied chords, tonal stability and temporal proximity of successive pitches forms asymmetries between the unstable and stable pitches. Unstable pitches tend to gravitate towards the proximate stable pitches. Consider the sequence C_4 E_4 F_4 $F\sharp_4$ which "yearns" towards the more stable, proximate G_4. Bharucha (1984, 1996) has outlined this asymmetry as the *melodic anchoring* principle (also Lerdahl, 1996), which is associated with cognitive reference points in cognitive psychology (Rosch, 1975). Lerdahl suggests that anchoring, might be culturally universal (1996, p. 361), which is also indirectly suggested by Bharucha (1996, p. 399), but no empirical data exist on their culture-transcending qualities. As a range of diverse tonal systems exist in cultures around the world, expectations of particular scale tones, tonal stabilities or chord progressions are most likely acquired through exposure to a particular culture although the strong proximity component that the anchoring principle incorporates is probably less culturally dependent, as observed earlier in the case of auditory stream segregation.

Melodic archetypes. Another layer of schema-driven musical features consist of melodic structures, to which Meyer has devoted considerable attention (1956, 1973; Rosner & Meyer, 1982, 1986). He has identified various archetypes, or prototypes that in turn have been used by others (Gjerdingen, 1988, 1991; Jones, 1981). Examples of these melodic structures range from "linear", "triadic", "complementary", and "changing-note", to "gap-fill" archetypes. Although some of these, notably gap-fill archetype, have been studied using cognitive experiments (Schmuckler, 1989; Schellenberg, 1997), they may not be so easily perceived, since the empirical results of these studies as well as those by Rosner and Meyer (1982, 1986) provided little evidence for them (von Hippel, 2000). Also, Narmour proposes that there are certain schema-driven cognitive rules, which may become evident only after dedicated music analysis (Narmour, 2000). Although music theoretical analysis might bring forward constructs showing how the composers created the music (for example, the specific mathematical way a twelve-tone pattern was generated), the perceptual validity of these constructs is dubious and therefore their effect on expectations is not considered in this thesis.

In sum, schema-driven knowledge has been demonstrated to affect expectations through learned pitch hierarchy, implied chords and melodic structures.

3.2.2 Temporal expectations

The Western notions of metre, grouping and metrical hierarchy, articulated by Lerdahl and Jackendoff (1983), have received considerable empirical attention (e.g., Palmer & Krumhansl, 1990; Bigand, 1993; Deliège, 1987; Todd, 1994). Palmer and Krumhansl (1990) found that participants expected the major events to occur at the most important beats in the metric hierarchy, followed by the lesser beats, which were then followed by the half-beat divisions and so on. However, the original formulation of the metrical hierarchy limits the theory exclusively to Western tonal music composed between c. 1600 and c. 1900 since much jazz and popular music employs dissonant rhythmic structures (Yeston, 1976) that are not well accounted for by the hierarchic model.

As detailed earlier in case of data-driven, temporal expectations, the Western notions of metre, grouping and metrical hierarchy may share some basic processes with other cultures but are mainly learned, specific patterns of Western culture (Lerdahl & Jackendoff, 1983). These notions have received empirical attention (Palmer & Krumhansl, 1990; Bigand, 1993) although not often by means of cross-cultural studies. The few explorations of rhythm by means of cross-cultural comparison have had problems of reconciling the previously-mentioned Western notions of metre and periodicity with the empirical findings. For example, Stobart and Cross (1994, 2000) explored pulse perception in recorded examples of the Easter music of the Northern Potosí (Bolivia), and found that European listeners tended to perceive different metre and position of the downbeat from that identified by Bolivian participants. From the European listeners' perspective the Bolivian musicians clapped on the "off-beats", which were also indicated by the foot-falls of the Bolivian dancers. The explanations derived from the language of the local culture, and the motor patterns involved in playing *guitarilla* or *charango* in the style prevalent in the *campesino* culture of Northern Potosí. The authors hold serious reservations as to the universal applicability of traditional Western music theory (see also Hughes, 1991, p. 330). Parallel evidence is available in various notations of African music, where many transcribers have indicated their assumptions about metre to be in contrast with the clapping provided by the African informants (Blacking, 1967; Jones, 1959; summarized by Temperley, 2000, p. 84-86). An empirical investigation by Magill and Pressing (1997) studied the tapping of West African rhythms by a Ghanaian percussionist. They compared an asymmetric timeline-ground model, which represented the traditional African understanding, and a pulse model based on Western ideas of regular metre and found that the African model fitted well with the performances. It was claimed to provide new insights into to the asymmetric and additive processes involved in producing multilayered African rhythms.

Igaga and Versey (1977, 1978) administered a series of rhythmic performance tests to Ugandan and English young people. Their studies showed that the Ugandan children showed superiority over the English children in the synchronization of rhythm, and the repetition of rhythm. Recently, Toiviainen and Eerola (2003) investigated the pulse perception of African and European listeners setting the task of tapping to melodic material consisting of both African and European folk melodies. They found no significant differences between the tapping behaviour of the two groups for the European melodies. In certain African melodies, however, there were differences that depended on the degree of conformity of their metric structure to Western metric hierarchy. The European participants had more difficulties in synchronizing with melodies containing rhythmic motives similar to the clave rhythm than did the African participants.

Hence, it appears that the way temporal patterns are organised in different cultures varies and that temporal patterns are influenced by the evolution of specific aesthetics in a particular culture, such as linguistic prosody or movement patterns, although nonetheless certain constants relating to the perceptual system are bound to exist. In summary, there have been a limited number of studies exploring the cross-cultural issues of temporal structures in music and they have provided evidence for the importance of stylistic knowledge and cultural cues in the perception of temporal patterns, although some culture-transcending factors have been observed as well.

3.3 Examples of melodic expectations

After a brief review of various expectations processes and models, concrete examples of melodic expectations are given. In the studies presented here, four different perspectives on melodic expectations were employed. The majority of the theoretical background and the empirical results concern (1) *discrete expectations*. These are expectations of particular tones at a particular point in time, defined by the structure of the previous context. To distinguish these "snapshots" from uninterrupted expectations, the term (2) *continuous expectations* is used. Continuous expectations cover the entire dynamic expectancy structure of a whole melodic sequence and are based either on discrete events, or on continuous (time-domain) representations of musical signals. A large-scale perspective of melodic expectations is obtained by examining the degree of violation in expectations across the melody, and this is taken as an index of (3) *melodic complexity*. Finally, melodic expectations are affected by recurring melodic segments, phrases, and motifs. The resemblance of these segments to previous segments is measured by (4) *melodic similarity*, which can be modelled by self-similarity of melodic contour and statistical features of the segments. To clarify these perspectives and the expectation models which are central to this thesis, examples of pitch-related expectations are given.

3.3.1 Discrete expectations

Discrete, event-based expectations have a central role in empirical work related to expectancy due to the principal experimental methodologies used in exploring listeners' musical expectations (probe-tone, reaction time and production methodologies, described in Chapter 4 which follows). These methodologies have drawn empirical attention to isolated moments of expectancy, which are typically constructed in the following way. A musical context is presented, consisting of a short melodic segment which is interrupted at a chosen point. At this point of interruption, listeners have formed specific expectations about how the context sequence might continue. These expectations can be explored by adding a continuation tone immediately after the context and asking how well the continuation fits with the listeners' expectations about how the melody might have continued. This process is then repeated with different possible continuation tones. The whole process is known as the *probe-tone method*. An analogous method can be applied to expectation models as well, as shown in Figure 3. In the figure, the musical context consists of a segment of *Läksin minä kesäyönä* tune (*On a summer's night*, a traditional Finnish folk song), which starts from the beginning of the tune but is interrupted in the middle of the second phrase.

Four possible continuation tones to the segment are given and evaluated by several expectation models.

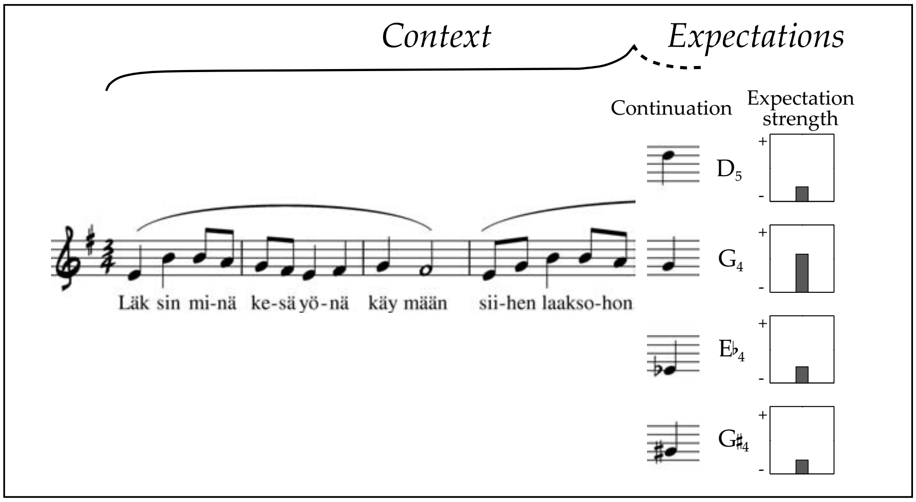


FIGURE 3 A segment of the *Läksin Minä Kesäyönä* song and the overall expectation strengths of four melodic continuations according to several data-driven and schema-driven expectation models.

In addition to displaying the musical context and the four proposed melodic continuations, Figure 3 displays the *expectation strengths* for the given melodic continuations. Expectation strength stands for a mean value taken from the various data-driven and schema-driven predictions for each of the four continuation tones. The tone that actually occurs next in the melody (G_4) receives the highest mean expectation strength score, whereas D_5 and the chromatic tones ($E\flat_4$ and $G\sharp_4$) receive considerably lower mean scores and are thus, according to the models, less expected continuations at this point of the tune. Expectation strength is intuitively clear even for a listener unfamiliar with the tune (i.e., one possessing no veridical knowledge of the song). The chromatic tones obviously do not fit the scale structure presented by the context and the D_5, which might otherwise be a suitable continuation, creates a sudden change in melodic direction and in the context involves a leap in pitch which makes it perhaps less than ideal. However, these intuitive comments can be better understood by examining the individual predictions of various expectation models with respect to the four possible continuations, as shown in Figure 4.

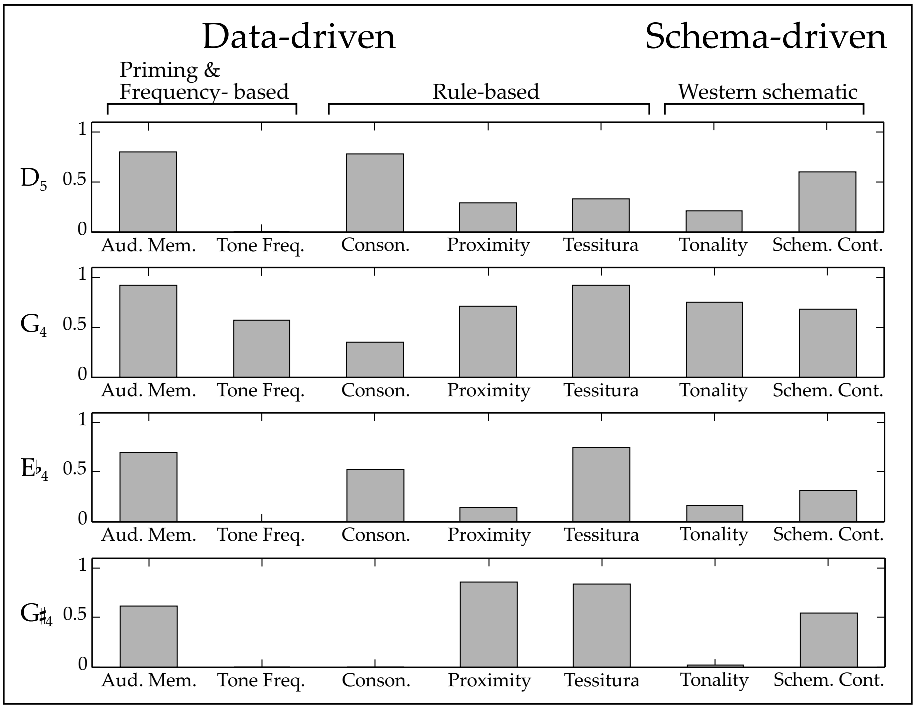


FIGURE 4 Individual components of expectation models for four melodic continuations to a segment of the song *Läksin Minä Kesäyönä*. From the left: *auditory memory model* (Leman, 2000), *tone frequency model*, two principles from Narmour's (1990) implication-realization model (*consonance* and *proximity*), *tessitura* (von Hippel, 2000), *tonality* (tonal stability values from Krumhansl & Kessler, 1982), and *Western schematic continuations* (Krumhansl, 1995b). Higher bars indicate higher expectation strength.

In Figure 4, the data-driven models are shown on the left and the schema-driven models on the right. The leftmost bar shows the prediction afforded by the short-term sensory context employing the *auditory memory model* (Leman, 2000), in which the auditory periphery and pitch completion process by the auditory system are simulated. The expectation strength is derived from a comparison of sensory and echoic memories at the time of the continuation tone. The second bar displays the expectation strength according to the tone frequencies that have occurred in the context. Both models predict highest values for G_4 as it has recently occurred in the context. The auditory memory model also shows some expectation for D_5, which shares many of its partials with the partials of the previous tones. Further data-driven predictions come from the rule-based heuristics. The next two principles are from Narmour's Implication-Realization model (*consonance* and *proximity*) and the third is inspired by Narmour's model (*tessitura*, von Hippel, 2000). According to the principle of *consonance*, D_5 and $E\flat_4$ are somewhat more expected than other tones as they form a consonant interval with the previous tone (perfect fourth). The principles of *proximity* and *tessitura* indicate high expectation strengths for $G\sharp_4$and G_4 as they are both close in pitch to the last tone of the context (A_4) and the same tones appear also in the expected melodic tessitura, since they are close to the median tone of the context ($F\sharp_4$). Finally, two schema-driven models indicate that the mediant tone (G_4) in E-minor is tonally more stable and thus more to be expected than the other tones residing in less tonally stable positions in the key hierarchy induced by the melodic context (*tonality*). According to the model that represents *Western schematic continuations*, G_4 and D_5 receive higher expectation strengths than the chromatic tones due to their correspondence with the ratings given to interval contexts by Western listeners (Krumhansl, 1995a).

In general, different principles yield different expectation strengths and it is an open empirical question how much each actually contributes the expectations formed by listeners. For example, empirical studies have found that a linear combination of certain – not all – principles is needed to account for the fitness ratings given by listeners. Also the exact weights of the principles may vary across musical styles and across musical backgrounds and the stylistic knowledge of the listeners. Note that the data-driven models work from a considerably shorter context (a few seconds or the previous two intervals) than the schema-driven models. For example, *tonality* infers the expectation strength of individual tones in relation to the key defined by the whole previous context. The applicability of the expectation principles and models may also vary across a single melody, depending on such factors as the salience of events influenced by metrical structure, phrase structure, and higher-level reductions, to name a few. Although all these aspects have not been thoroughly studied yet, one way to approach these questions is to look at expectations across a melody.

3.3.2 Continuous expectations

The previous example considered melodic expectation as a frozen snapshot of a certain position in a melody. However, expectations are far from static and shift constantly whilst music progresses. For this reason, it is useful to consider how the expectation models can be applied across time. There are two techniques for applying expectation models across time, depending on the model's reliance on continuous time-domain representation or on discrete note events (event-based representation). In the first case, the musical material (acoustic representation, tone onsets, etc.) is processed within a running window that represents the psychological present. In this case, the short-term memory or memories is simulated so that events that are further back in time from the present moment are less strongly represented. Thus the length of the running window is typically a few seconds and often contains exponential decay parameters. The *auditory memory model* (Leman, 2000) and *pitch-class distribution entropy model* (see Study VI) use this technique. In the second case, the model is applied separately to discrete, consecutive events (tones, intervals, or pairs of intervals). Narmour's Implication-realization model (1990), for example, uses this kind of technique, having a "window" of two intervals, where the degree of fitness of the realized interval to the implications set by the first interval is evaluated according to the model's principles. Figure 5 demonstrates an application of two continuous models (panels A, B), one discrete model (panel C) and in addition listeners' continuous ratings of predictability (panel B, dotted line). The example melody is the first eight bars of Charles Ives's song *Immortality*, which contains interest-ing and sudden departures from the conventions of Western tonal music. The particular melody and the models have been taken from Study VI.

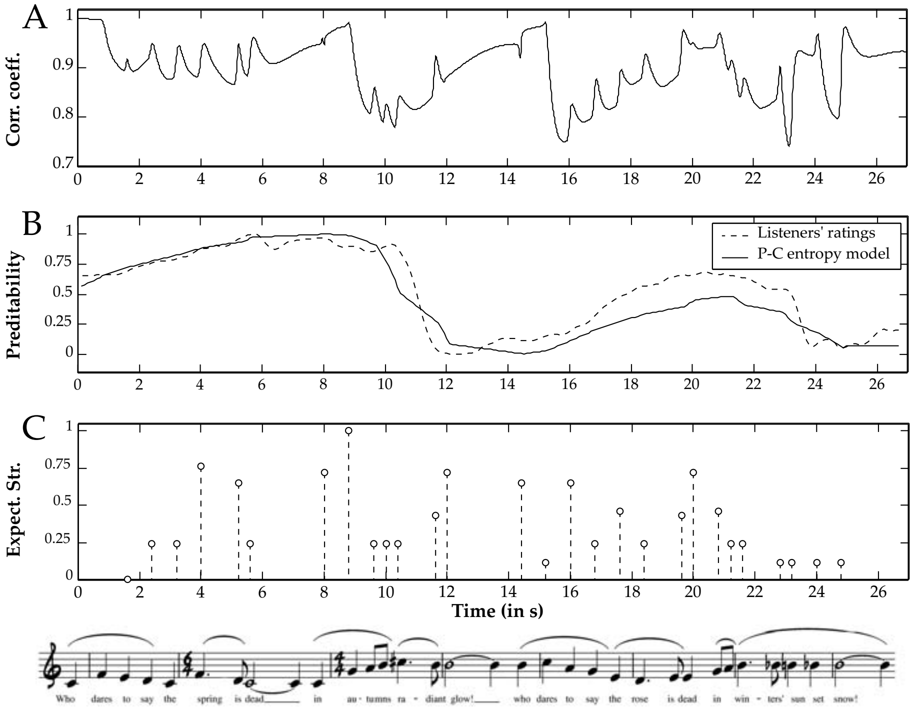


FIGURE 5 Continuous expectations for a segment of *Immortality* by Charles E. Ives (Song for voice & piano, S. 273, K. 6B67a). High value designates high expectation strength. Panel A represents the *auditory memory model* (Leman, 2000). Panel B shows the *pitch-class distribution entropy model* and listeners' continuous predictability ratings for the melody (both from Study VI). Panel C displays the mean of *Narmour's implication-realization model* (1990).

In Figure 5, panel A represents the prediction of the *auditory memory model*, which is a proper continuous model as it is based on audio signal (although technically, even an audio signal in the digital domain is based on discrete events, samples). The model depicts the commonality between acoustic sensory memory and longer echoic memory (half-decay times of 0.1 and 1.5 seconds, respectively) across the duration of the melody. This commonality measures how similar the current moment is acoustically to the previous moment and the model has been used to explain the fitness ratings obtained from listeners in probe-tone studies (Leman, 2000). Panel B represents the *pitch-class distribution entropy model*, which is an event-based continuous model. Entropy is a measure of disorder in a distribution. In this case, the tone distribution for each moment in time (within a 2-second window) is sampled and the entropy of the distribution is taken as a measure of predictability. This measure corresponds fairly well to listeners' continuous ratings of predictability, which are also shown in panel B (the dotted line denotes the mean value of participants' predictability ratings in Study VI). At the 10-second point in time, listeners' ratings of melodic predictability sharply decrease as the melody suddenly goes out of the tonal frame (appearance of C♯5 and ending the phrase to B4). However, the melody re-establishes the more conventional melodic progression during the next three bars, as indicated also by the increased values of the model. Consequently, the listeners' ratings also show an increase in predictability. In the last two bars, unexpected melodic twists occur again (alternation of B4 and B♭4), indicated again by the sudden drop in the predictability ratings by the listeners. Both the auditory memory model and the pitch-class distribution entropy model show sharp dips in their predictions of melodic expectation at these unexpected moments in the melody. Panel C shows the mean using the principles of *Narmour's implication-realization model*, which is based on discrete events. At the onset of each tone, the model's principles are evaluated on the basis of the previous two tones (implicative interval and realized interval). For example, tone repetitions show as peaks in the graph as they receive high expectation values by the principles of *proximity* and *consonance*. However, this simple demonstration does not take into account the fact that the principles in Narmour's model should be weighted in a certain way to obtain a plausible expectation structure. It is also the case that all the other models shown in Figure 4 could be implemented as continuous models using the techniques described above. At this point, however, it is sufficient to note that continuous models highlight the dynamic, time-dependent nature of melodic expectations.

3.3.3 Expectations and complexity

Measuring the overall fluctuations in melodic expectations also has the advantage of capturing the global "expectedness" or predictability of a melody. This index, called *melodic complexity* has been a focus of a number of studies relating to musical preferences (Berlyne, 1971; Smith & Cuddy, 1986; North & Hargreaves, 1995) and culture-transcending components of music that convey emotions in music (Balkwill & Thompson, 1999). In this thesis melodic complexity is taken to represent the extent to which the melody violates listeners' expectations. In other words, if listeners' expectations on the principle of pitch proximity are violated the melody will be judged to be more complex (e.g., Cuddy & Lunney, 1995). Similarly, if the melody is tonally ambiguous it is experienced as melodically incoherent and hence complex. Similarly, deviant temporal patterns, such as syncopation and uneven beat division are more difficult to produce (Clarke, 1985; Povel, 1981; Boltz & Jones, 1986), and tend to be perceived as more complex (Povel & Essens, 1985; Essens, 1995). The continuous perspective to expectation makes it possible to assess the degree of violation in expectation by taking the inverse of the mean of the expectation strengths across the whole melody. For example, the opening segment of the song *Immortality* by Charles Ives, shown in Figure 5, has presumably higher melodic complexity than the opening segment of the tune of *Läksin minä kesäyönä*, as the former melody violates listeners' expectations more than the conventional folk melody. In this case, the differences lie mostly in the tonal structures of the segments but of course this is not the only way expectations can be violated. Insights into the components of complexity may be discovered by comparing the listeners' evaluations of melodic complexity and a set of variables representing deviations from the expected pitch, rhythm and temporal structures (Study IV).

3.3.4 Expectations and similarity

The expectation models used in the previous examples do not explicitly take into account repetition of melodic phrases and motifs that give cues to the listener as to what to expect in future. Psychological studies of similarity relations in music have demonstrated that repeated or similar melodic patterns influence music perception considerably (Dowling & Bartlett, 1981; Cuddy, Cohen & Miller, 1979; Pollard-Gott, 1983; Smith & Cuddy, 1989). In *Läksin minä kesäyönä*, the second phrase resembles the first and therefore listeners' expectations about how the melody continues in the second phrase are probably influenced by the recognition of the similarity between the phrases. This is largely what Krumhansl (1997) found when applying melodic expectations to higher hierarchical levels and to phrase repetition.

This issue can be examined using two methods, s*elf-similarity of melodic contour* and *similarity of statistical features of melodies*. The first operates with the shape of the melodic contour, which can be defined by simple encoding using the so-called *Parson code* (1975), where each pair of consecutive tones is coded as "U" for upward intervals, "D" for downward intervals, and "R" for repeated tones. In this thesis, a representation of contour that retains the sizes of the intervals and the durations of the tones, is used instead of the simple encoding. This richer representation is shown in panel A of Figure 6. Melodic contour is a useful representation of a melody and found to be easier to remember than exact interval information (Dowling, 1978; Dowling & Fujitani, 1971; Edworthy, 1985; Dowling & Bartlett, 1981). Also numerous music information retrieval systems use melodic contour to find specific melodies from large music databases (e.g., Kim, Chai, Garcia & Vercoe, 2000; Lemström, Wiggins & Meredith, 2001). To find out whether any example of melodic contour contains repeated patterns, an autocorrelation technique can be used (details in study IV). A time series is autocorrelated if it is possible to predict its value at a given point in time from knowledge of its value at other points in time. The autocorrelation function of a time series is obtained by correlating the series with a delayed copy of itself, using different time lag values. If the melodic contour can be predicted from the previous points in time, it can be called *self-similar*. Panel B in Figure 6 shows the self-similarity in the melodic contour (panel A) of *Läksin minä kesäyönä* based on the autocorrelation function.

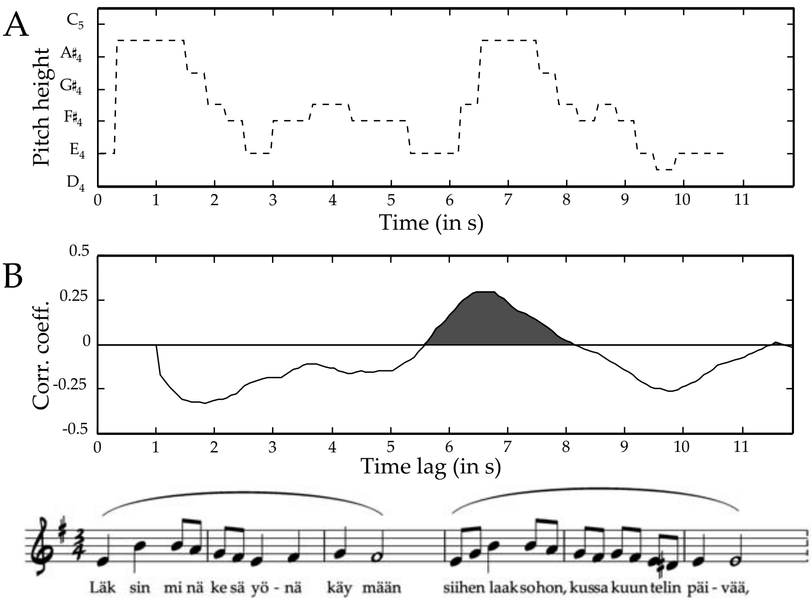


FIGURE 6 The melodic contour (panel A) of the first two phrases of *Läksin Minä Kesäyönä* tune and the autocorrelation across the melodic contour (panel B). The shaded area shows the location and the magnitude of the self-similarity in the melodic contour. In the lower panel, the values for the time lags between 0 and 1 second have been omitted as the autocorrelation always correlates perfectly with itself (time lag 0).

The shaded area in Figure 6 shows the self-similarity of the melodic contour, showing high values at a time lag of 6 to 7 seconds. These time lags correspond with the start of the second phrase, which indicates that the second phrase is similar to the first phrase. Repeated melodic segments such as these, will most likely generate expectations for the imitation of the previous segment. Also, highly self-similar melodic contours that are easier to predict are rated as melodically simpler than those which are not (see Study IV).

The second measure of melodic similarity is based on the comparison of the statistical features extracted from the melodies. The idea is that if two melodic segments have similar distributions in the extracted features, the two will probably be perceived as similar. The choice of the features is therefore crucial. Ideally, the features would correspond with those that a listener might use in evaluating similarity. Often zero-order statistics, for example, the frequency of tone onsets, are used, as it has been shown that listeners are sensitive to tone frequencies (Oram & Cuddy, 1995). Other zero-order statistics include interval and duration frequencies. However, higher-order statistics (tone-, interval, and duration transitions, three-tone patterns, etc.), are usually necessary to avoid confusions between melodies that have identical zero-order distributions but different temporal order within them. In other words the tune of *Läksin minä kesäyönä* and a scrambled version of the same melody, where the notes have been randomly reordered, would exhibit identical tone and duration distributions (zero-order statistics). Similarity between the profiles can be measured by calculating the distance between the profiles, using, for example, Euclidean or city block distance metrics, which are dissimilarity measures. In Figure 7, an example of similarity in the statistical features of melodies is given. The phrases of *Läksin minä kesäyönä* are compared with each other by calculating the similarity between six statistical features (zero- and first-order tone-, interval- and duration distributions).

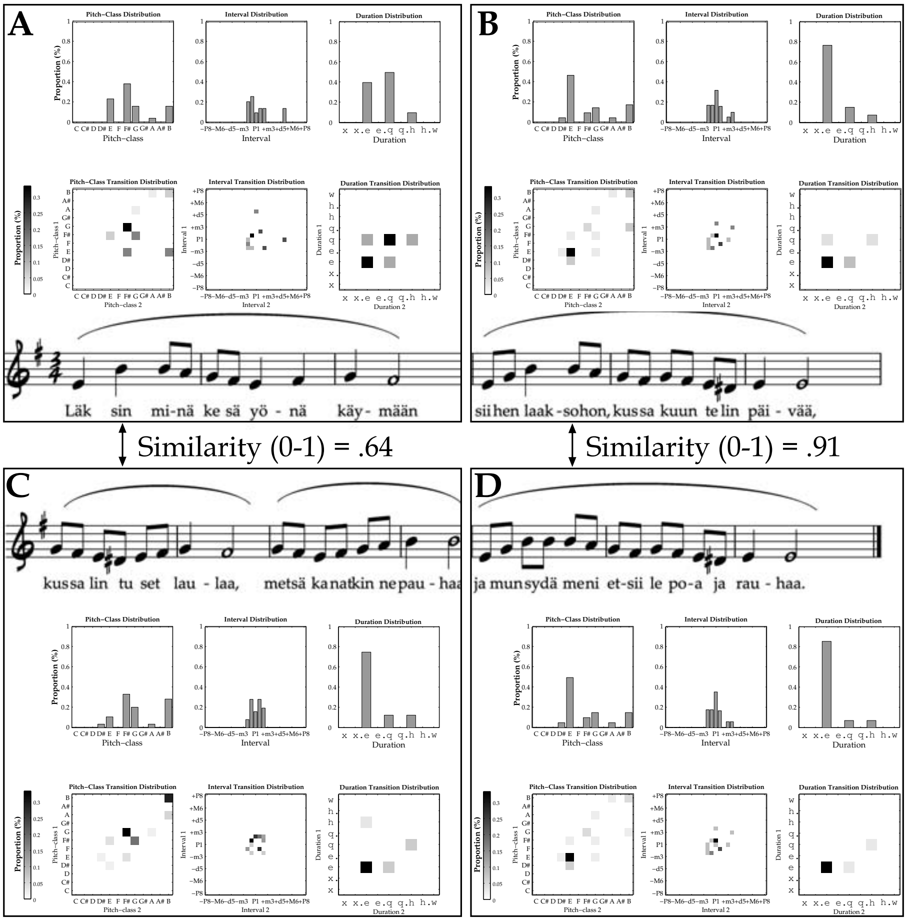


FIGURE 7 The similarity between phrases A and B (.64, modest similarity) and A' and A'' (.91, high similarity) of *Läksin minä kesäyönä* using zero- and first-order tone-, interval-, and duration distributions. The similarity measure is the mean city block distance of the distributions, scaled to between 0 (no similarity) and 1 (perfect similarity).

Figure 7 displays six statistical distributions related to each phrase of *Läksin minä kesäyönä*. By looking at the notation of the first (A) and the third phrase (B), it is apparent that the phrases are somewhat different even though largely the same tones are used in both. Despite this similarity in the pitch-class distributions, the correspondence between all the six distributions is rather modest (0.64). On the other hand the similarity between the second phrase (A') and the fourth phrase (A'') is high (0.91) as the two phrases only differ marginally. The similarities between numerous melodic segments or complete melodies, consisting of multiple statistical features, can be visualized using such techniques as multidimensional scaling (Kruskal & Wish, 1978) or self-organizing maps (Kohonen, 1997). Using these projection methods, melodies with similar features will be projected in proximate areas (see Study V and Toiviainen & Eerola, 2001, 2002).

For listeners' melodic expectations, detection of melodic similarity has three implications. Firstly, listeners are sensitive to the frequency of occurrence of melodic events and hence the reappearance of similar melodic segments will increase expectation for these events. Secondly, the recognition of longer patterns, such as reoccurrence of parallel phrases in a melody, will form expectations for further repetitions of these patterns. To put it another way, listeners can have *veridical expectations* of the melody based on repetition of melodic segments even though they are not previously familiar with the melody. However, the detection and recognition of melodic segments is a complicated process requiring segmentation of a melody and will not be covered here. Thirdly, similarity is also the basis for using appropriate schematic expectations of familiar musical style. The recognition and activation of schematic knowledge requires the detection of similarity relations between the music that is heard and long-term, schematic representations.

In sum, melodic similarity, based either on self-similarity of melodic contour or on statistical features of the sequences, is an important component of melodic expectations.

4 EMPIRICAL METHODOLOGIES

Research in music cognition assumes that understanding of the perceptual and cognitive processes in music is obtained from theoretical models corroborated by empirical studies, where pertinent data are collected from listeners using a wide variety of experimental methods. The exact empirical methodologies vary considerably within music cognition (including cross-cultural approaches). The following chapters introduce the assortment of empirical techniques and methodologies that have been used in determining expectations in music.

4.1 Music analytic methods

At least since Rameau (1722/1971) suggested that certain harmonic movements imply certain harmonic closures, expectations have featured prominently in music analytical treatises in the Western tradition. For example, Schenker's famous theory (1906/1954, 1935/1979) involves expectations of abstract tonal movements where underlying tonal movement is directed towards certain goals. Also Piston's typology of chord progression from very common to uncommon (1941/1978) suggested which chord orders are more predictable, and this hypothesis has led to empirical assessments (Povel & Jansen, 2001, 2002). Work by Meyer (1956, 1973) provided numerous examples of implications arising from melodic, rhythmic and harmonic structures occurring in Western art music between c. 1600 and c. 1900. Although empirical evidence has been found for certain melodic archetypes (Rosner & Meyer, 1982, 1986), these findings have recently been critically re-evaluated and simpler explanations have been proposed (von Hippel & Huron, 2000). Nevertheless, Meyer has had a strong influence on other music theorists, who have provided examples and theoretical details of how expectations work in Western music. For example, Gjerdingen (1988) analyzed changing-note melodic pattern (1-7 ... 4-3) and Lerdahl and Jackendoff's theory (1983) also portrays the formation of hierarchical, temporal expectations.

Another way of analysing expectations in music arises from the notions of information theory, which had their heyday in the 1950s. Statistical analyses of melodies have provided information about the statistical properties of music. These studies usually involve counting pitch-classes and intervals (Watt, 1924; Densmore, 1972/1932; Cohen, 1962; Pinkerton, 1956) although more sophisticated sets of properties have been tallied as well. For example, Marillier (1983) calculated the tonal progressions in Haydn's symphonies, while other frequencies in music were studied by Hofstetter (1979), Morehen (1981), Pont (1990), Nettheim (1993), Simonton (1980), Knopoff & Hutchinson (1981, 1983), and Snyder (1990). It can be assumed that occurrences of tone events will influence listeners' expectations, either through their being exposed to music which conforms to certain regularities or as a result of the composers working within human cognitive limitations and predispositions. For example, Vos and Troost (1989; also Watt, 1924) demonstrated that in folk songs and classical music, large intervals more commonly ascend and small ones descend. Huron (1996) sought evidence for the predominance of the arch shape in melodic phrases of Western folk songs. Perhaps the most direct link to musical expectations was proposed by Coons and Kraehenbuehl in 1958 when they outlined a dynamic procedure that adjusted the probabilities of the musical system as the music progressed.

Further indirect evidence for musical expectations has been accumulated in more recent studies of melodic structures that have tested the expectation model of Narmour (1990) using an analysis of Bohemian folk melodies (Thompson & Stainton, 1996, 1998). Moreover von Hippel (2000) proposed enhancements to Narmour's model by redefining pitch proximity in terms of two melodic constraints called *tessitura* and *mobility* based on an analysis of folk melodies around the globe. All these music analytical methods raise a question about whether the musical material influences our learning of expectations, or whether conversely certain common human predispositions for processing auditory signals produce the patterns revealed by music analytical studies.

4.2 Production methodology

A direct measure of melodic expectation is obtained by asking listeners to produce continuations for melodic fragments. In production studies the participants either sing what they feel is the most likely continuation (Hershman, 1995; Cohen, 1991; Carlsen, Divenyi & Taylor, 1970; Carlsen, 1981; Unyk & Carlsen, 1987; Larson, 1997), play the continuation on a keyboard or other apparatus (Auhagen, 1994) or even compose the continuations (Schmuckler, 1988). The participants can also indicate their expectations by relating a continuation tone to other tones (Brown & Butler, 1981). While this method is musically relevant and ecologically valid in many cases, the drawback lies in its dependence on production skills. For example, participants may not be able to sing all the tones they expect and predict. Moreover the apparatus, for example, the keyboard may direct the production of melodic continuations in specific ways. The methodology also requires conscious attention to expectations, which may in fact be largely unconscious. Despite these shortcomings, producing expected continuations is a natural task to perform in music and should be further developed because of its high ecological validity.

4.3 Probe-tone methodology

Sequential and chord-related expectations have been studied using the *probe-tone method* (Krumhansl, 1990; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979). This technique consists of evaluation of probe tones with respect to their fit for a given context and hence the task is *retrospective*. The context is usually a chord sequence establishing the key, although various melodic contexts have been used as well. After the presentation of the context a probe-tone follows. Each of the probe tones is rated by listeners in terms of how well it completes or fits with the context. After a number of probe-tones – from twelve to twenty-four – have been evaluated by the listeners, a profile describing the fitness of the probe-tones to the context emerges. If the key-defining context uses all 12 pitch-classes, these results can be taken to represent tonal stability of the pitch-classes (Krumhansl, 1990). There have been numerous applications of this method, ranging from the original melodic and harmonic contexts to rhythmic contexts (Palmer & Krumhansl, 1990). The task has been performed by young and old listeners, musicians and non-musicians (for a summary, see Cuddy & Badertscher, 1987).

Criticism of tonal hierarchy theory and the probe-tone method has been raised by Butler and Brown (Butler, 1989, 1990; Butler & Brown, 1984), whose moot point was that the tonal hierarchies do not account for the temporal order of tones in a musical context. Although the control of temporal order does not seem to be essential for the emergence of the hierarchies (Smith & Schmuckler, 2000), this points to a drawback of the probe-tone methodology. A recent variant of the probe-tone method, termed continuous probe-tone method (Krumhansl & Toiviainen, 2000) amends the method by incorporating sequential aspects to it as listeners concurrently rate the fit of the probe to the sequence. The sequential order of tone events and the limitations of short-term memory have been incorporated in recent models of tonality that can predict concurrent probe-tone ratings of sequences (Toiviainen & Krumhansl, 2003).

Another criticism of probe-tone methodology concerns the role of short-term effects related to the stimuli instead of reflecting the long-term memory abstractions (Huron & Parncutt, 1993). Simulation of the early probe-tone experiments of Krumhansl & Kessler (1982) by Leman (2000) showed that echoic images of periodicity pitch patterns might account for the results obtained in the experiments.

4.4 Reaction time and detection methodology

Expectations aid the formation of responses to events. Strong and definite expectations yield faster responses. Thus reaction time offers another way of measuring expectations. Bharucha and Stoeckig (1986) used the following technique to obtain rapid identification of a target chord (minor/major or mistuned) that had been preceded by a *prime chord*. A prime chord is a previously heard chord which is made more accessible – *primed* – in the listeners mind. Those chord combinations that were primed were identified more quickly than unexpected target chords. The same technique has been used by Schmuckler and Boltz (1994) to determine harmonic and rhythmic influences on expectations.

A related indirect way of studying expectations comes from the study of the errors listeners make when detecting violations in melodic rules (Jones, Boltz & Kidd, 1982). However, criticism of both techniques calls attention to the simplicity of the stimuli, which need to be highly constrained in order to vary one particular musical feature at a time. The advantage of the reaction time methodology is that the reactions obtained are little affected by conscious reflections and therefore implicit, automatic expectations are revealed.

4.5 Continuous measurement methodology

In the 1930s, Kate Hevner noticed in her well-known studies of musical emotions that a single label denoting the emotional character of music is not sufficient to capture the changing nature of musical emotions. She developed a method (1936) where listeners indicated their emotional labels for premarked sections of music and by doing this paved the way for continuous measurements. There are other, more recent studies using semi-continuous or incremental measurements of expectations. Krumhansl and Kessler (1982) presented probe-tone tasks in a sequential manner, where the first chord was presented, followed by the probe, and then another chord was added and so on until all the chords in the sequence were incrementally evaluated. Similar, though somewhat more advanced, incremental stimulus presentation has been used to study key mode inference (Vos & Verkaart, 1999), and melodic expectations (Aarden, 2002; Povel & Jansen, 2001, 2002). In these studies expectations are probed at various locations within the melody by means of rapid reaction time test.

Actual continuous measurements of expectations are in short supply in the experimental literature. Nielsen (1983) collected ratings of musical tension by means of a pair of tongs and the results were registered on a polygraph. Subsequently, segmentation, musical ideas and tension ratings have been studied using continuous behavioural measurements (Krumhansl, 1996, 1998; Madsen, 1997) where listeners indicate their rating of something (usually emotion or tension) using a mouse connected to a computer. While continuous rating methodology is able to come to terms with the dynamic nature of musical processing, the difficulty lies in the analysis of time-series data (Schubert, 1999, 2001a, b). However, new analytical techniques from the field of signal processing are opening up new possibilities for dealing with continuous measurements.

4.6 Methods of cognitive neuroscience

Several neuropsychological indicators of musical expectancy have been established in cognitive neuroscience. The most widely used way of capturing the temporal and spatial properties of brain processes is to measure the electrical fields generated by brain activity. The techniques for measuring these fields – *electroencephalography* (EEG) and *magnetoencephalography* (MEG) – have long been known. To investigate time-locked events such as deviations from an expected continuation of sequence, discrete, short stimuli are presented numerous times while brain responses are recorded. The mean of the responses (waveforms) is usually taken, and this is known as *event-related potential* (ERP) or *event-related magnetic field* (ERF) waveform, which can be analysed in terms of amplitude, peak location, width and latency (review of methods in Tervaniemi & Huotilainen, in press).

In EEG studies, the ERPs are one of the most common indicators of deviations in sequential stimuli, including music. Several studies have examined the ERPs elicited by deviations inserted into well-known melodies (Besson & Macar, 1987; Paller, McCarthy & Wood, 1992; Verleger, 1990). Even more studies have focused on the ERPs evoked by the deviations in unfamiliar musical contexts (Besson & Faïta, 1995; Besson, Faïta, Peretz & Bonnel, 1998; Hantz, Kreilick, Kananen & Swartz, 1997; Janata, 1995; Patel, Gibson, Ratner, Besson & Holcomb, 1998; Granot & Donchin, 2002; Schmidt, Gunter & Kotz, 2002). The specific component in the ERP which has been found in these studies to be the best indicator of violation of musical expectations, both in familiar and unfamiliar melodies, is the late positive component (usually P300, i.e., a peak in amplitude occurs 300 ms after the event). Deviations in timbre, pitch direction or contour in sequential stimuli may also be apparent in mismatch negativity (MMN) component related to memory- and change-detection processes (Tervaniemi, Ilvonen, Karma, Alho & Näätänen, 1997; Tervaniemi, Winkler & Näätänen, 1997; Tervaniemi, 2001). In harmonic violations the component is often P600 or another music-specific component (RATN, ERAN etc., see Patel, Gibson, Ratner, Besson & Holcomb, 1998; Koelsch, Gunter, Friederici & Schröger, 2000; Maess, Koelsch, Gunter & Friederici, 2001; Janata, 1995). The amplitude of the component is positively related to the degree of expectation violation. Some effects of musical training have been observed, in particular that musically-trained subjects have larger and better defined amplitude peaks than musically naïve subjects (Besson & Faïta, 1995; Tervaniemi, 2001; Besson, Faïta & Requin, 1994; Koelsch et al., 1999; Regnault et al., 2001).

While EEG and MEG recordings can be used to locate the activated areas in the brain that have caused the recorded signals by using various source techniques (dipole and distributed source models), hemodynamic methods tap directly into anatomical correlates of the brain processes. These methods, namely *positron emission tomography* (PET) and *functional magnetic resonance imaging* (fMRI) detect the metabolic changes associated with neural activity and thus identify the areas responsible for processing various tasks. Patel and Balaban (2000) used fMRI to explore the brain processes involved in processing melodic sequences that varied in their predictability. They introduced amplitude modulation to pure tones, which resulted in detectable stimulus-related cortical activity in a brain scan. Hemodynamic methods also allow to study purely imagined musical imagery, generated from the long-term memory as contrasted to what normally occurs in music listening in which the memory-dependent processes interact with representations of incoming auditory input. PET scans of imagined music listening have revealed activation in the very same structures processing the incoming audio signals (Halpern & Zatorre, 1999; Zatorre, Halpern, Pery, Meyer & Evans, 1996; Janata, 2001), suggesting that mental imagery and musical expectations are processed in the same way as real heard excerpts would be, i.e., they are heard in "the mind's ear".

The methods of cognitive neuroscience have the desirable qualities of portraying listeners' expectations in their raw, unprocessed form. However, this is also a limitation, as the mechanisms and the explanations for the observed results are still difficult to interpret. As this highly technical field progresses from its infancy to adulthood, it is assumed that greater convergence will occur between behavioural results, music-theoretical predictions and brain responses. It should also be noted that various behavioural measurements that have been described previously are often used in conjunction with brain recordings (e.g., Tervaniemi, Ilvonen, Sinkkonen, Kujala, Alho, Huotilainen & Näätänen, 2000; Schoen & Besson, 2002). For example, priming effects, detailed earlier in reaction time methodology, have recently been further refined by fMRI (Tillmann, Janata & Bharucha, 2003).

5 AIM OF THE STUDIES

This work aims to investigate musical expectancy, a field of study which provides ways of understanding the processes and knowledge listeners use to structure, interpret, remember and perform music. These consist of data-driven and schema-driven processes, which are investigated in a series of empirical studies concerning melodic expectations (I-III), complexity (IV), similarity (V), and predictability (VI). All the models used in the empirical studies in the thesis are available as Matlab toolbox (Eerola & Toiviainen, 2003).

Substantial similarities in responses to simple auditory stimuli have been observed among listeners from different cultures. Agreement between people from various different cultures, however, starts to diminish once the focus of the studies turns towards culturally learned properties of music. Melodic expectancy may be positioned somewhere between universal and culturally determined processes, since the literature to date appears to demonstrate both similarities and differences between listeners from different cultures. Some of the similarities have been captured in a cognitively oriented music-theoretical model of melodic expectations (Narmour, 1990). Another data-driven explanation of melodic expectations relies on the listener's sensitivity to statistical properties in the melodies, a factor which has also been shown to influence the formation of melodic expectations (Oram & Cuddy, 1995). Both of these models are presumed to be universal, and the claim of universality has been investigated in cross-cultural contexts, focusing on tonal hierarchies (Castellano et al., 1984; Kessler et al., 1984) and interval contexts (Carlsen, 1981; Unyk & Carlsen, 1987) but the specific expectation models and their data-driven and schema-driven qualities have been little studied (Krumhansl, 1995b).

Studies I-III, the intention was to tease apart the role of schema-driven and data-driven processes in the formation of melodic expectations by using probe-tone methodology and several groups of listeners, varying in their musical backgrounds. Study I, a rare variety of Finnish spiritual folk hymns were evaluated by expert and non-expert Western listeners. However, comparison of these two groups did not permit exploration of the role of different degrees of stylistic knowledge in expectancy formation. The aim of study II was to extend the scope of the previous study by using another highly idiosyncratic style of music (North Sami yoiks), the same methodology, and three groups in the comparison (expert, semi-expert and non-expert listeners). Previous cross-cultural studies of expectations, as also Studies I and II, have employed groups of Western listeners of different nationalities (German, Finnish, North-American, Hungarian, etc.) who have been significantly exposed to Western music. Because of this, the possible culture-transcending qualities of expectation models have not been seriously put to the test. This is amended in Study III, where South African traditional healers, who have been minimally exposed to Western music, evaluate continuations of the excerpts from North Sami yoiks, which are in a musical style completely unfamiliar to them. The goal of Study IV was to explore both the effects of stylistic familiarity and culture-transcendent factors in the perception of melodic complexity by comparing African and Western listeners' ratings of African and Western folk melodies.

The aim of the second series of studies (V-VI) was to examine supplementary questions relating to melodic expectations. The first of these is melodic similarity (V), which has an important influence on melodic expectation, as recognition of familiar patterns requires perception of similarity relations in music. The aim of Study V was to model melodic similarity by using statistical features of melodies and to evaluate these similarity measures using behavioural data from a similarity rating experiment consisting of European folk melodies.

Expectancy is essentially a temporal, dynamic process, in which pitch and temporal pattern promote the formation of hypotheses about ensuing events. This dynamic nature of expectations was approached in Study VI, where the aim was to develop the experimental methodology by collecting continuous measurements of melodic expectations. An attempt was made to assess the continuous ratings by dynamic models, which were based on a combination of data-driven and schema-driven probabilistic information.

6 MATERIALS, METHODS, AND RESULTS

Details of the materials, methods and results are set out at length in research papers I-VI. The following summary deals only with the central findings of the research.

6.1 Cross-cultural approach to melodic expectations (I-IV)

Studies I and II, probe-tone experiments were conducted using Finnish spiritual Folk Hymns (I) and North Sami Yoiks (II). In Study I, behavioural responses were obtained from listeners with generally similar cultural and musical backgrounds. Two groups of Finnish listeners participated: those who were familiar with the musical style represented by the excerpts (experts) and those who were not (non-experts). In Study II, three groups of listeners participated in the experiment. These were Sami yoikers (experts), Finnish music students (semi-experts) and Central European music students (non-experts).

The ratings obtained from the expert and non-expert listeners in both studies supported the notion that expectations are governed by common, data-driven psychological principles although schema-driven, stylistic knowledge has a distinct effect on these principles. Although the results obtained on the principles of Narmour's implication-realization model (1990) were different from those of previous studies (Krumhansl, 1995b; Schellenberg, 1996, 1997), the principles predicted expert and non-expert ratings to a comparable degree. Also, there were differences between the experiments in the contribution made by the principles of Narmour model. For the hymns (Study I) the principle of *proximity* was the most important component in the model, while for the yoiks (Study II) *consonance* played a considerable role in the ratings. In both cases the differences in the emphasis of the Narmour model's principles reflected the characteristic features of the musical material. However, a direct comparison of the weights derived from the implication-realization model between the experiments or with previous studies is not feasible as the musical contexts in each experiment are entirely different.

In addition the importance of statistical information in expectancy formation was demonstrated. This was achieved by performing statistical analysis of the relevant, larger corpora (spiritual folk hymns and yoiks). As predicted, the experts' ratings were especially influenced by the regularities of the style in question. Zero-, first- and second-order tone distributions were also considered and the results showed that reliance on these features in the formation of expectations was dependent on the level of stylistic knowledge: the experts used more detailed information, higher-order statistics, whereas the non-experts relied generally on zero-order statistics. Nevertheless, all listeners were not only sensitive to tone distributions, as demonstrated in previous research (Oram & Cuddy, 1995) but also to higher-order statistics. With regard to the hymns, the experts were less critical than the non-experts in distinguishing the tones that define the minor or major key, implying further schema-driven differences, as the distinction between major and minor modes is less clear in spiritual folk hymns. In yoiks, the Finnish listeners demonstrated highly similar responses to those of Sami listeners in the yoiks they were *familiar* with, suggesting that familiarity can compensate for lack of stylistic knowledge. The effect of stylistic knowledge was nevertheless apparent, as schematic knowledge of Western music had greater influence on the non-expert listeners than on the other two groups. In both experiments the experts' ratings reflected veridical expectations, as to the correct continuation tones in the hymns and in the yoiks. All in all, to a considerable degree listeners' expectations could be explained by the models: prediction rates using all variables ranged from 70% to 90% for non-experts and experts in Study I, and 70%, to 80% in Study II.

Study III extended the scope of the findings obtained in Studies I and II to non-Western listeners (traditional healers from South Africa) and to the principles of auditory processing. In previous cross-cultural studies, as also in Studies I and II, the participants have been significantly exposed to Western music and hence research into the universality of Narmour's principles has not been very firmly based. In Study III, South African traditional healers, who are minimally exposed to Western music, evaluated continuations to North Sami yoiks (a subset of the yoiks used in Study II). In addition the role of Western schematic knowledge, captured in the model of Krumhansl (1995b), was critically evaluated. The results of the experiment indicated that data-driven models are robust in explaining expectations, regardless of the cultural background of the listeners. It was also found that event-frequency models had more explanatory power than the auditory memory model. Again, support was found for the same principles in Narmour's implication-realization model (1990) that predicted the ratings in Studies I and II.

Although probe-tone methodology posed challenges in the choice, presentation, and rating of the musical material, the one overarching outcome was that all groups in Studies I to III produced responses largely similar to each, suggesting that listeners adapt easily to different ways of structuring music of an unfamiliar style. This finding is in line with previous cross-cultural studies (Castellano et al., 1984; Krumhansl, 1995b). In sum, all listeners seem to use data-driven heuristics consisting of dynamic strategies (frequency of events and the acoustic information provided by the contexts) and innate or at least culture-transcending heuristics (the principles of the implication-realization model).

In Study IV, a more generic approach was taken to exploring melodic expectations. The notion of melodic complexity, which represents the degree of expectation violation experienced by listeners, was investigated theoretically and empirically. The role of schema-driven and data-driven processes in expectation and complexity formation was explored by developing a range of melodic complexity measures and carrying out two experiments using Western and African folk songs. The complexity measures included entropy-based measure of event distributions in music and deviations from Western temporal structures as well as measures derived from the self-similarity of melodic contour.

The results obtained from the two groups of listeners (Western and African) illustrated no large overall differences although some rhythmic aspects of the African melodies brought out culturally-determined differences between the groups. In the complexity ratings of the Western folk songs the differences were small but some evidence of the type of stylistic knowledge the groups possessed was evident. The African folk songs, which were stylistically familiar only to the African group, were less well predicted by the variables, suggesting that the Western perspectives built into measures rendered them incapable of accounting for the full range of melodic complexity in African music. The idea that ratings of melodic complexity transcend cultural boundaries, as proposed by Balkwill and Thompson (1999), was supported with reservations, as stylistic knowledge was also observed to have an effect upon the ratings.

Although these three empirical studies are amongst the few cross-cultural explorations into musical processing, it is acknowledged that the questions, empirical paradigms and the models in each study represent largely Western musical concepts. Despite this, the reactions and comments collected from all participant groups, coming from diverse musical cultures and separated by a distance of 10 000 km, suggested that melodic expectation was a topic that they could all immediately respond meaningfully to.

6.2 Statistical approach to melodic expectations (V-VI)

The second series of studies (V-VI) examined supplementary questions relating to melodic expectations. In Study V, the question of melodic similarity was approached by modelling similarity using the statistical features of melodies (zero- and first-order statistics). Reliance on statistical features was motivated by findings in language processing (Saffran et al., 1999) and cross-cultural studies on music (Castellano et al., 1984; Kessler et al., 1984), both of which have shown the importance of frequency information in cognitive processing. An experiment was conducted in which listeners rated the pairwise similarities between 15 folk melodies representing five folk music styles (Finnish, Greek, Irish, German, and Sami). Moderate success (40% of the variance) was achieved in explaining listeners' similarity ratings by similarities between the statistical distributions of the melodies. Melodies possessing similar rhythms (duration and duration transition distributions) and similar tone transitions were judged to be similar. Different weighting schemes were also investigated by weighting the melodic events according to their durations and creating hierarchical reductions of the melodies. This resulted in a slight increase in the fit between the perceptual judgments and the computational similarity between the melodies.

The link between melodic similarity and expectancy was not explicitly tested but the perceived similarities of the melodies could be explained by descriptive variables that were derived from expectation literature. These variables, such as number of tones, rhythmic variability, and melodic predictability accounted for a larger (55%) portion of the similarity ratings than the frequency-based variables. In other words, melodies possessing similar expectation structures were perceived as more similar than those which differed in this respect. Correspondingly, expectations are influenced by perceiving similarity relations within the melody (phrase and pattern repetition, for example) and between the melody and schematic knowledge.

Study VI examined the dynamic nature of melodic expectations. First of all new experimental methodology was developed for collecting continuous measurements of melodic expectations. In this method listeners rated the predictability of a melody simultaneously while the melody was being played. This approach made it possible to capture the dynamic nature of musical expectations and permitted the use of realistic musical contexts. Secondly the continuous ratings were predicted by dynamic models based on a combination of data-driven and schema-driven musical information.

Two behavioural experiments were carried out in which the melodies varied in their predictability: Experiment 1 contained 40 real melodies, mostly folk melodies from the Essen folk song collection, although song melodies by Charles Ives were used as well, whereas Experiment 2 consisted of 27 artificial, computer-generated isochronous sequences. Listeners' predictability ratings of the melodies in both experiments were predicted with on the basis of three models. Model 1 was based on discrete musical events and consisted of the principles in Narmour's implication-realization model (1990) supplemented by tonal stability (Krumhansl & Kessler, 1982) and melodic anchoring (Lerdahl, 1996). The two other models were also based on discrete events, with reference to pitch-class and interval respectively, but the models were conceived as continuous time-domain representations. Both models had a data-driven and a schema-driven component. The data-driven component represented the events of the melody as it progresses and imitated short-term memory by incorporating exponential decay of the events prior to the current moment. The second component corresponded to the knowledge of typical patterns (either pitch-class or interval distributions) in music, implemented as probabilities of events in the particular data-driven feature in question. Hence, the interplay of these two components created dynamics that were not present in the rule-based model. Model 2 was called the *interval distribution entropy model*, in which the data-driven component was based on local interval distribution of the stimulus with decay function and the schema-driven component was the probabilities of intervals derived from a large corpus of folk melodies. Model 3 (*pitch-class distribution entropy model*) was a variant of Model 2, in which the pitch-class distribution was used instead of the interval distribution. In both cases, the predictability of each moment in time was defined as the inverse of the entropy of the sum of the components.

The results suggested that continuous rating of melodic predictability is a feasible methodology as intersubject agreement amongst the participants was high. The combination of the rule-based heuristics (Model 1) was outperformed by the two dynamic models. In Experiment 1, the explanatory rates of the models were 5%, 15%, and 41%, for Models 1, 2 and 3, respectively. The same values for Experiment 2 were 8%, 22%, and 53%. Moreover, a linear combination of dynamic models accounted for substantial portions of the variance in the predictability ratings of experiments 1 and 2 (44% and 67%, respectively). The success of the dynamic models over the rule-based heuristics was concluded to lie in their better context-sensitivity. It is worth mentioning that the lower predictive rate of all the models in the case of real melodies (Experiment 1) was presumed to occur due to the cues to predictability provided by such structural features as pattern repetitions, cadences, and phrase structures. These factors draw attention to one of the central limitations of the models, namely the problems of segmentation, key-finding and pattern detection, factors which were not considered by the models although they seemed to have an effect on the predictability ratings given by the participants.

7 CONCLUSIONS

Musical expectancy is a fundamental aspect of the experience of listening to music. Expectancy is a multi-faceted phenomenon consisting of different musical features generated by various cognitive processes that operate on various levels and interact in intricate ways, bestowing upon musical experiences intensity and variety. In this chapter, conclusions are first drawn from the empirical studies, then limitations and methodological considerations are presented, and finally some applications of musical expectations are discussed.

The aim of the series of studies reported in this thesis was to investigate musical expectancy, a field of research which provides ways to understand the processes and knowledge listeners use to structure, interpret, remember and perform music. The central issues can be divided into (1) evaluation of the expectation models and their dependence on data-driven and schema-driven processes (cultural background and stylistic knowledge) (I-IV), (2) assessment of melodic complexity and similarity that either result from expectation violations (IV) or influence expectations (V) using a statistical approach, and (3) a dynamic approach to expectations (VI).

The first issue concerns expectation processes and how they are affected by cultural background. It is assumed that the structural design of the human information processing system relies on processing principles which are assumed to be common to the members of all cultures. The results of the studies support the view that expectations are governed by common data-driven processes although schema-driven knowledge has a distinct effect on these processes. The former were represented by the model of melodic expectation devised by Narmour (1990), event-frequency models and the auditory memory model. Of these the first two were robust in explaining the expectations experienced by listeners from distinct cultures in several experiments. Schema-driven knowledge was found to be important in forming melodic expectations, both when listening to music that is familiar and unfamiliar, employing different schematic expectations in each case. Various statistical models derived from the analysis of musical styles or empirical experiments represented schema-driven knowledge. Schema-driven, stylistic knowledge can be regarded as a slow form of dynamics where the expectations change as the experience of a musical style is accumulated. In future, more effort should be put into developing more comprehensive, style-specific models.

In approaching the second issue expectations were used as a tool for assessing melodic complexity. This investigation found commonalities between listeners from separate cultural backgrounds. Melodic similarity was modelled using self-similarity of melodic contour and statistical features of melodic events that are useful in creating expectations. Although the statistical measures were able to capture some basic aspects of similarity relations in music, these low-level features may not account for higher-level schemata that are often used in similarity judgements. For example, for European listeners, the similarity between *La Marseillaise*, *God Save the Queen*, and *O Tannenbaum* may be identified on the basis of their functions, thereby rendering the first two national anthems apparently similar despite the actual dissimilarity in their low-level features. For American listeners, the similarity relations between these three songs may be somewhat different, as the Christmas carol *O Tannenbaum* has been adopted as the state song by four states in the Unites States. Melodic similarity also has implications for data-driven processes that rely on finding parallel relations in music such as segmentation and grouping.

With respect to the third issue, considerations was given to the dynamic aspects of expectations, whereby information about expected events in music is updated across time. The dynamic concept also covers different levels of expectations, as short-term dynamics (acoustic memory) and long-term dynamics (pitch patterns, motifs, melodic archetypes) create differing predictions. The necessary research into these interactions remains to be done and is contingent upon the understanding of the role of the individual components. Although experiments using continuous rating methodology were conducted only with relatively limited musical material, important consequences for future methodological development were brought to light by the study and the results showed up ways towards a more comprehensive account of the dynamic processes in musical expectancy.

The issue of processing predispositions (universals) in melodic expectations was briefly considered in studies I-IV as the cross-cultural comparisons brought out certain principles and models that operated regardless of the cultural background of the listeners. While it is plausible that certain universal predispositions for processing music might exist – especially considering the ubiquity of music in all cultures (Blacking, 1995), the discoveries of specialized systems for music processing using the methods of cognitive neuroscience (Peretz, 2001) and the robustness of infant's musical perception (Trehub, 2000) – the empirical findings of the studies presented in this thesis do not provide support for the notion of universality as such. This is because the heuristics captured by these so-called universal principles may reflect structural properties that are prevalent in the acoustic environment and especially in speech. In other words, the so-called universal principles may be learned similarly by listeners in different cultures due to similar regularities in the environment they are exposed to and similar physical constraints operative in the deciphering of regularities in the environment. Regardless of the possible universal principles, there is a great diversity of principles by which different cultures manipulate musical expectations, and the cultural context and the particular features that contribute to these should be carefully considered.

Cognitive principles found in melodic expectations seem to share features with general cognitive processes. The phenomena of expectancy in music ties in with the notion of mind as an expectation-generator that enables better survival in a dangerous environment. Music certainly makes great use of these fundamental anticipatory processes. Melodic expectations also show a relationship between data- and schema-driven processes similar to that which has been observed with language. Understanding the speech of another person entails the detection of vowels, consonants and syllables from the raw sound on the basis of their acoustic features, a feat that is aided by knowledge of what these features typically sound like. Simultaneously words and phrases are assembled from these acoustic representations in accordance with the regularities of the particular language, syntax and idiom, until finally the content of the speech is understood. In music the end product is rather different but the combination of the processes is comparable. As in music, speech perception shows similar signs of being based on robust processes that many researchers call innate or, less controversially, processes that humans have a predisposition towards learning (Pinker, 1994).

7.1 Methodological considerations

Inevitably in a topic of this size, some important themes have not been addressed. The empirical methodologies themselves contain significant inherent limitations as the experiments depend upon highly-controlled situations. Also the models are raw approximations of the assumed information. However, these very same features that can be considered limitations are the strengths of the cognitive approach, where alternative explanations are eliminated and simple and testable models are proposed. Another advantage of the cognitive approach to music research is to combine several methods and types of data, and gather participants from different musical backgrounds and cultures. Although in this work the range of musical backgrounds represented in the experiments was fairly large, the common factor between the participants was musical expertise. This is widely found in the literature, as the music students are easily available for experimental research related to music and they are well versed in the musical concepts used in the instructions for the experiments. However, the drawback is that the perspective represented by naïve music listeners is not covered. In work relating to expectations no large differences are assumed to exist between musicians and non-musicians in the case of data-driven processes (e.g., Krumhansl, 1995a, b, Schellenberg, 1996), if stylistic knowledge is discounted. As one of the key issues was to study the effects on expectations of increased knowledge of the musical style, it was convenient to use participants steeped in the particular music culture and style. However, it would have been possible to study non-musicians using the same methods, as the probe-tone method has been found meaningful and reliable with such listeners (Oram & Cuddy, 1995; Cuddy & Thompson, 1992; Cuddy & Badertscher, 1987).

The study of music cognition is dominated by a Western, or Eurocentric perspective. Although the work presented here attempted to broaden the scope by performing cross-cultural comparisons, the very same bias is apparent in the methodology and the choice of the musical material. This bias may also affect the very notions of melodic expectations and complexity as heard and rationally evaluated by listeners. For example, the excursions into the expectations of North Sami and South African listeners, while successful and providing meaningful and consistent data, may have covered only superficial details of their musical processing. For one thing body movements and motor theories were not considered at all, despite their apparent role in the musical processing, especially in the case of African music (a point already made over seven decades ago by Hornbostel, 1928). Thus the scope of the research ideas needs to be enlarged to cover more refined aspects of music including intonation, lyrics, language, movement, dance, and various contextual factors.

Another challenge was to study ongoing psychological processes. Probe-tone methodology requires interrupting the listening experience and it is difficult to estimate how much this methodological factor confounds the experience of expectations as compared with uninterrupted listening. Continuous methodologies obviously do not have this limitation but still the conscious rating of musical experience will not do justice to the many-sided characteristics of the ongoing experience. In future, different continuous measurements used together with other methodologies, such as brain recordings, might be especially fruitful in distilling the key aspects of musical expectations across time.

7.2 Applications

The topic of musical expectancy has theoretical implications for music cognition in general but also for work related to music and emotions, music theory, and music information retrieval.

Expectancy has been connected to musical preferences and emotions. According to Meyer (1956), denial of expectations is assumed to cause an affect. Since then supporting evidence for this idea has been provided by Sloboda (1991) and on the connection between emotions and melodic complexity by Balkwill & Thompson (1999). However, the association between expectancy and emotions is still largely theoretical as expectancy is not comprehensively understood. Expectancy also have another link with affects. Three decades ago, students in the field of empirical aesthetics sought to clarify the connection between complexity and affect by connecting complexity with the preference for stimulus (Berlyne, 1971; and for reviews of work in the musical domain, Fung, 1995). Investigations of these ideas have tended to support Berlyne's claim that there is what he called an inverted-U relationship between liking and complexity (North & Hargreaves, 1995). This link between complexity and preference offers other applications too. For example, a high degree of complexity interferes with information processing and interferes with the performance of cognitive tasks (e.g., Kiger, 1989; Furnham & Allass, 1999), including driving (e.g., Beh & Hirst, 1999), as well as changing subjective estimation of time (e.g., Bueno, Firmino & Engelman, 2002). From another angle, research into expectations may be employed to uncover the brain processes involved in listening to music (e.g., Patel & Balaban, 2000).

Expectancy features strongly in music theory (Schenker, Piston, Lerdahl & Jackendoff). The approach taken by music cognition may develop music theory by addressing the basic questions of how music is organised and how this can be rigorously tested. This would help music theory to avoid using music-related concepts for which there is little perceptual support. On the other hand, musical analysis and the notions of music theory have provided analytical techniques that have helped to illuminate the psychological foundations underlying musical structure. Without music-theoretical descriptions of musical structure, music cognition would not have had such a wealth of ideas and hypotheses as to how music works. Melodic expectations are a good example of this.

Melodic expectations also have practical repercussions for music information retrieval. A proper understanding of human music processing tendencies is essential to music information retrieval systems that aim to find music in human terms. Taking into account our cognitive operations – whether they reflect the conventions of Western music, or more style-specific constraints or merely the limitations enforced by our auditory systems – would not only save precious bandwidth by decoding the material in a more compact way but also provide a more meaningful way of retrieving music from databases. In addition, interactive music systems need to possess mechanisms to anticipate future events in order to process and plan responses ahead. The groundwork for such anticipatory systems can be based upon data-driven and schema-driven expectations.

7.3 Future directions

The dynamic and acoustic aspects of expectations deserve further attention in the future. Furthermore, an explicit connection between melodic contour, similarity and expectations needs to be established as the topic was only briefly touched upon here. With respect to the approach and the methodology, it would be ideal to work with more cultures in examining whether the claim as to culture-transcending properties of music really holds in different contexts. Regarding the modelling paradigm, a shift from an abstract, symbolic approach towards a position that takes more account of the properties of the auditory system as well as the body as a whole would be welcome. There is now a considerable body of evidence on expectations so it would be appropriate to focus next on the application of the findings, especially to the emotions.

The approach taken in this study is relevant to music scholarship in general. For a music historian, the cognitive approach offers the possibility to shed light on the lost practices of music history (e.g., Toivanen, 2001; Veltman & Huron, 2002). The role of expectations in such work may be connected with what listeners automatically attend to in music, how these processes operate and how they interact with the stylistic practices of a certain period. Many music anthropologists, musicologists and ethnomusicologists emphasize that the essential features of musical culture as a phenomenon do not lie in the similarities between cultures but rather in the unique features of a culture that can only be studied from within that culture. The cultural context is important, and in order to reduce ethnocentricity in the explanatory models, the research questions should also be derived from the unique properties of the musical cultures themselves.

Cross-cultural music cognition may also help to draw attention to musical practices and cultures that are poorly known and highlight their unique features. For music anthropologists, ethnomusicologists, and music theorists, music cognition offers powerful methods to understand the minds of others and how their similarities and differences can be made discernible in music. This reflects the motivation of cognitive research into musical behaviour, in which the study of common mental mechanisms is an appealing topic, despite –and due to – the fact that there are large differences in musical parameters across musical styles and cultures.

REFERENCES

Aarden, B. 2002. Expectancy vs. retrospective perception: Reconsidering the effects of schema and continuation judgments on measures of melodic expectancy. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th international conference on music perception and cognition* (pp. 469-472). Adelaide: Causal Productions.

Adachi, M., & Carlsen, J. C. 1994. Melodic expectancy development in musical children. In I. Dèliege (Ed.), *Proceedings of the Third International Conference for Music Perception and Cognition* (pp. 143-144). Liege, Belgium: ESCOM.

Adachi, M., & Carlsen, J. C. 1995. Measuring melodic expectancies with children. *Council for Research in Music Education, 127*, 1-7.

Anderson, J. R. 1983. *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Aslin, R. N., Saffran, J. R., & Newport, E. L. 1998. Computation of conditional probability statistics by human infants. *Psychological Science, 9*, 321-324.

Auhagen, W. 1994. *Experimentelle untersuchungen zur auditiven tonalitätsbestimmung in melodien*. Regensburg: Gustav Bosse Verlag.

Baddeley, A. 1986. *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. 1974. Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. VIII, pp. 47-90). New York: Academic Press.

Baily, J. 1985. Music structure and human movement. In P. Howell, I. Cross & R. West (Eds.), *Music structure and cognition* (pp. 237-258). London: Academic Press.

Baldwin, J. M. 1896. A new factor in evolution. *American Naturalist, 30*, 441-451.

Balkwill, L. L., & Thompson, W. F. 1999. A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*, 43-64.

Bartlett, F. C. 1932. *Remembering: A study in experimental and social psychology* (reprinted in 1977 ed.). Cambridge, UK: Cambridge University Press.

Bartlett, J. C., & Dowling, W. J. 1988. Scale structure and similarity of melodies. *Music Perception, 5*, 285-314.

Baruch, C., & Drake, C. 1997. Tempo discrimination in infants. *Infant Behavior and Development, 20*, 573-577.

Becker, J. 2000. Anthropological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 135-160). Oxford: Oxford University Press.

Beh, H. C., & Hirst, R. 1999. Performance on driving-related tasks during music. *Ergonomics, 42*, 1087-1098.

Berlyne, D. E. 1971. *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Berry, J. W., Poortinga, Y. H., Segall, M., & Dasen, P. R. 1992. *Cross-cultural psychology: Research and applications*. Cambridge, UK: Cambridge University Press.

Besson, M., & Faïta, F. 1995. An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1278-1296.

Besson, M., Faïta, F., Peretz, I., Bonnel, A. M., & Requin, J. 1998. Singing in the brain: Independence of lyrics and tunes. *Psychological Science, 9*, 494-498.

Besson, M., Faïta, F., & Requin, J. 1994. Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience Letters, 168*, 101-105.

Besson, M., & Macar, F. 1987. An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology, 24*, 14-25.

Bey, C., & McAdams, S. 2002. Schema-based processing in auditory scene analysis. *Perception and Psychophysics, 64*, 844-854.

Bharucha, J. 1984. Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16*, 485-518.

Bharucha, J. J. 1987. Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5*, 1-30.

Bharucha, J. J. 1991. Pitch, harmony and neural nets: A psychological perspective. In P. M. Todd & D. G. Loy (Eds.), *Music and Connectionism* (pp. 84-99). Cambridge, MA: MIT Press.

Bharucha, J. J. 1994. Tonality and expectation. In R. Aiello (Ed.), *Musical Perceptions* (pp. 213-239). Oxford: Oxford University Press.

Bharucha, J. J. 1996. Melodic anchoring. *Music Perception, 13*, 383-400.

Bharucha, J. J. 1999. Neural nets, temporal composites and tonality. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 413-440). New York: Academic Press.

Bharucha, J. J., & Stoeckig, K. 1986. Reaction time and musical expectancy: priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 403-410.

Bharucha, J. J., & Stoeckig, K. 1987. Priming of chords: spreading activation or overlapping frequency spectra? *Perception and Psychophysics, 41*, 519-524.
Bharucha, J. J., & Todd, P. M. 1989. Modeling the perception of tonal structure with neural nets. *Computer Music Journal, 13*, 44-53.
Bigand, E. 1993. The influence of implicit harmony, rhythm and musical training on the abstraction of `tension-relaxation schemas' in tonal musical phrases. *Contemporary Music Review, 9*, 123-137.
Bigand, E., Madurell, F., B., T., & Pineau, M. 1999. Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 184-197.
Bigand, E., & Parncutt, R. 1999. Perceiving musical tension in long chord sequences. *Psychological Research, 62*, 237-254.
Bigand, E., Parncutt, R., & Lerdahl, F. 1996. Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics, 58*, 125-141.
Bigand, E., & Pineau, M. 1997. Context effects on musical expectancy. *Perception and Psychophysics, 59*, 1098-1107.
Bigand, E., Poulin, B., Tillmann, B., Madurell, F., & D'Adamo, D. A. 2003. Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 159-171.
Bissell, A. D. 1921. *The role of expectation in music.* Unpublished doctoral dissertation, Yale University, New Haven.
Blacking, J. 1967. *Venda children's songs: A study in ethnomusicological analysis*. Johannesburg: Witwatersrand University Press.
Blacking, J. 1995. *Music, culture and experience*. London: University of Chicago Press.
Blum, S. 1991. European musical terminology and the music of Africa. In B. Nettl & P. V. Bohlman (Eds.), *Comparative musicology and anthropology of music: Essays on the history of ethnomusicology* (pp. 3-36). Chicago: University of Chicago Press.
Bod, R. 2002. Memory-based models of melodic analysis: challenging the gestalt principles. *Journal of New Music Research, 31*, 27-37.
Boden, M. A. 1990. Has AI helped psychology? In D. Partridge & Y. Wilks (Eds.), *The foundations of artificial intelligence. A source book* (pp. 108-111). Cambridge: Cambridge University Press.
Boltz, M. 1991. Some structural determinants of melody recall. *Memory & Cognition, 19*, 239-251.
Boltz, M. 1993. The generation of temporal and melodic expectancies during musical listening. *Perception and Psychophysics, 53*, 585-600.
Boltz, M. G., & Jones, M. R. 1986. Does rule recursion make easier to reproduce? If not, what does? *Cognitive Psychology, 18*, 389-431.
Bregman, A. S. 1990. *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
Bregman, A. S., & Campbell, J. 1971. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89*, 244-249.
Brown, H., & Butler, D. 1981. Diatonic trichords as minimal tonal cue-cells. *In Theory Only, 5*, 39-55.
Brown, S. 2000. The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music* (pp. 271-300). Cambdridge, MA: MIT Press.
Brown, S., Merker, B., & Wallin, N. L. 2000. An introduction to evolutionary musicology. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music* (pp. 3-24). Cambdridge, MA: MIT Press.
Bueno, J. L. O., Firmino, E. A., & Engelman, A. 2002. Influence of generalized complexity of a musical event on subjective time estimation. *Perceptual & Motor Skills, 94*, 541-547.
Butler, D. 1989. Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception, 6*, 219-242.
Butler, D. 1990. Response to Carol Krumhansl. *Music Perception, 7*, 325-338.
Butler, D. A., & Brown, H. 1984. Tonal structure versus function: Studies of the recognition of harmonic motion. *Music Perception, 2*, 6-24.
Cambouropoulos, E. 1997. Musical rhythm: A formal model for determining local boundaries, accents and metre in a melodic surface. In M. Leman (Ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology* (pp. 277-293). Berlin: Springer Verlag.
Camras, L., Holland, E. A., & Patterson, M. J. 1993. Facial expressions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 199–208). New York: Guildford.
Camurri, A., & Leman, M. 1997. Gestalt-based composition and performance in multimodal environments. In M. Leman (Ed.), *Music, Gestalt, and computing - studies in cognitive and systematic musicology* (pp. 495-508). Berlin: Springer-Verlag.

Carlsen, J. C. 1981. Some factors which influence melodic expectancy. *Psychomusicology, 1*, 12-29.
Carlsen, J. C., Divenyi, P. L., & Taylor, J. A. 1970. A Preliminary study of perceptual expectancy in melodic configurations. *Council for Research in Music Education Bulletin*4-12.
Carterette, E. C., & Kendall, R. A. 1999. Comparative music perception and cognition. In D. Deutsch (Ed.), *The psychology of music* (pp. 725-791). San Diego, CA: Academic Press.
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.
Chew, E. 2000. *Towards a mathematical model of tonality.* Unpublished doctoral dissertation, MIT, Cambridge, MA.
Chiappe, P., & Schmuckler, M. A. 1997. Phrasing influences the recognition of melodies. *Psychonomic Bulletin Review, 4*, 254-259.
Chun, M. M., & Jiang, Y. 1999. Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10*, 360-365.
Clarke, E. F. 1985. Structure and expression in rhythmic performance. In P. Howell, I. Cross & R. West (Eds.), *Musical Structure and Cognition* (pp. 209-236). London: Academic Press.
Clarke, E. F. 1987. Levels of structure in the organization of musical time. *Contemporary Music Review, 2*, 211-238.
Clarke, E. F. 1988. Generative principles in music performance. In J. Sloboda (Ed.), *Generative Processes in Music* (pp. 1-26). Oxford: The Clarendon Press.
Clarke, E. F. 1999. Rhythm and timing in music. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 473-500). New York: Academic Press.
Cohen, A. J. 1991. Tonality and perception: Musical scales primed by excerpts from the Well-Tempered Clavier of J. S. Bach. *Psychological Research, 28*, 255-270.
Cohen, A. J., Thorpe, L. A., & Trehub, S. E. 1987. Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology, 41*, 33-47.
Cohen, J. E. 1962. Information theory and music. *Behavioural Science, 7*, 137-163.
Cole, M. 1996. *Cultural psychology: A once and future discipline.* Cambridge, MA: Harvard University Press.
Cook, N. 1987. The perception of large-scale tonal closure. *Music Perception, 5*, 197-206.
Cook, N. 1994. Perception: A perspective from music theory. In R. Aiello & J. Sloboda (Eds.), *Musical Perceptions* (pp. 64-95). New York: Oxford University Press.
Cook, P. R. (Ed.). 1999. *Music, cognition, and computerized sound: An introduction to psychoacoustics*. Cambridge, MA: Mit Press.
Coons, E., & Kraehenbuehl, D. 1958. Information as a measure of structure in music. *Journal of Music Theory, 2*, 127-161.
Cooper, G. W., & Meyer, L. B. 1960. *The rhythmic structure of music*. Chicago, IL: University of Chicago Press.
Cosmides, L., & Tooby, J. 2000. Consider the source: The evolution of adaptations for decoupling and metarepresentations. In D. Sperber (Ed.), *Metarepresentations: A Multidisciplinary Perspective* (pp. 53-115). Oxford: Oxford University Press.
Cuddy, L. L. 1993. Melody comprehension and tonal structure. In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 19-38). Hillsdale, NJ: Lawrence Erlbaum Associates.
Cuddy, L. L., & Badertscher, B. 1987. Recovery of tonal hierarchy: Some comparisons across age and levels of musical expertise. *Perception and Psychophysics, 41*, 609-620.
Cuddy, L. L., Cohen, A., & Mewhort, D. J. 1981. Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance, 7*, 869-883.
Cuddy, L. L., Cohen, A. J., & Miller, J. 1979. Melody recognition: the experimental application of musical rules. *Canadian Journal of Psychology, 33*, 148-157.
Cuddy, L. L., & Lunney, C. A. 1995. Expectancies generated by melodic intervals: Perceptual judgments in melodic continuity. *Perception & Psychophysics, 57*, 451-462.
Cuddy, L. L., & Thompson, W. F. 1992. Asymmetry of perceived key movement in chorale sequences: converging evidence from a probe-tone analysis. *Psychol Res, 54*, 51-59.
Cutler, A., Mehler, J., Norris, D., & Segui, J. 1987. Phoneme identification and the lexicon. *Cognitive Psychology, 19*, 141–177.
Daoussis, L., & McKelvie, S. J. 1986. Musical preferences and effects of music on a reading comprehension test for extraverts and introverts. *Perceptual and Motor Skills, 62*, 283-289.
Darwin, C. J., Turvey, M. T., & Crowder, R. G. 1972. An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory store. *Cognitive Psychology, 3*, 255-267.
Dawe, L. A., Platt, J. R., & Racine, R. J. 1994. Inference of metrical structure from perception of iterative pulses within time spans defined by chord changes. *Music Perception, 12*, 57-76.
Davidson, L., & Torff, B. 1992. Situated cognition in music. *World of Music, 34*, 120-139.
De Poli, G., Piccialli, A., & Roads, C. (Eds.). 1991. *Representations of musical signals*. Cambridge, MA: MIT Press.

Deliège, I. 1987. Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception, 4*, 325-360.
Deliège, I. 1996. Cue abstraction as a component of categorisation processes iin music listening. *Psychology of Music, 24*, 131-156.
Dennett, D. C. 1996. *Kinds of minds: Toward an understanding of consciousness*. New York: Harper Collins.
Densmore, F. 1972. *Yuman and Yaqui Music* (Original work published 1932). New York: Da Capo Press.
Deutsch, D., & Feroe, F. 1981. Internal representation of pitch sequence in tonal music. *Psychological Review, 88*, 503-522.
DeWitt, L. A., & Samuel, A. G. 1990. The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General, 119*, 123-144.
Dowling, W. J. 1967. *Rhythmic fission and the perceptual organization of tone sequences.* Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
Dowling, W. J. 1978. Dichotic recognition of musical canons: effects of leading ear and time lag between ears. *Perception and Psychophysics, 23*, 321-325.
Dowling, W. J. 1990. Expectancy and attention in melody perception. *Psychomusicology. Special Issue: Music expectancy, 9*, 148-160.
Dowling, W. J., & Bartlett, J. C. 1981. The importance of interval information in long-term memory for melodies. *Psychomusicology, 1*, 130-149.
Dowling, W. J., & Fujitani, D. S. 1971. Contour, interval, and pitch recognition in memory for melodies. *Journal of Acoustical Society of America, 49*, 524-531.
Dowling, W. J., Lung, K. M., & Herrbold, S. 1987. Aiming attention in pitch and time in the perception of interleaved melodies. *Perception and Psychophysics, 41*, 642-656.
Drake, C. 1998. Psychological processes involved in the temporal organization of complex auditory sequences: Universal and acquired processes. *Music Perception, 16*, 11-26.
Drake, C., & Baruch, C. 1995. From temporal sensitivity to models of temporal organization: Hypotheses and data on the acquisition of temporal abilities. *Année Psychologique, 95*, 555-569.
Dupoux, E., & Mehler, J. 1990. Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Journal of Memory and Language, 29*, 316-335.
Edworthy, J. 1985. Interval and contour in melody processing. *Music Perception, 2*, 375-388.
Eerola, T., & Toiviainen, P. 2003. *MIDI toolbox: MATLAB tools for music research*. Jyväskylä, Finland: University of Jyväskylä.
Essens, P. 1995. Structuring temporal sequences: comparison of models and factors of complexity. *Perception and Psychophysics, 57*, 519-532.
Feld, S. 1982. *Sound and sentiment*. Philadelphia: University of Pennsylvania Press.
Feld, S. 1984. Sound structure as social structure. *Ethnomusicology, 28*, 383-410.
Fodor, J. 1983. *The modularity of mind*. Cambridge, MA: MIT Press.
Forster, K. I., & Davis, C. 1984. Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 680-698.
Fraisse, P. 1982. Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 149–180). New York: Academic Press.
Fung, C. V. 1995. Music preference as a function of musical characteristics. *The Quarterly Journal of Music Teaching and Learning, 6*, 30-45.
Furnham, A., & Allass, K. 1999. The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts. *European Journal of Personality, 13*, 27-38.
Gabrielsson, A., & Juslin, P. N. 2003. Emotional expression in music. In H. H. Goldsmith, R. J. Davidson & K. R. Scherer (Eds.), *Handbook of affective sciences* (pp. 503-534). New York: Oxford University Press.
Gardner, H. 1987. *The Mind's New Science*. New York: Basic Books.
Gjerdingen, R. O. 1988. *A classic turn of phrase: Music and the psychology of convention*. Philadelphia: University of Pennsylvania Press.
Gjerdingen, R. O. 1991. Defining a prototypical utterance. *Psychomusicology, 10*, 127-139.
Gjerdingen, R. O. 1992. Learning syntactically significant temporal patterns of chords: a masking field embedded in an ART 3 architecture. *Neural Networks, 5*, 551-564.
Gotlief, H., & Konecni, V. J. 1985. The effects of instrumentation, playing style, and structure in the Goldberg variations by Johann Sebastian Bach. *Music Perception, 3*, 87-102.
Gourlay, K. A. 1984. The non-universality of music and the universality of non-music. *The world of music, 26*, 25-36.

Granot, R., & Donchin, E. 2002. Do re mi fa sol la ti—constraints, congruity, and musical training: an event-related brain potentials study of musical expectancies. *Music Perception, 19*, 487–528.
Gregory, A. H., & Varney, N. 1996. Cross-cultural comparisons in the affective response to music. *Psychology of Music, 24*, 47-52.
Griffith, N., & Todd, P. M. (Eds.). 1998. *Musical networks: Parallel distributed perception and performance*. Cambridge, MA: MIT Press.
Halpern, A. R., & Zatorre, R. J. 1999. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex, 9*, 697-704.
Handel, S. 1989. *Listening: An introduction to the perception of auditory events*. Cambridge, MA: The MIT Press.
Hantz, E. C., Kreilick, K. G., Kananen, W., & Swartz, K. P. 1997. Neural responses to melodic and harmonic closure: An event-related-potential study. *Music Perception, 15*, 69-98.
Hasher, L., & Zacks, R. T. 1984. Automatic processing of fundamental information. *American Psychologist, 39*, 1372–1388.
Herndon, M., & Mcleod, N. 1981. *Music as culture* (2nd ed.). Darby, PA: Norwood Editions.
Hershman, D. P. 1995. Rhythmic factors in tonality. *Psychomusicology, 14*, 4-19.
Hewlett, W. B., & Selfridge-Field, E. (Eds.). 1998. *Melodic similarity: Concepts, procedures, and applications*. Cambridge, MA: MIT Press.
Hevner, K. 1936. Experimental studies of the elements of expression in music. *American Journal of Psychology, 48*, 246-268.
von Hippel, P. 2000. Redefining pitch proximity: Tessitura and mobility as constraints on melodic interval size. *Music Perception, 17*, 315-327.
von Hippel, P. 2002. Melodic-expectation rules as learned heuristics. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002* (pp. 315-317). Adelaide: Causal Productions.
von Hippel, P., & Huron, D. 2000. Why do skips precede reversals? The effect of tessitura on melodic structure. *Music Perception, 18*, 59-85.
Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. 1989. Development of rehearsal in short term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology, 7*, 347-362.
Hofstetter, F. T. 1979. The nationalist fingerprint in nineteenth-century Romantic chamber music. *Computers and the Humanities, 13*, 105-119.
Holleran, S., Jones, M. R., & Butler, D. 1995. Perceiving implied harmony: The influence of melodic and harmonic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 737–753.
Honing, H. 1993. Issues in the representation of time and structure in music. *Contemporary Music Review, 9*, 221-239.
Hornbostel, E. M. v. 1928. African Negro Music. *Africa, 1*, 30-62.
Hughes, D. 1991. Grammars in non-Western musics: a selective survey. In P. Howell, R. West & I. Cross (Eds.), *Representing musical structure* (pp. 327-362). London: Academic Press.
Huron, D. 1992. Design principles in computer-based music representation. In A. Marsden & A. Pople (Eds.), *Computer representations and models in music* (pp. 5-39). London: Academic Press.
Huron, D. 1996. The melodic arch in Western folksongs. *Computing in Musicology, 10*, 3-23.
Huron, D. 2001. Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception, 19*, 11-64.
Huron, D., & Parncutt, R. 1993. An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology, 12*, 154-171.
Igaga, J. M., & J., V. 1977. Cultural differences in rhythmic performance. *Psychology of Music, 5*, 23-27.
Igaga, J. M., & Versey, J. 1978. Cultural differences in rhythmic perception. *Psychology of Music, 6*, 61-64.
Jackendoff, R. 1987. *Consciousness and the computational mind*. Cambridge, MA: MIT Press.
Janata, P. 1995. ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience, 7*, 153-164.
Janata, P. 2001. Neurophysiological mechanisms underlying auditory image formation in music. In R. I. Godøy & H. Jørgensen (Eds.), *Elements of Musical Imagery* (pp. 27-42). Lisse: Swets & Zeitlinger Publishers.
Janata, P., Birk, J. L., Tillmann, B., & Bharucha, J. J. 2002. Online detection of tonal pop-out in modulating contexts. *Music Perception, 20*, 283-305.

Jansen, E. L., & Povel, D. J. L. 1999. Mechanisms in the perception of accented melodic sequences. In *Proceedings of the 1999 Conference of the Society for Music Perception and Cognition* (pp. 29). Evanston, IL: SMPC.

Jansen, E. L., & Povel, D. J. L. 2000. The role of implied harmony in the perception of brief tone sequences. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 927-934). Keele, Staffordshire, UK: Keele University.

Johnson, M. H., & Morton, J. 1991. *Biology and cognitive development: The case of face recognition*. Oxford, UK: Blackwell.

Jones, A. M. 1959. *Studies in African Music*. London: Oxford University Press.

Jones, M. R. 1981. Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology, 1*, 34-51.

Jones, M. R. 1982. Music as a stimulus for psychological motion: Part 2. An expectancy model. *Psychomusicology, 2*, 1-13.

Jones, M. R. 1987. Dynamic pattern structures in music: Recent theory and research. *Perception and Psychophysics, 41*, 621-634.

Jones, M. R. 1990. Learning and the development of expectancies: An interactionist approach. *Psychomusicology. Special Issue: Music expectancy, 9*, 193-228.

Jones, M. R., & Boltz, M. 1989. Dynamic attending and responses to time. *Psychological Review, 96*, 459-491.

Jones, M. R., Boltz, M. G., & Kidd, G. 1982. Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics, 32*, 211-218.

Jones, M. R., & Pfordresher, P. Q. 1997. Tracking melodic events using Joint Accent Structure. *The Canadian Journal of Experimental Psychology, 51*, 271-291.

Jusczyk, P. W., & Krumhansl, C. L. 1993. Pitch and rhythm patterns affecting infant's sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 627-640.

Karno, M., & Konecni, V. J. 1992. The effects of structural interventions in the first movement of Mozart's symphony in G minor K.550 on aesthetic preference. *Music Perception, 10*, 63-72.

Keil, C., & Feld, S. 1994. *Music grooves: Esssays and dialogues*. Chicago, IL: University of Chicago Press.

Kelly, M. H., & Martin, S. 1994. Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua, 92*, 105–140.

Kessler, E. J., Hansen, C., & Shepard, R. N. 1984. Tonal schemata in the perception of music in Bali and the West. *Music Perception, 2*, 131-165.

Kiger, D. M. 1989. Effects of music information load on a reading-comprehension task. *Perceptual and Motor Skills, 69*, 531-534.

Kim, Y. E., Chai, W., Garcia, R., & Vercoe, B. 2000. Analysis of a contour-based representation for melody. In *International Symposium on Music Information Retrieval*. Plymouth, MA: Indiana University.

Knopoff, L., & Hutchinson, W. 1983. Entropy as a measure of style: The influence of sample length. *Journal of Music Theory, 27*, 75-97.

Knoppoff, L., & Hutchinson, W. 1981. Information theory for musical continua. *Journal of Music Theory, 25*, 17-44.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. 2000. Brain indices of music processing: 'non-musicians' are musical. *Journal of Cognitive Neuroscience, 12*, 520–541.

Koelsch, S., Schroger, E., & Tervaniemi, M. 1999. Superior pre-attentive auditory processing in musicians. *Neuroreport, 10*, 1309-1313.

Kohonen, T. 1997. *Self-organizing maps* (2nd ed.). Springer Series in Information Sciences. Berlin: Springer.

Krumhansl, C. 1979. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11*, 346-374.

Krumhansl, C. L. 1990. *Cognitive foundations of musical pitch*. Volume 17 of Oxford Psychology Series. Oxford: Oxford University Press.

Krumhansl, C. L. 1991. Music Psychology: Tonal Structures in Perception and Memory. *Annual Review of Psychology, 42*, 277-303.

Krumhansl, C. L. 1995a. Effects of Musical Context on Similarity and Expectancy. *Systematische musikwissenschaft, 3*, 211-250.

Krumhansl, C. L. 1995b. Music psychology and music theory: Problems and prospects. *Music Theory Spectrum, 17*, 53-90.

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. An exploratory study of musical emotions and physiology. *Canadian Journal of Experimental Psychology, 51*, 336-352.

Krumhansl, C. L. 1998. Topic in music: An empirical study of memorability, openness and emotion in Mozart's string quintet in C major and Beethoven's string quartet in A minor. *Music Perception, 16*, 119-134.

Krumhansl, C. L. 2000. Tonality induction: A statistical approach applied cross-culturally. *Music Perception, 17*, 461-479.

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

Krumhansl, C. L., & Keil, F. C. 1982. Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition, 10*, 325-334.

Krumhansl, C. L., & Kessler, E. J. 1982. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334-368.

Krumhansl, C. L., Sandell, G. J., & Sergeant, D. C. 1987. The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception, 5*, 153-184.

Krumhansl, C. L., & Shepard, R. N. 1979. Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance, 5*, 579-594.

Krumhansl, C. L., & Toiviainen, P. 2000. Dynamics of tonality induction: a new method and a new model. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 1504-1513). Keele, Staffordshire, UK: Keele University.

Kruskal, J. B., & Wish, M. 1978. *Multidimensional scaling*. Beverly Hills, CA: Sage University Series.

Lamont, A., & Cross, I. 1994. Children's cognitive representations of musical pitch. *Music Perception, 12*, 27-55.

Lang, P. J., Simons, R. F., & Balaban, M. T. (Eds.). 1997. *Attention and orienting: Sensory and motivational processes*. Hillsdale, NJ: Erlbaum.

Large, E., & Kolen, J. 1994. Resonance and the perception of musical meter. *Connection Science, 6*, 177-208.

Large, E. W., & Kolen, J. F. 1999. Resonance and the perception of musical meter. In N. Griffith & P. Todd (Eds.), *Musical Networks: Parallel Distributed Perception and Performance* (pp. 279-312). Cambridge, MA: MIT Press.

Larson, S. 1993. Modeling melodic expectation: Using three "musical forces" to predict melodic continuations. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 629–634). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larson, S. (1994). *Musical forces, step collections, tonal pitch space, and melodic expectation.* Paper presented at the Third International Conference on Music Perception and Cognition, Liège.

Larson, S. 1997. Continuations as completions: Studying melodic expectation in the creative microdomain Seek Well. In M. Leman (Ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology* (pp. 321-334). Berlin: Springer Verlag.

Larson, S. 2002. Musical forces, melodic expectation, and jazz melody. *Music Perception, 19*, 351-385.

Laske, O. 1988. Introduction to Cognitive Musicology. *Computer Music Journal, 12*, 43-57.

Leman, M. 1993. Symbolic and subsymbolic description of music. In G. Haus (Ed.), *Music processing* (pp. 118–164). Madison, WI: A-R Editions.

Leman, M. 1994. Schema-based tone center recognition of musical signals. *Journal of New Music Research, 23*, 169-204.

Leman, M. 1995. *Music and schema theory: Cognitive foundations of systematic musicology*. Berlin: Springer.

Leman, M. (Ed.). 1997. *Music, Gestalt, and computing - studies in cognitive and systematic musicology*. Berlin: Springer-Verlag.

Leman, M. 1999. Naturalistic approaches to musical semiotics and the study of causal musical signification. In I. Zannos (Ed.), *Music and Signs – Semiotic and Cognitive Studies in Music* (pp. 11-38). Bratislava: ASCO Art and Science.

Leman, M. 2000. An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception, 17*, 481-509.

Leman, M., & Carreras, F. 1997. Schema and Gestalt: Testing the hypothesis of psychoneural isomorphism by computer simulation. In M. Leman (Ed.), *Music, Gestalt, and computing - studies in cognitive and systematic musicology* (pp. 144-168). Berlin: Springer-Verlag.

Lemström, K., Wiggins, G. A., & Meredith, D. 2001. A Three-Layer Approach for Music Retrieval in Large Databases. In *Proceedings of ISMIR 2001, 2nd Annual International Symposium on Music Information Retrieval* (pp. 13-14). Bloomington, Indiana: ISMIR.

Lerdahl, F. 1996. Calculating tonal tension. *Music Perception, 13*, 319-363.

Lerdahl, F., & Jackendoff, R. 1983. *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Longuet-Higgins, H. C., & Steedman, M. J. 1971. On interpreting Bach. *Machine Intelligence, 6*, 221-241.
Lynch, M. P., & Eilers, R. E. 1991. Children's perception of native and nonnative musical scales. *Music Perception, 9*, 121-132.
Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. 1990. Innateness, experience, and music perception. *Psychological Science, 1*, 272-276.
Madsen, C. K. 1997. Emotional response to music as measured by the two-dimensional CRDI. *Journal of Music Therapy, 34*, 187-199.
Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. 2001. Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience, 4*, 540-545.
Magill, J. M., & Pressing, J. L. 1997. Asymmetric cognitive clock structures in West African rhythms. *Music Perception*, *15*, 189-222.
Mandler, G. 1975. *Mind and emotions*. New York: Wiley.
Mandler, J. M. 1984. *Stories, scripts and scenes: Aspects of schema theory*. New Jersey: Lawrence Erlbaum associates.
Mandler, J. M., & Johnson, N. S. 1977. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*, 111-151.
Marillier, C. G. 1983. Computer Assisted Analysis of Tonal Structure in the Classical Symphony. *Haydn Yearbook, 14*, 187-199.
Marr, D. 1982. *Vision: A computational investigation into human representation and processing of visual information*. San Francisco: W. H. Freeman.
McAdams, S. 1989. Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review, 4*, 181-198.
Merriam, A. P. 1964. *The anthropology of music*. Chicago: Northwestern University Press.
Merriam, A. P., Whinery, S., & Fred, B. G. 1956. Songs of a Rada community in Trinidad. *Anthropos, 51*, 157-174.
Meyer, L. B. 1956. *Emotion and meaning in music*. Chicago: University of Chicago Press.
Meyer, L. B. 1973. *Explaining music: Essays and explorations*. Chicago: University of Chicago Press.
Meyer, L. B. 1989. *Style and music: Theory, history, and ideology*. Philadelphia: University of Pennsylvania Press.
Meyer, R. K., Palmer, C., & Mazo, M. 1998. Affective and coherence responses to Russian laments. *Music Perception*, *16*, 135-150.
Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
Miller, G. A., & Heise, G. A. 1950. The trill threshold. *Journal of Acoustical Society of America, 22*, 637-638.
Mitroudot, L. 2001. Infant's melodic schemas: analysis of song productions by 4 and 5-year old subjects. *Musicae Scientiae, 5*, 83-104.
Monahan, C. B., Kendall, R. A., & Carterette, E. C. 1987. The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics, 41*, 576-600.
Morehen, J. 1981. Statistics in the analysis of musical style. In *Proceedings of the Second International Symposium on Computers and Musicology* (pp. 169-183). Paris: CNRS.
Munroe, R. H., Munroe, R. L., & Whiting, B. B. 1981. *Handbook of cross-cultural human development*. New York: Garland STPM Press.
Nam, U. 1998. Pitch distribution in Korean court music: Evidence consistent with tonal hierarchies. *Music Perception, 16*, 243-247.
Narmour, E. 1990. *The Analysis and cognition of basic melodic structures: The Implication-realization model*. Chicago, IL: University of Chicago Press.
Narmour, E. 1991. The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception, 9*, 1-26.
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* (2nd ed., pp. 441-472). San Diego: Academic Press.
Narmour, E. 2000. Music expectation by cognitive rule-mapping. *Music Perception, 17*, 329-398.
Neisser, U. 1976. *Cognition and reality*. San Francisco, CA: Freeman.
Nettheim, N. 1993. The pulse in German folksong: a statistical investigation. *Musikometrika, 5*, 69-89.
Nielsen, F. V. 1983. *Oplevelse af musikalsk spænding [The experience of musical tension]*. Copenhagen, Denmark: Akademisk Forlag.
North, A. C., & Hargreaves, D. J. 1995. Subjective complexity, familiarity and liking for popular music. *Psychomusicology, 14*, 77-93.
North, A. C., & Hargreaves, D. J. 1997. Liking for musical styles. *Musicae Scientiae, 1*, 109-128.

Näätänen, R., & Winkler, I. 1999. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull, 125*, 826-859.

Olson, J. M., Roese, N. J., & Zanna, M. P. 1996. Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211-238). New York: Guilford.

Oram, N., & Cuddy, L. L. 1995. Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research, 57*, 103-118.

Oxford English Dictionary. 1989. (2nd ed.). OED Online: Oxford University Press.

Paller, K. A., McCarthy, G., & Wood, C. C. 1992. Event-related potentials elicited by deviant endings to melodies. *Psychophysiology, 29*, 202-206

Palmer, C. 1996. Anatomy of a performance: sources of musical expression. *Music Perception, 13*, 433-454.

Palmer, C., & Krumhansl, C. L. 1987. Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 116-126.

Palmer, C., & Krumhansl, C. L. 1990. Mental representation for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 728-741.

Papousek, M. 1996. Intuitive parenting: a hidden source of musical stimulation in infancy. In I. Deliège & J. Sloboda (Eds.), *Musical beginnings* (pp. 88-112). Oxford: Oxford University Press.

Parncutt, R. 1989. *Harmony: a psychoacoustical approach*. London: Springer-Verlag.

Parncutt, R. 1994. A perceptual model of pulse salience and metrical accent in musical rhythm. *Music Perception, 11*, 409-464.

Parncutt, R., & Bregman, A. 2000. Tone profiles following short chord progressions: Top-down or Bottom-up? *Music Perception, 18*, 25-58.

Parsons, D. 1975. *The directory of tunes and musical themes*. Cambridge, UK: Spencer Brown.

Patel, A. D., & Balaban, E. 2000. Temporal patterns of human cortical activity reflect tone sequence structure. *Nature, 404*, 80-84.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. 1998. Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience, 10*, 717–733.

Peretz, I. 1990. Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113*, 1185-1205.

Peretz, I. 2001. Brain specialization for music: New evidence from congenital amusia. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (Vol. 930, pp. 153-165). New York: New York Academy of Sciences.

Peretz, I., & Kolinsky, R. 1993. Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Quarterly Journal of Experimental Psychology, 46*, 301-325.

Pinker, S. 1994. *The language instinct*. New York: William Morrow.

Pinker, S. 2002. *The blank slate. The modern denial of human nature*. New York: Viking.

Pinkerton, R. C. 1956. Information theory and melody. *Scientific American, 194*, 77-86.

Piston, W. 1978. *Harmony* (4th ed.). New York: W. W. Norton.

Plamondon, L. 1995. *Melodic expectancies and familiarization in music perception.* Unpublished doctoral dissertation, University of California, Berkeley.

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

Pont, G. 1990. Geography and human song. *New Scientist, 20*, 38-41.

Povel, D. J. 1981. Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception & Performance, 7*, 3-18.

Povel, D. J., & Essens, P. 1985. Perception of temporal patterns. *Music Perception, 2*, 411-440.

Povel, D. J., & Jansen, E. 2001. Perceptual Mechanisms in Music Processing. *Music Perception, 19*, 169-198.

Povel, D. J., & Jansen, E. 2002. Harmonic factors in the perception of tonal melodies. *Music Perception, 20*, 51-85.

Pressing, J. 1983. Cognitive isomorphisms in world music: West Africa, the Balkans, Thailand and Western tonality. *Studies in Music, 17*, 38-61.

Rameau, J. P. 1971. *Treatise of harmony* (P. Gosset, Trans. Original work published 1722). New York: Dover.

Regnault, P., Bigand, E., & Besson, M. 2001. Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory Event Related brain Potentials. *Journal of Cognitive Neurosciences, 13*, 1-17.

Riemann, H. 1895. *Harmony simplified* (H. Bewerunge, Trans.). London: Augener.

Ringer, A. L. (2003). *Melody. Definition and origins.* Grove Music Online, (Ed.) L. Macy. Retrieved 13 March, 2003, from http://www.grovemusic.com

Roh, S. K., & Yi, S. W. 2000. The effects of musical styles and experience on melodic expectancy. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 1635-1643). Keele, Staffordshire, UK: Keele University.

Rosch, E. 1975. Cognitive reference points. *Cognitive Psychology, 7*, 532-547.

Rosch, E., & Lloyd, B. B. 1978. *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosner, B. S., & Meyer, L. B. 1982. Melodic processes and the perception of music. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 317-341). New York: Academic Press.

Rosner, B. S., & Meyer, L. B. 1986. The perceptual roles of melodic process, contour, and form. *Music Perception, 4*, 1-39.

Rumelhart, D. E. 1975. Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding* (pp. 211–236). New York: Academic Press.

Rumelhart, D. E. 1980. Schemata: The building blocks of cognition. In R. Spiro, B. Bruce & W. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Russel, P., Deregowski, J., & Kinnear, P. 1997. Perception and aesthetics. In J. W. Berry, P. R. Dasen & T. S. Saraswathi (Eds.), *Handbook of Cross-Cultural Psychology, Volume 2: Basic Processes and Human Development* (pp. 107-136). Boston: Allyn and Bacon.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition, 70*, 27-52.

Sampat, K. S. 1978. Categorical perception in music and music intervals. *Journal of the Indian Musicological Society, 9*, 32-35.

Schank, R. C., & Abelson, R. P. 1977. *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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 expectancy. *Music Perception, 14*, 295-318.

Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. 2002. Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General, 131*, 511-537.

Schenker, H. 1954. *Harmony* (E. M. Borgese, Trans. Original work published 1906). Chicago, IL: University of Chicago Press.

Schenker, H. 1979. *Free Composition* (E. Oster, Trans. Original work published 1935). New York: Longman.

Schmidt, B.-H., Gunter, T. C., & Kotz, S. A. 2002. Expectancy in familiar melodies: an ERP Investigation. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002* (pp. 716-719). Adelaide: Causal Productions.

Schmuckler, M. A. 1988. *Expectation in music: Additivity of melodic and harmonic processes.* Unpublished doctoral dissertation, Cornell U, NY, USA.

Schmuckler, M. A. 1989. Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*, 109-150.

Schmuckler, M. A. 1990. The performance of global expectations. *Psychomusicology, 9*, 122-147.

Schmuckler, M. A. 1997. Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology, 51*, 292-305.

Schmuckler, M. A., & Boltz, M. G. 1994. Harmonic and rhythmic influences on musical expectancy. *Perception and Psychophysics, 56*, 313-325.

Schoen, D., & Besson, M. 2002. Processing pitch and duration in music reading: A RT-ERP study. *Neuropsychologia, 40*, 868-878.

Schoenberg, A. 1978. *Theory of Harmony* (R. E. Carter, Trans. Original work published in 1911). London: Faber.

Schubert, E. 1999. Measuring emotion continuously: Validity and reliability of the two dimensional emotion space. *Australian Journal of Psychology, 51*, 154-165.

Schubert, E. 2001a. Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 393-414). London: Oxford University Press.

Schubert, E. 2001b. Correlation analysis of continuous emotional response to music: Correcting for the effects of serial correlation. *Musicae Scientiae, Special Issue: Current trends in the study of music and emotion*, 213-236.

Shepard, R. 1982. Geometrical approximations to the structure of musical pitch. *Psychological Review, 89*, 305-333.

Shweder, R. A. 1991. *Thinking through cultures – expeditions in cultural psychology*. Cambridge, MA: Harvard University Press.

Simonton, D. K. 1980. Thematic fame and melodic originality in classical music: a multivariate computer-content analysis. *Journal of Personality, 48*, 206-219.
Sloboda, J. A. 1985. *The musical mind. The cognitive psychology of music.* Oxford Psychology Series No. 5. Oxford: Clarendon Press.
Sloboda, J. A. 1991. Music structure and emotional response: Some empirical findings. *Psychology of Music, 19*, 110-120.
Sloboda, J. A. 2000. Musical performance and emotion: Issues and developments. In S. W. Yipp (Ed.), *Music, Mind, and Science* (pp. 220-238). Soeul, Korea: Western Music Research Institute.
Sloboda, J. A., & Parker, D. H. H. 1985. Immediate recall of melodies. In P. Howell, I. Cross & R. West (Eds.), *Musical structure and cognition* (pp. 143-167). London: Academic Press.
Smith, K. C., & Cuddy, L. L. 1986. The pleasingness of melodic sequences: Contrasting effects of repetition and rule-familiarity. *Psychology of Music, 14*, 17-32.
Smith, K. C., & Cuddy, L. L. 1989. Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance, 15*, 457-471.
Smith, N. A., & Schmuckler, M. A. 2000. Pitch-distributional effects on the perception of tonality. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 563). Keele, Staffordshire, UK: Keele University.
Snyder, J. L. 1990. Entropy as a measure of musical style: the influence of a priori assumptions. *Music Theory Spectrum, 12*, 121-160.
Speer, J. R., & Meeks, P. U. 1985. School children's perception of pitch in music. *Psychomusicology, 5*, 49-56.
Stobart, H., & Cross, I. 1994. Aspects of rhythmic structure and perception in the music of northern Potosí, Bolivia. In *Proceedings of the Third International Conference on Music Perception and Cognition* (pp. 393-395). Liège, Belgium: ICMPC.
Stobart, H., & Cross, I. 2000. The Andean Anacrusis? Rhythmic structure and perception in Easter songs of Northern Potosí, Bolivia. *British Journal of Ethnomusicology, 9*, 63-94.
Taraban, R., & McClelland, J. L. 1988. Constituent attachment and thematic role assignment in sentence processing: influences of content-based expectations. *Journal of Memory and Language, 27*, 597-632.
Tekman, H. G., & Bharucha, J. J. 1998. Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 252-260.
Tekman, H. G., & Hortaçsu, N. 2002. Aspects of stylistic knowledge: What are different styles like and why do we listen to them? *Psychology of Music, 30*, 28-47.
Temperley, D. 2000. Meter and grouping in African music: A View from music theory. *Ethnomusicology, 44*, 65-96.
Temperley, D. 2001. *The cognition of basic musical structures*. Cambridge: MIT Press.
Temperley, D., & Sleator, D. 1999. Modeling meter and harmony: A preference-rule approach. *Computer Music Journal, 23*, 110-127.
Tenney, J., & Polansky, L. 1980. Temporal gestalt perception in music. *Journal of Music Theory, 24*, 205–241.
Tervaniemi, M. 1999. Pre-attentive processing of musical information in the human brain. *Journal of New Music Research, 28*, 237-245.
Tervaniemi, M. 2001. Musical sound processing in the human brain. Evidence from electric and magnetic recordings. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music* (Vol. 930, pp. 259-272). New York: New York Academy of Sciences.
Tervaniemi, M., & Huotilainen, M. in press. Neuroscience of music – methods and discoveries.
Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. 1997. The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters, 226*, 1-4.
Tervaniemi, M., Ilvonen, T., Sinkkonen, J., Kujala, A., Alho, K., Huotilainen, M., et al. 2000. Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioral evidence. *Neuroscience Letters, 279*, 29-32.
Tervaniemi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. 1999. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage, 9*, 330-336.
Tervaniemi, M., Winkler, I., & Näätänen, R. 1997. Pre-attentive categorization of sounds by timbre as revealed by event-related potentials. *Neuroreport, 8*, 2571–2574.
Thompson, E. F., Cuddy, L. L., & Plaus, C. 1997. Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody production task. *Perception and Psychophysics, 59*, 1069-1076.

Thompson, W. F., & Stainton, M. 1996. Using Humdrum to analyze melodic structure: An assessment of Narmour's Implication-Realization model. *Computing in Musicology, 10*, 24-33.

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.

Tillmann, B., Bharucha, J. J., & Bigand, E. 2000. Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*, 885-913.

Tillmann, B., Bigand, E., & Pineau, M. 1998. Effects of global and local contexts on harmonic expectancy. *Music Perception, 16*, 99-117.

Tillmann, B., Janata, P., & Bharucha, J. J. 2003. Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16*, 145–161.

Todd, N. P. M. 1994. The auditory primal sketch: A multi-scale model of rhythmic grouping. *Journal of New Music Research, 23*, 25-70.

Todd, N. P. M., Lee, C. S., & J., O. B. D. 2002. A sensory-motor theory of temporal tracking and beat induction. *Psychological Research, 66*, 26-39.

Toivanen, P. 2001. *The Pencerdd's toolkit - Cognitive and musical hierarchies in medieval Welsh harp music*. Jyväskylä Studies in the Arts nro 78. Jyväskylä, Finland: Jyväskylä University.

Toiviainen, P. 1996. *Modelling musical cognition with artificial neural networks*. Jyväskylä Studies in the Arts 51. Jyväskylä, Finland: University of Jyväskylä.

Toiviainen, P., & Eerola, T. 2001. A method for comparative analysis of folk music based on musical feature extraction and neural networks. In *Proceedings of the VII International Symposium on Systematic and Comparative Musicology and III International Conference on Cognitive Musicology* (pp. 41-45). Jyväskylä, Finland: ICSCM.

Toiviainen, P., & Eerola, T. 2002. A computational model of melodic similarity based on multiple representations and self-organizing maps. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002* (pp. 236-239). Adelaide: Causal Productions.

Toiviainen, P., & Eerola, T. 2003. Where is the Beat?: Comparison of Finnish and South African Listeners. In R. Kopiez, A. C. Lehmann, I. Wolther & C. Wolf (Eds.), *Proceedings of the 5th Triennial ESCOM Conference* (pp. 501-504). Hanover, Germany: Hanover University of Music and Drama.

Toiviainen, P., & Krumhansl, C. L. 2003. Measuring and modeling real-time responses to music: the dynamics of tonality induction. *Perception, 32*, 741-766.

Trainor, L. J., Desjardins, R. N., & Rockel, C. 1999. A comparison of contour and interval processing in musicians and nonmusicians unsing Event-related potentials. *Australian Journal of Psychology, 51*, 147-153.

Trainor, L. J., & Trehub, S. E. 1992. A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception & Performance, 18*, 394-402.

Trehub, S. 2000. Human processing predisposition and musical universals. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music* (pp. 421-448). Cambdridge, MA: MIT Press.

Trehub, S., Schellenberg, E., & Hill, D. 1997. The origins of music perception and cognition: A developmental perspective. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 103-128). East Sussex, UK: Psychology Press.

Trehub, S. E., & Thorpe, L. A. 1989. Infants' perception of rhythm: categorization of auditory sequences by temporal structure. *Canadian Journal of Experimental Psychology, 43*, 217-229.

Trehub, S. E., & Trainor, L. J. 1993. Listening strategies in infancy: The roots of music and language development. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 278-327). New York: Oxford University Press.

Treisman, A. M., & Rostron, A. B. 1972. Brief auditory storage: A modification of Sperling's paradigm applied to audition. *Acta Psychologica, 36*, 41-60.

Tulving, E., & Schacter, D. L. 1990. Priming and human memory systems. *Science, 247*, 301-306.

Unyk, A. M., & Carlsen, J. C. 1987. The influence of expectancy on melodic perception. *Psychomusicology, 7*, 3-23.

Walker, R. 1996. Open peer commentary: Can we understand the music of another culture? *Psychology of Music, 24*, 103-130.

Walker, R. 2002. Culture is the real driver in music perception. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002* (pp. 122-124). Adelaide: Causal Productions.

Van Noorden, L. 1975. *Temporal coherence in the perception of tone sequences.* Unpublished doctoral dissertation, Eindhoven University of Technology, Netherlands.

Van Noorden, L., & Moelants, D. 1999. Resonance in the perception of musical pulse. *Journal of New Music Research, 28*, 43-66.
Watt, H. J. 1924. Functions of the size of interval in the songs of Schubert and of the Chippewa and Sioux Indians. *British Journal of Psychology, 14*, 370-386.
Veltman, J., & Huron, D. 2002. A Cognitive Approach to the Medieval Modes. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 277-280). Adelaide, Australia: Causal Productions.
Verleger, R. 1990. P3-evoking wrong notes: Unexpected, awaited, or arousing? *International Journal of Neuroscience, 55*, 171-179.
Vos, P. G., & Troost, J. M. 1989. Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception, 6*, 383-396.
Vos, P. G., van Assen, M., & Franek, M. 1997. Perceived tempo change is dependent on base tempo and direction of change: Evidence for a generalized version of Schulze's (1978) internal beat model. *Psychological Research/Psychologische Forschung, 59*, 240-247.
Vos, P. G., & Verkaart, P. P. 1999. Inference of mode in melodies. *Music Perception, 17*, 223-239.
Yeston, M. 1976. *The stratification of musical rhythm*. New Haven, CT: Yale University Press.
Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. 1996. Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience, 8*, 29-46.
Zbikowski, L. M. 1999. Musical coherence, motive, and categorization. *Music Perception, 17*, 5-42.

YHTEENVETO (SUMMARY)

Musiikillisten odotusten tarkastelu kulttuurienvälisten vertailujen ja tilastollisten mallien avulla

Odotukset ja eri asioiden ennakoiminen ovat ensiarvoisen tärkeitä toimintamme kannalta, niin musiikissa kuin muussakin elämässä. Musiikilliset odotukset luovat kuulijalle ajatuksia erilaisista vaihtoehdoista, miten musiikki voisi jatkua. Odotukset ovat olennaisia kuuntelutilanteessa, mutta niillä on merkittävä rooli myös soitettaessa, improvisoitaessa ja säveltäessä. Tämä väitöskirja tutkii, mitkä asiat vaikuttavat kuulijoiden melodisiin odotuksiin.

Melodiaa käsittelevät odotukset ovat osittain kulttuurisidonnaisia, eli riippuvaisia kuulijan musiikillisesta taustasta, kokemuksista ja opitusta tyylintuntemuksesta. Tällöin esimerkiksi luterilaisten virsien odotetaan etenevän musiikillisesti eri tavoin kuin vaikkapa blues-sävelmän.

Toisaalta melodisia odotuksia ohjaavat myös kulttuuritaustasta riippumattomat, universaalit lainalaisuudet. Nämä psykologiset periaatteet ovat automaattisia ja liittyvät sellaisiin tekijöihin kuten sävelten ajallisen etäisyyden, sävelkorkeuksien läheisyyden tai vaikkapa hyvän jatkuvuuden ja symmetrian periaatteiden hahmottamiseen.

Lisäksi odotuksiin vaikuttaa lyhytkestoinen muisti, kaikki se mitä musiikissa on juuri tapahtunut musiikin tyylistä tai odotusten automaattisista periaatteista huolimatta. Tämän, musiikin tilastollisiin piirteisiin perustuvan tapahtumamuistin avulla ihmisille syntyy odotuksia dynaamisesti, äskettäin kuulemaansa pohjaten. Näin esimerkiksi tietyn melodisen rakenteen toistuminen lisää kuulijan odotuksia siitä, että samanlaista rakennetta tullaan käyttämään myös jatkossa. Tämä musiikinkuuntelun aikana muodostuva käsitys musiikin etenemisestä yhdessä opittujen tyylillisten säännönmukaisuuksien sekä erilaisten automaattisten havaintoperiaatteiden kanssa vaikuttaa siis siihen, kuinka kuulija havaitsee musiikillisia tapahtumia ja odottaa niiden jatkuvan. Juuri näiden erilaisten tietämisen lajien – opittujen, automaattisten ja sävelmateriaalista nousevien tietorakenteiden – vaikutusta kuuntelutapahtumaan tutkittiin tässä väitöksessä.

Väitöskirja perustuu kuuteen erilliseen tutkimukseen, joissa melodioiden synnyttämiä odotuksia tutkittiin neljän eri kulttuuripiirin musiikilla: saamelaisjoiuilla, Länsi-Suomen rukoilevaisten virsillä, keskieurooppalaisilla kansansävelmillä sekä eteläafrikkalaisilla kansanlauluilla. Samoin olivat myös koehenkilöt peräisin näistä äsken mainituista kulttuuripiireistä poiketen toisistaan lisäksi ko. tyylintuntemuksen osalta. Nämä vertailut tarjosivat mahdollisuuden etsiä melodisiin odotuksiin vaikuttavia seikkoja, jotka olivat joko kulttuurisista tekijöistä riippuvaisia tai niistä riippumattomia.

Tutkimuksen ensimmäinen osio käsittelee tyylintuntemuksen psykologisten lainalaisuuksien sekä melodisen tapahtumamuistin vaikutusta melodisiin odotuksiin. Näitä tietämisen lajeja on tutkittu kulttuurienvälisen vertailun avulla osatutkimuksissa I-IV. Tämän kaltaisella vertailulla voidaan osoittaa, millaisia odotuksia koetaan samalla tavoin. Nämä universaalit, eri kulttuuritaustan omaavien henkilöiden samanlaiset odotukset voidaan erottaa itse musiikista nousevista tai tyylintuntemukseen liittyvistä odotuksista. Tutkimukset I-III to-

teutettiin nk. *koetinsävelmenetelmää* käyttäen, jossa kuulijat arvioivat, kuinka hyvin jokin sävel heidän mielestään jatkaa annettua sävelmää. Kokeiden perusteella huomattiin yleisen, psykologisiin lainalaisuuksiin pohjautuvan mallin sekä musiikin tilastollisten piirteiden selittävän verrattain hyvin kuulijoiden odotuksia, tosin tyylin tietämys vaikutti näiden kahden selitysmallin painotuksiin. Tutkimuksessa IV käsiteltiin sitä, millä tavoin odotusten rikkominen hankaloittaa melodian hahmottamista. Kokeiden perusteella huomattiin mm. melodiakaarrosten toistuvuuteen ja rytmiin liittyvien tekijöiden merkitys melodioiden hahmottamisessa. Vertailtaessa afrikkalaisten ja eurooppalaisten kuulijoiden vastauksia eroja havaittiin erityisesti liittyen rytmisiin tekijöihin. Yleisesti ottaen oli positiivista huomata, että erilaisista kulttuuritaustoista huolimatta koehenkilöiden oli mielekästä ja helppoa vastata annettuihin tehtäviin.

Tutkimuksen toinen osio tarkentaa erilaisia odotusten syntymiseen liittyviä lisäkysymyksiä. Melodisten hahmojen muodostuminen ja niiden toiston havaitseminen vaikuttavat myös odotuksiin. Näitä hahmoja tunnistetaan samanlaisuussuhteiden avulla, mitä selvitettiin osatutkimuksessa V, jossa kuulijat arvioivat melodiaparien samankaltaisuutta. Melodioiden samanlaisuusarvioiden voidaan katsoa syntyvän musiikillisten piirteiden yhtäläisyyksistä, joita mallinnettiin musiikin tilastollisten piirteiden avulla. Osatutkimuksessa VI aiemmin mainittua koetinsävelmenetelmää kehitettiin dynaamisempaan suuntaan siten, että kuulijat arvioivat musiikin ennalta-arvattavuutta musiikin soidessa. Näille nk. *jatkuville arvioille* etsittiin selitystä malleilla, joissa yhdistettiin lyhytkestoinen, sävelmateriaalista nouseva tietämys pitkäkestoisiin tietorakenteisiin, eli tonaliteettiin sekä tyypillisiin intervalli- ja motiivirakenteisiin nojaavaan musiikilliseen tietämykseen. Nämä dynaamiset mallit kykenivät selittämään verrattain hyvin kuulijoiden vastauksia musiikin ennustettavuudesta, mikä selittyy mallien kyvyllä ottaa huomioon musiikin tyypilliset rakenteet ja yhdistää ne musiikilliseen kontekstiin.

Väitöskirjan tuloksilla on teoreettista merkitystä musiikillisen kognition ymmärtämisessä, erityisesti selitettäessä, miten kulttuurisidonnaiset ja yleiset periaatteet tuottavat musiikillisia odotuksia. Odotusten eri periaatteita koskevia lainalaisuuksia voidaan yleistää sellaisiin sovelluksiin, joissa musiikillista tietoa on tarve digitaalisesti pakata ja etsiä tapoja, joilla entistä paremmin otetaan huomioon kuulijoiden tapa hahmottaa musiikkia. Kenties mielenkiintoisin näkökulma musiikillisten odotusten sovelluksista liittyy musiikin aiheuttamien tunnereaktioiden ymmärtämiseen, sillä odotuksilla leikittelyllä on keskeinen rooli mielenkiinnon ylläpitämisessä ja tunnekokemusten synnyttämisessä musiikkia kuunnellessa.

JYVÄSKYLÄ STUDIES IN HUMANITIES

1 KOSTIAINEN, EMMA, Viestintä ammattiosaamisen ulottuvuutena. - Communication as a dimension of vocational competence. 305 p. Summary 4 p. 2003.

2 SEPPÄLÄ, ANTTI, Todellisuutta kuvaamassa – todellisuutta tuottamassa. Työ ja koti televisi-on ja vähän radionkin uutisissa. - Describing reality – producing reality. Discourses of work and home in television and on a small scale in radio news. 211 p. Summary 3 p. 2003.

3 GERLANDER, MAIJA, Jännitteet lääkärin ja potilaan välisessä viestintäsuhteessa. - Tensions in the doctor-patient communication and relationship. 228 p. Summary 6 p. 2003.

4 LEHIKOINEN, TAISTO, Religious media theory - Understanding mediated fith and christian applications of modern media. - Uskonnollinen mediateoria: Modernin median kristilliset sovellukset. 341 p. Summary 5 p. 2003.

5 JARVA, VESA, Venäläisperäisyys ja ekspressiivisyys suomen murteiden sanastossa. - Russian influence and expressivity in the lexicon of Finnish dialects. 215 p. 6 p. 2003.

6 USKALI, TURO, "Älä kirjoita itseäsi ulos" Suomalaisen Moskovan-kirjeenvaihtajuuden alkutaival 1957–1975. - "Do not write yourself out" The beginning of the Finnish Moscow-correspondency in 1957–1975. 484 p. Summary 4 p. 2003.

7 VALKONEN, TARJA, Puheviestintätaitojen arviointi. Näkökulmia lukioikäisten esiintymis- ja ryhmätaitoihin. - Assessing speech communication skills. Perspectives on presentation and group communication skills among upper secondary school students. 310 p. Summary 7 p. 2003.

8 TAMPERE, KAJA, Public relations in a transition society 1989-2002. Using a stakeholder approach in organisational communications and relation analyses. 137 p. 2003.

9 EEROLA, TUOMAS, The Dynamics of Musical Expectancy. Cross-Cultural and Statistical Approaches to Melodic Expectations. - Musiikillisten odotusten tarkastelu kulttuurienvälisten vertailujen ja tilastollisten mallien avulla. 84 p. (277 p.) Yhteenveto 2 p. 2003.