Temporal Expectations

and their Violations

Olivia Ladinig

Temporal Expectations

and their Violations

ILLC Dissertation Series DS-2009-09

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 904
1098 XH Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

Temporal Expectations

and their Violations

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. dr. D.C. van den Boom
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op dinsdag 27 oktober 2009, te 10.00 uur

door

Olivia Ladinig

geboren te Klagenfurt, Oostenrijk

Promotiecommissie:

Promotor: Prof. dr. R. Scha
Co-promotor: Dr. H. Honing

Overige leden:
Prof. dr. R. Bod
Prof. dr. R. de Groot
Prof. dr. C.J. ten Cate
Dr. B.H. Repp
Drs. M. Sadakata

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The investigations were supported by the EmCAP (Emergent Cognition through Active Perception) project funded by the European Commission (FP6-IST, contract 013123).

Copyright © 2009 by Olivia Ladinig

Cover design by Olivia Ladinig.
Printed and bound by Printpartners Ipskamp.

ISBN: 978-90-5776-200-0

Contents

Acknowledgments

I am very grateful to

Henkjan Honing, my supervisor, for teaching me how to become a researcher, always believing in me, and making this thesis possible; Leigh Smith, my unofficial second daily supervisor, for countless talks about the significance of doing research and for sharing his vast knowledge about the world; Remko Schah, my promotor, for his valuable suggestions on the conceptual level and his liberal way of advising; the EmCAP project, for supporting me financially and providing inspiring interactions with an extraordinary group of researchers; the ILLC, for accepting people surfing on the margin of their wave; Bruno Repp, the most concise and disciplined person I know, for helping to make my scribbles look proper; István Winkler and Gábor Háden from the MTIP in Budapest, and Glenn Schellenberg, Patrick Hunter and Stephanie Stalinsky from the Department of Psychology in Toronto, for hosting me during my research visits and providing great grounds for collaborations; all the students that participated in a project or class I taught; all fellow PhD students, researchers, and awesome administrative staff from the ILLC, especially my dozen office mates; Alex, Amelie, Andreas, Ansten, Charlotte, Gaëlle, Herman, Jacob, Jonathan, Loesha, Luc, Nicole, Rebecca, Reinco, Salvador, Tracy, and Teun, for sharing energy, music, thoughts, travels, dances, and houses in the last four years; Alexandra, Arthur, Elena, Gabriele, Ilja, Marijan, Monika, Oliver, and Sabine, for keeping me connected to my motherland; the Occii, for promoting the alternative; Harry Partch, Wevie Stonder, and Frederic Chopin, for their other-wordly music; my father, mother, and brothers, for being loving, understanding, and supportive; Arnaud, for being a zebra without a single vice de fabrication; my wonderful grandmother, Josefa, for showing me how to live and die.

Amsterdam Olivia Ladinig
October, 2009.

Chapter 1

Introduction

1.1 The role of time in music listening

The term that is most commonly used to refer to the temporal dimension of music is *rhythm*. As Fraisse (1982) pointed out more than two decades ago, "the task of those who study rhythm is a difficult one, because a precise, generally accepted definition of rhythm does not exist. The difficulty derives from the fact that rhythm refers to a complex reality in which several variables are fused" (p. 149). His observation seems still to be valid, considering that even today authors still quote this remark, and establish their own definitions of rhythm that focus on the component that is relevant to their particular research question. For example, Patel (2008), after stating that "Unfortunately, there is no universally accepted definition of rhythm" (p. 96), defines rhythm as "the systematic patterning of sound in terms of timing, accent, and grouping" (p. 96). On the most basic and purely physical level, musical rhythm can be described in terms of inter-onset-intervals (IOIs), which are distances of onsets of sounding events along the temporal dimension (Clarke, 1999; Fraisse, 1982). A second parameter of rhythm, though usually seen as less important, is the duration of events, defined as the temporal distance between event onset and offset.

Not surprisingly, physical features are not alone responsible for what is perceived by a listener. The cognitive system tries to organise input in order to facilitate processing and memory. Thus the same sequence of events can be perceived differently - by different listeners, and even by one and the same listener at different times and in different contexts (Desain & Honing, 2003; Repp, 2007). For example, the context in which a rhythm is presented can result in the subjective emphasis of different events. In certain instances a sequence of events will receive subjective accentuation of every second or fourth event; in other instances every third or sixth event will receive subjective emphasis (see discussion of meter below). Furthermore, tempo is a crucial factor and determines if lower- or higher-order relationships between events are established. Notwithstanding these factors that can cause different interpretations of the same rhythmic pattern, the cognitive system is able to abstract invariances from a variety of features and can categorise two rhythms as being the same even if they are played in different tempi, or if they have different expressive timing.

One crucial mechanism that allows us to hear a sequence of events as a rhythm is the tendency to perceive a beat in a pattern of events, usually referred to as *beat induction*. (For a recent overview of research see Patel, 2008.) The ability to perceive a beat seems to be universal in humans and enables them to entrain to music, and to coordinate perception and behaviour. Although some animals show periodic behaviour and are able to produce relatively isochronous (i.e., equally spaced) sounds, it was for a long time believed that they are not able to synchronise with an audio signal, or to pick up regularities from a signal as complex as a musical rhythm. (For the recently fast growing discussion on that subject see Large, Velasco, & Gray, 2008; Patel, Iversen, Bregman, & Schulz, 2009; Schachner, Brady, Pepperberg, & Hauser, 2009.) Interestingly, not all occurrences of a perceived beat have to coincide with an actual event onset, but some can also be perceived in a moment of silence, a so-called *loud rest* in a rhythm. If people are asked to produce beats at a tempo they like (called spontaneous tapping rate, personal tempo, or referent level), the distribution of inter-onset-intervals produced by adults is usually centred around 600 ms (Drake, Jones, & Baruch, 2000; Fraisse, 1982). In this region listeners are also most sensitive in making perceptual assessments - for example, in judging slight tempo differences between two rhythmic patterns (Drake et al., 2000).

Perceiving a beat in a rhythm means to perceive certain events as accented or *salient* events, which do not necessarily have to be (but of course can be) physically accented. Besides an optional physical accentuation, the perception of beats is based on subjective accents that can arise solely from the listener's interpretation of a particular rhythmic structure. Lerdahl and Jackendoff (1983) distinguish three types of accent: accents that arise from the musical surface, such as tone duration or intensity (phenomenal accents), accents that arise from the musical structure, such as boundaries introduced by musical phrases or harmonic progressions (structural accents), and accents that themselves arise from the perception of a beat (metric accents). The first two kinds of accent are physical accents, since they are based on physical cues in the music. The last kind of accent, metric accents, can be seen as subjective accents, since often no physical cues are apparent in metrically accented events.

It is generally believed that based on the perception of subjective accents, each event of a rhythm is perceived as belonging to one particular level of salience, and beats are perceived as having higher salience than the events that occur between them, which in this dissertation are termed *subbeats*. This cognitive tendency of creating a structure with different levels of salience can be observed even with very simple rhythms, such as the ticking of a clock or metronome, where every other event (more rarely, every third or fourth) often receives a metric accent

(Bolton, 1894; Brochard, Abecasis, Potter, Ragot, & Drake, 2003). This phenomenon is termed subjective metricisation (London, 2004). In most common Western rhythms, subbeats (the weaker elements) divide the beat periods into two, three, or four equal intervals.

When at least two periodic levels of perceptual salience can be distinguished, as in the case of subjective metricisation, one speaks of meter perception or *meter induction* (Yeston, 1976). Lerdahl and Jackendoff (1983) state that "fundamental to the idea of meter is the notion of periodic alternations of strong and weak beats" (p. 19). (It should be noted that this terminology differs from the one used in this dissertation. In the present context, 'strong beats' are simply termed 'beats', and 'weak beats' are termed 'subbeats'.) The respective structural levels to which events are perceived as belonging are characterised by their *metric salience*. (Metric salience has to be distinguished from *event salience*, which depends on all three types of accent.) Musicological models of meter (Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1984) assign metric salience to *metric positions* in a hierarchical and recursive way, by dividing the musical measure into equal subparts: The more subdivisions have to be made to assign a position to a structural level, the lower the salience of an event occurring in this position. The first position, usually termed the downbeat, is assigned the highest possible salience of a measure, and the position right after it is by definition always one of the positions with the lowest possible salience. The musical measure itself is, of course, a notational representation of assumed or intended metric units, usually comprising several beats.

1.2 Formation and violation of temporal expectations

Beat induction exemplifies the active role of the listener in the interpretation of rhythm, which involves assigning different perceptual saliences to certain events, even if they are physically identical. This active role becomes even more striking if we consider the phenomenon of a loud rest, which is caused by an expectation for a particular position to have an event, which then does not happen. Many other examples can be found that show that listening to music is not purely passive. What makes perception active is that the cognitive system constantly makes predictions about future events, which amounts to having expectations. One crucial element in our appreciation of music is the balance between expectation fulfilment and violation (Berlyne, 1970; Huron, 2006; Meyer, 1965; Narmour, 1990). Highly predictable music might not be interesting for long, since it is too simple. Music that violates too many expectations will be too complicated for the average listener and sound random, thus also not being perceived as interesting.

Foundations for the study of rhythmic expectations in particular were laid by Mari Riess Jones and colleagues (Jones, 1976; Jones & Boltz, 1989; Large &

Jones, 1999), and are being further developed with neuroscientific methods. (For an overview of research see Zanto, Snyder, & Large, 2006.) Expectation violations result from the relation of event patterns to listeners' underlying cognitive schemata. The more expectations are violated, the higher the perceived complexity of the music. Events with a higher salience are expected more strongly, are better memorised and recalled, and receive the most immediate attention. Consequently the absence of salient events or the disproportionally high occurrence of non-salient events will lead to the impression of complexity (Fitch & Rosenfeld, 2007). In other words, perceived complexity is a good indicator of underlying expectations determined by the metric salience of events.

No generally accepted theory of rhythmic complexity exists yet. (For an overview see Streich, 2007.) However, Smith and Honing (2006) showed that the concept of *syncopation* developed by Longuet-Higgins and Lee (1984) can explain a considerable part of perceived complexity as reflected in listeners' judgments. Admittedly, the concept of complexity can not be entirely reduced to syncopation. Other factors like note density, tempo variation, absence of periodicity and repetition certainly play a role in perceived complexity. However, if as many of these factors as possible are held constant, strength of syncopation may explain the perception of rhythmic complexity. Syncopation is used throughout this dissertation as a way of operationalising complexity of rhythmic patterns. For musicians, syncopation has a relatively clear meaning: a subjective accent in a metrically unaccented position, or a rest in a metrically accented position. A formal definition is given by Longuet-Higgins and Lee (1984), who specify also which parts of a musical measure are accented, to what degree, and how strong a certain syncopation will be. The model of syncopation by Longuet-Higgins and Lee is based on the specification of metric salience for each event in a rhythm by hierarchically and recursively subdividing the measure into equally sized subparts. The metric saliences are then in turn used to calculate the strength of syncopation: A syncopation occurs if a sounding event occurs in a position that has lower metric salience than an immediately preceding position in which no event occurs. The strength of a syncopation is calculated as the difference in metric salience of those events.

The concepts discussed so far (beat, meter, syncopation) presume hierarchical processing of events. Another factor relevant to rhythm perception that was considered in the course of this dissertation is *serial position* within a relatively short rhythmic sequence (Martin, 1972; Jongsma, Desain, & Honing, 2004). The importance of events in comparison with each other may not only be based on hierarchical metric division of the musical measure, but also be related to their absolute placement within one sequence. Effects of serial position are known mainly from memory research (Acheson & MacDonald, 2009; Ebbinghaus, 1885), with events at the beginning and at the end of a sequence being recalled easier (termed *primacy* and *recency* effects, respectively). In this thesis serial position was considered as an additional factor that may determine salience of rhythmic events in a musical measure. The hypothesis was that events at the beginning and at the end of a musical measure would have a higher salience than events in the middle. These issues are addressed in Chapter 4.

1.3 Competence in music listening

There are ways to assess and classify someone's ability regarding music performance. One obvious source of information is the amount of time spent on practice and study of the particular skill in question, as assessed by questionnaires. Less established, however, is a way to assess *listening competence*, or the various capacities for processing music. In their recent review of musical capacities that do not depend on formal music training, Bigand and Poulin-Charronnat (2006) define listening competence as "perceiving the relationships between a theme and its variations, perceiving musical tensions and relaxations, generating musical expectancies, integrating local structures in large-scale structures, learning new compositional systems and responding to music in an emotional (affective) way" (p. 100) as examples of processing musical structure that underlies the musical surface. More systematic research has to be done to evaluate if listening competence grows mainly as a result of formal music training, or if extensive exposure to music can be sufficient to acquire high levels of musical competence via implicit learning.

Most measures used to assess a person's level of expertise are based on the years and intensity of continuous formal music education and practice, as well as on the initial age when the training started. An evaluation of measures used in previous research can be found in Ollen (2006), who reports that 'musical sophistication' is mainly acquired through formal music training and can be assessed by questions about the duration and intensity of the training, the age when the training started, the involvement in music theory courses, and the number of concerts attended on average. However, the question remains if these measures assess someone's competence in music listening. Apart from formal music training, every listener accumulates implicit knowledge based on extracting statistical regularities from simple exposure to music. Ethnomusicological and cross-cultural investigations consider the fact that the culture and musical tradition a listener is exposed to influence his or her perception and cognition, and effects of exposure to different musical traditions are commonly acknowledged. However, effects of exposure can also exist on a smaller scale, within one and the same culture, provided by the many different musical styles or genres available. Despite the fact that musical styles and genres from within one music tradition share basic tonal and rhythmic conventions, variations can be found on a more subtle level, for example in the use of timbre (e.g., acoustic vs. electronic music in contemporary Western culture) or the typical conventions of expressivity specific to different musical styles (for example drum rolls around or after the second snare sound in Breakcore, or the *pump up*, a rise of the key by one step, after a chorus in Pop music). Regarding rhythm, *expressive timing* refers to subtle variations in performance along the temporal dimension, for example 'tempo rubato', the local slowing down and speeding up (Repp, 1990, 1995), or playing slightly before or after the expected time of a note (Desain & Honing, 2003). Unusual expressive timing can make a particular performance distinctive. In Chapter 3 of this thesis, sensitivity of listeners to expressive timing in familiar and unfamiliar styles is assessed as a form of implicit learning.

Removing all effects of formal music training as well as long-term exposure would theoretically allow us to study innate and universal mechanisms regarding the perception and cognition of music. Though impossible to achieve this in laboratory settings with adult participants, in recent years the opportunity to study newborn infants has arisen. Such research can address questions about the origins of music, and about the contributions of nature and nurture to music perception and cognition. Contrary to what has long been believed, human minds do not seem to be tabulae rasae at the time of birth. Recent studies with babies a few days old have revealed that certain perceptual styles and cognitive predispositions govern our perception from very early on (Trehub, 2003; Winkler et al., 2003). Chapter 6 in this dissertation reports an attempt to determine if a basis for the perception of metric salience can be found already in newborn babies.

1.4 Research methods used in this thesis

This thesis reports research that employs both behavioural and electrophysiological measures. One reason for using a variety of methods is the need to assess processes occurring with and without the benefit of attention. Even though it is fruitful to study the effects of different foci of attention on perception and cognition, it is also highly informative to construct experimental conditions that lead to a withdrawal of attention from the material under study, in order to gain insights into automatic processes. Behavioural methods were used to obtain listeners judgments (Chapters 3 and 4) and to measure processing speed and discrimination accuracy (Chapter 5). In the research of Chapters 5 and 6, electrophysiological measures were used to investigate parts of the same processes in adults and infants, respectively, without the benefit of subjects' attention. The analysed component of the induced event-related potentials was the Mismatch Negativity (MMN), which in most cases is not modulated by different attentional states (Sussman, 2007). The next section gives a brief overview of the relevant research methods.

1.4.1 Behavioural methods

One behavioural method used in this dissertation was to collect listeners' judgments, which give indications about conscious assessments and choices regarding the presented stimuli. Judgments are based on reflective thoughts, and there exists the possibility to revise initial decisions before a final answer is reported. In the research described in Chapter 3, listeners made a decision in a two-alternative forced choice task according to the perceived naturalness of two stimuli, and in the research described in Chapter 4 listeners ranked elements within a set of stimuli according to the subjective degree of complexity.

A complementary approach to collecting judgements is to measure listeners' processing speed. Processing speed can be assessed by measuring the reaction time (RT) in experimental tasks. The speed of detecting a deviant stimulus reflects perception and processing while the participant is in an attentive state (Chapter 5). Reaction time is the duration between the presentation of a stimulus and the respective observable response. A fast reaction time can indicate that a deviation was stronger, or that a stimulus was more expected.

Listeners' sensitivity to differences between events can be expressed with the sensitivity index *d'*. This measure takes the correct responses (hit rates) as well as the incorrect responses (false alarms) into account. The higher the *d'*, the more sensitive is a subject in the detection of deviants or in discriminating different events.

1.4.2 Electrophysiological methods

Complementary to traditional behavioural methods, electrophysiological measurements provide a way to study processing without any behavioural responses. Electroencephalography (EEG) measures electricity from the scalp, which is generated by neuronal activity involved in sensation as well as cognition. EEG is an ideal method when studying temporal processing, since it provides high temporal resolution. However, spatial resolution is rather coarse. Electric activity that can be directly related to an event (a thought or a percept) is called an event-related potential (ERP). ERPs are visible as slow waves starting roughly at 100 msec after the event onset. Based on their time course, such waves can be decomposed into various components that have a specific time of occurrence and direction of amplitude, and are seen as stereotypical indicators of certain processes. In practice, ERPs can only be observed after averaging the time-locked data of many trials, since the level of noise (e.g., stemming from the EEG system itself, from spontaneous brain activity, or from body movements) is far beyond the level of the actual signal that has an amplitude of only a few millivolts.

One important component of the ERP was discovered by Näätänen, Gaillard, and Mäntysalo (1978) and termed the Mismatch Negativity (MMN). The MMN signals the detection of an irregular event in a string of regular events, and occurs

150-250 msec after the onset of the deviation. The common view is that the MMN reflects the process of testing a perceptual model against incoming stimuli, with a violation of the model being reflected in the MMN. The auditory MMN can be observed if the deviating event differs from the regular events in terms of frequency, amplitude, duration, or location. Bigger differences between a standard and a deviating event cause bigger and earlier MMN responses. An MMN response can also be observed in case an expected event is simply omitted. By comparing the latency and the amplitude of MMN responses to two different deviations, one can infer which one was more salient. For an overview of paradigms, results, and implications of research using MMN, see Näätänen, Paavilainen, Rinne, and Alho (2007).

The MMN is not only elicited by simple violations of physical properties. Deviations from higher-order regularities can also be reflected. For example, Saarinen, Paavilainen, Schröger, Tervaniemi, and Näätänen (1992) observed MMN responses to descending tone pairs occurring in a series of ascending tone pairs, while the starting tones of the pairs varied. Paavilainen, Simola, Jaramillo, Näätänen, and Winkler (2001) found responses to violations of rules based on relationships between sound attributes, for example "the higher the frequency of a stimulus, the higher the amplitude". Deviants in this case were either low frequency stimuli with high amplitude, or high frequency stimuli with low amplitude.

In our two studies using the MMN (reported in Chapters 5 and 6) we investigated this ability to detect regularities in abstract patterns, and used a set of rhythms with no omissions or with omissions in low-salience positions as standards and a set of rhythms with omissions in high-salience positions as deviants. If the brain creates hierarchical representations of the rhythmic sequences, omission of the most salient event (the downbeat) is expected to elicit stronger responses from participants than omission of a metrically less salient event.

Another useful feature of the MMN component is that it does not require attention to the task, and in most cases does not even benefit from it. This allows use of this method with sleeping newborns (Chapter 6), or with adults who perform some distraction task (Chapter 5).

1.5 Thesis purpose and outline

The purpose of this thesis was to study temporal expectations indirectly by observing reactions to violations of expectations. Expectations on different structural levels were considered, with expressive timing violations testing small-scale expectations, and event omissions testing larger-scale expectations. Special attention was given to listener-specific variables, namely level of formal music training, exposure, and developmental stage, as well as to the role of attention, using methods that monitor pre-attentive as well as attentive processing.

Chapter 2 is a short commentary on the validity of Internet experiments in music perception research. The use of the Internet for data collection is growing in various areas. Still, many scientific journals are sceptical towards this method, and until recently some had a policy to simply reject submissions that report data gathered on the Internet. We argue that the accessibility of a broad range of participants, the saving of time and money, and the increase of ecological validity together with the reduction of experimenter bias are points in favour of that method of data collection.

Chapter 3 reports an Internet experiment that explored the effects of extensive exposure to a certain musical style and of formal music training on a particular aspect of listening competence, namely the sensitivity to small deviations in performance. Scaling musical performances by speeding them up or slowing them down proportionally scales their expressive timing. Previous research has shown that proportionally scaled expressive timing does not sound natural. Participants, either expert musicians or non-musicians, and either familiar or unfamiliar with certain musical styles, were to distinguish between a tempo-transformed and a non-transformed musical performance of the same piece by focusing on the expressive timing. The hypothesis was that formal music training would not be a strong predictor of performance, but that the crucial factor would be the familiarity with a particular musical genre.

Chapter 4 describes an attempt to look at differences between musicians and non-musicians regarding the hierarchical depth of meter perception, expressed in terms of metric saliences. Also based on an Internet experiment, listeners rated the perceived complexity of various rhythms containing omissions in more or less salient positions. The degree of perceived complexity was expected to indicate the degree of expectation violation based on the underlying metric salience. In addition to metric processing, the effect of serial position was studied.

An alternative approach to study metric salience is reported in Chapter 5. Adult non-musicians were required to detect rhythms with two different strengths of syncopation within a string of non-syncopated rhythms, and their reaction time and discrimination sensitivity were measured. Electrophysiological data were gathered for the same stimuli, for two different attentional conditions. The hypothesis was that listeners would be faster and more accurate in behaviourally detecting stronger compared to weaker syncopations, and that stronger MMN responses would be observed for stronger compared to weaker syncopations.

The study reported in Chapter 6 employed a reduced version of the above-mentioned paradigm to study the highest level of meter violation, i.e., downbeat omission, in sleeping newborn infants, to see whether expectations based on metric salience are already active at a very early stage in human development.

The dissertation concludes with a discussion and an outlook on future research.

All chapters contain material that is either already published or has been peer-reviewed and is now under revision. The references to published or submitted articles are given below the respective chapter titles. In cases where articles require additions or corrections, these are reported as *Notes* at the end of the text of each chapter.

Chapter 2

Web-based music cognition research

Honing, H. & Ladinig, O. (2009). The potential of the Internet for music perception research: A comment on lab-based versus web-based studies. *Empirical Musicology Review, 3,* 4-7.

Abstract

While the discussion on the integrity of data obtained from Web-delivered experiments is mainly about issues of method and control (Mehler, 1999; McGraw, Tew, & Williams, 2000; Auditory, 2007), this comment stresses the potential that Web-based experiments might have for studies in music perception. It is argued that, due to some important advances in technology, Web-based experiments have become a reliable source for empirical research. Next to becoming a serious alternative to a certain class of lab-based experiments, Web-based experiments can potentially reach a much larger, more varied and intrinsically motivated participant pool. Nevertheless, an important challenge to Web-based experiments is to control for attention and to make sure that participants act as instructed; Interestingly, this is not essentially different from experiments that are performed in the laboratory. Some practical solutions to this challenge are proposed.

2.1 Commentary

Web-based experiments are not novel. Since the availability of the Internet several initiatives have been developed using it as an alternative to lab-based experiments (Birnbaum, 2004; Johnson-Laird & Savary, 1999; Klauer, Musch, & Naumer, 2000; Musch & Reips, 2000). However, in the domain of vision and audition the potential of Web-based experiments is still little used, if not simply avoided.

There are at least two reasons for this. First, there is some doubt on how much control there is over the participant population and their sampling, as compared to orthodox experiments (cf. Johnson-Laird & Savary, 1999, p. 221). In orthodox experiments much biographical data is available about the participants, while in Web experiments, it is argued, there is no such control, potentially allowing participants to 'conspire' to generate the date one needs (Mehler, 1999). A second reason for some conservatism in doing Web experiments is the issue of replicability. Especially in the fields of experimental psychology and psychophysics there are serious concerns about the (apparent) lack of control one has in Web experiments as opposed to those performed in the laboratory. Where in the lab most relevant factors, including all technical issues, are under control of the experimenter (i.e., have a high internal validity; Campbell & Stanley, 1963) it is argued that Web experiments lack this important foundation of experimental psychology (Mehler, 1999). As a result of the first issue, it often proves to be problematic to convince University Review Panels to give permission when there is little insight in the environment in which participants tend to do these experiments (Auditory, 2007). As a result of the second issue, some high-impact journals made it a policy decision not to publish Web-based studies, as such discouraging Web experiments to be performed.

Skeptics of Web-based studies are mainly concerned with the question of how sure an experimenter can be that participants do not conspire towards a certain result, or in other ways try to deceive the purpose of the experiment (Mehler, 1999, p. 188). However, it is not clear what would motivate participants to deceive. In a laboratory experiment as well, a typical paid participant could well, for example, just press buttons and take little care in doing the instructed task. Actually, it can be an advantage that there is no experimenter present, because having participants completing tasks in spite of their anonymity, which would make it easy at any point to drop out, can be considered a valuable sign of motivation. The absence of the experimenter also minimizes the performance according to social desirability, and eliminates possible experimenter biases or Pygmalion effects.

Nevertheless, in a Web experiment - just like in an orthodox experiment - one has to make sure the participants are doing what you asked them to do. One way of solving this issue is to ensure there is little reason for the participant to deceive.[1] Make the experiment challenging and fun to do, do not reward good answers (but simply participation), and make certain the participants feel involved. In music perception research this turns out to be relatively easy. Music lovers tend to like listening experiments and are usually very motivated, resulting in large numbers of responses (see Honing, 2006b, 2007; Honing & Ladinig, 2006,

[1] For most of the observations made in this comment, empirical support is (e.g., Reips, 2002) or can be made available. However, due to space restrictions we will not do this here.

2009).[2] Furthermore, to make sure that a potential participant (who is typically sitting at home behind a computer screen) gets involved in the experiment, one can, for instance, use a screen cast:[3] a video showing what to expect and that presents the instructions in a compelling way. It generally makes the participants feel more involved and motivates them in really taking note of the instructions, more than when they are simply asked to read text from the screen. And finally, as an extra incentive, we often use a raffle of gift certificates among all respondents, independent of their responses.

However, in Web experiments there is always a participant group that is either just curious, didn't intend to do the full experiment, or that is simply not serious. Hence, one of the tasks of the Web experimenter is to distinguish between serious and unserious responses. Dropout of non-serious participants (typically around 30-40%; Reips, 2002) includes people that did not finish the experiment, did not fill-in all required information, did it too quickly (e.g., didn't listen through the full sound examples when instructed as such), or those that did the experiment too hastily. By, for instance, inserting foils (e.g., a question to test whether a participant is paying attention), consistency checks, and time logs in the experiment (e.g., using Javascript) one can easily check for this and filter out the serious from the non-serious responses. To minimize the dropout of serious respondents, it is generally a good idea to (1) put a platform, browser and audio check first (to avoid frustration about possible technical problems later in the experiment, 2) put biographical questions at the beginning of the Web experiment (they tend to make the respondent feel more involved), followed by 3) engaging instructions (making sure the participants know what to expect), and 4) the actual experiment (that should not last longer than about fifteen minutes). In addition, it is important to give feedback about the results and the research context (what was the relevance of the experiment), along with the question of whether the respondent would like to participate in future experiments. The latter turns out to be a good index of seriousness, as well as an opportunity to build up a motivated participant pool that has the appropriate technical setup and a sincere interest in the topic area one is studying.

While for more socially oriented research the Web was used early on (Musch & Reips, 2000), for vision, audition and music perception studies the internet is only sparsely used. The main reason for this is, next to the issues discussed above, of a technical nature. One could, generally, not be sure of the video and sound quality at the user-end. However, several recent studies (e.g., Krantz, 2000; McGraw et al., 2000) showed that these technicalities are currently less of an issue. As an example, McGraw et al. (2000) found that, for some classical experiments in the visual domain, Web-based experiments are able to replicate the results ob-

[2] For the URLs to the online listening experiments on which these studies are based, see the reference section.

[3] For an example, see URL http://www.musiccognition.nl/e4/

tained in the laboratory. They concluded that "existing technology is adequate to permit Web delivery of many cognitive and social psychological experiments [..] The added noise created by having participants in different settings using different computers is easily compensated for by the sample sizes achievable with Web delivery." (McGraw et al., 2000, p. 502). With comparable advances in audio presentation over the internet (using, for instance, file formats like MPEG-4 that minimize loading time and guarantee optimal sound quality on different computer platforms at different transmission rates) there is little reason to think that Web experiments in music perception would be less reliable. Especially when the aim is to study music perception (as opposed to some psychophysical task that might have additional demands to the presentation of audio), the Web can even be preferred over lab-based experiments.

Web-based experiments have a much greater external validity as compared to lab-based experiments. While this sometimes results in losing some internal validity (cf. Auditory, 2007), in music perception studies this might actually be desirable: i.e., to have listeners respond in an environment in which they normally listen, including e.g., its noisiness and/or use of low-quality headphones or loudspeakers. For such a listener a lab situation - with high-quality audio, a soundproof booth, focused listening, and the pressure of having to perform - might be quite unnatural. In short, having listeners take part from their home in an arguably more natural environment and, as such, being less stressed, might actually positively influence the ecological validity of the results.

Another criticism of Web-based studies is that of the sample of participants. One can argue that participants with access to Internet belong to a special, technologically versatile, subgroup of the population, and that this could affect the representativity of a certain result. While this is changing rapidly (especially in the Western countries), it should be noted that such a (potential) restriction on the sample also applies to the typical psychology-student pool that biases most studies in psychology. As such this is not a specific drawback of Web-based studies.

In conclusion, while there are still plenty of challenges to online data collection, we believe that Web experiments generally do not generate more problems than an orthodox lab-based experiment. It might actually be a rich source for more ecologically valid and truly engaging studies in music perception. In the end it is up to the researcher to decide, and argue for each specific case, where to draw the line with regard to the trade-off between a higher internal validity of laboratory settings compared to a higher external validity of Web-based setups.

Chapter 3

Timing sensitivity

Honing, H. & Ladinig, O. (2009). Exposure influences timing judgments in music. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 281-288.[1]

Abstract

This study is concerned with the question whether, and to what extent, listeners' previous exposure to music in everyday life, and expertise as a result of formal musical training, play a role in making expressive timing judgments in music. This was investigated by using a Web-based listening experiment in which listeners with a wide range of musical backgrounds were asked to compare two recordings of the same composition (15 pairs, grouped in three musical genres), one of which was tempo-transformed (manipulating the expressive timing). The results show that expressive timing judgments are not so much influenced by expertise levels, as is suggested by the *expertise hypothesis*, but by exposure to a certain musical idiom, as is suggested by the *exposure hypothesis*. As such, the current study provides evidence for the idea that some musical capabilities are acquired through mere exposure to music, and that these abilities are more likely enhanced by active listening (exposure) than by formal musical training (expertise).

3.1 Introduction

The ability to make, perceive, and enjoy music is generally regarded as an evolutionary by-product of more important functions, such as those involved in lan-

[1]The experiment can be found at
http://cf.hum.uva.nl/mmm/drafts/EEE-online/EEE-index.html

guage (Pinker, 1997). However, there is increasing evidence that humans are born with musical biases and predispositions that are unique to human cognition (Hannon & Trehub, 2005; Peretz, 2006; Zatorre, 2005). Although it remains unclear whether this evidence can be interpreted as a sign that a capacity for music is rooted in nature, rather than nurture, there is little controversy around the idea that musical competence is a special human capacity that is shared across ages and cultures (Blacking, 1974; Jackendoff & Lerdahl, 2006; Mithen, 2005; Sloboda, 2000; Trehub, 2003). In the present article we concentrate on the question whether *musical competence* - the perceptual skills and musical knowledge that are required to perceive and appreciate musical input - is influenced by extensive formal musical training (explicit knowledge), or whether it can also be interpreted as a result of mere exposure to music (implicit knowledge).

Although some older studies argue that musical competence is a special, innate talent ('musicians are born, not made'; cf. Sloboda, 1994), the most common view is that musical abilities are shaped mostly by intense musical training (Dienes & Longuet-Higgins, 2004; Sloboda, 1994; Smith, 2002; Wolpert, 2000) and that they remain rather rough in untrained listeners (Jackendoff & Lerdahl, 2006). Some authors even suggest that after the age of 10, musical abilities no longer evolve without explicit musical training (Francès, Zenatti, & Imberty, 1979). These studies give support to the common idea that musicians, due to their specific musical talent and training, are more aware of musical detail (such as nuances in *expressive timing*,[2] discussed in the present study) than are average listeners (Sloboda, 1994). We refer to this view as the *expertise hypothesis*, in which explicit knowledge and extensive musical training are considered the main contributors to musical competence.

Another, more recent view is that listeners without formal musical training, when given sufficient exposure (e.g., listening to music in everyday life, moving and dancing to music, attending concerts) to a certain musical idiom, might actually perform similarly in a task when compared with musically trained listeners (Levitin, 2006; Schellenberg, 2006), especially when they are asked to do a musical task that uses realistic and ecologically valid stimuli. With regard to the latter, it could be argued that the differences in musical competence between musicians and non-musicians, as suggested by the literature, could well be an artifact of tasks using explicit naming - a situation in which musically trained listeners would have an advantage over untrained listeners. We refer to this view as the *exposure hypothesis*, in which implicit knowledge as a result of mere exposure (e.g., listening to one's preferred music) is considered the main contributor to musical competence.

An example in support of the exposure hypothesis is a study by Bigand and Poulin-Charronnat (2006), who discovered that non-musicians can be as sensitive

[2] *Expressive timing* is the term used to refer to the minute deviations from regularity that contribute to the quality of a music performance (Clarke, 1999; Palmer, 1997).

as musicians to subtle aspects of music harmony, suggesting that musical training and explicit knowledge of music theory are unnecessary to acquire sophisticated knowledge about melody and harmony (Bigand, Tillman, Poulin, D'Adamo, & Madurell, 2001).[3] Furthermore, prolonged exposure to a specific musical idiom seems to allow nonmusicians, without explicit knowledge about a certain musical genre, to internalize the rules that are typical to such a genre and do almost equally well as musicians in a comparison task. Dalla Bella and Peretz (2005) found that all listeners - musicians and nonmusicians alike - are sensitive to styles of Western classical music, arguing that this is supported by cross-cultural perceptual processes that allow for discrimination of key perceptual features.

In the present study we are interested in whether these recent findings (i.e., the effect of exposure on making sophisticated musical judgments) also hold in the temporal domain of music cognition.

3.2 Listening Experiment Using a Comparison Task

To study the effect of exposure and expertise in the temporal domain, we used a listening task that allows for testing the effect of different listener groups and different expertise levels on temporal sensitivity. In this task, participants were asked to compare two performances of the same composition (15 pairs, grouped in three musical genres: classical, rock, and jazz; see Table 3.1). Each stimulus pair consisted of an audio recording by one artist and a manipulated, tempo-transformed audio recording by another artist. The tempo-transformed version was originally performed at a different tempo, but was scaled to be similar in overall tempo to the other performance of the pair. This resulted in stimulus pairs that have the same tempo, one of which is not manipulated, the other tempo-transformed. The participants had to indicate which of the two stimuli sounds more "natural" or musically plausible by focusing on the expressive timing that could have been manipulated as a result of the tempo-transformation.

This particular task was used for a number of reasons. First, the use of different musical genres (rock, jazz, classical) allows every participant to be either explicitly or implicitly competent, through either formal training or listening experience, in at least one musical genre. Second, expressive timing tends to be characteristic for a particular genre.[4] In fact, Dalla Bella and Peretz (2005) showed that temporal variability can serve as an index to mark a certain musical

[3]This is not to say that no differences exist between musicians and nonmusicians but that these differences remain tiny in light of the considerable difference in the amount of explicit training that exists between both groups.

[4]For instance, tempo rubato (local speeding-up and slowing-down in a performance) is often used in classical music (e.g., Hudson, 1994), whereas in jazz and rock it is more common to use timing deviations that are early or late with respect to a constant tempo (e.g., Ashley, 2002).

style. Both aspects suggest that expressive timing could serve as an indicator of temporal sensitivity to a musical idiom. Third, because expressive timing was shown not to be perceptually invariant under tempo-transformation (Honing, 2006b), as such it can function as a cue for listeners to decide whether or not a performance is tempo-transformed. Fourth, the time-scale algorithm used to make the tempo-transformed stimuli (Bonada, 2000) allows for manipulating the temporal information while maintaining musically realistic stimuli. This algorithm manipulates expressive timing while the original sound quality (e.g., attack transients and timbre) is kept perceptually invariant.[a] And finally, the task (i.e., comparing the quality of the expressive timing used in a performance) is similar to the "blindfold test" that is quite popular in media that review new CD recordings (such as magazines and radio shows). In such a test, a panel of music professionals is asked to compare and comment on the musical quality of a number of different recordings of the same music. Music lovers tend to find such a task attractive and challenging (Honing & Ladinig, 2008).

A previous study (Honing, 2006b) showed that experienced listeners are quite good at this comparison task and can distinguish between a real and a tempo-transformed performance. In the current study we investigate whether this is expert behavior or whether listeners without formal musical training, but with sufficient exposure to a certain musical idiom (e.g., jazz, rock, or classical music), can do this equally well. The expertise hypothesis predicts that experts should do better, independent of musical genre. The exposure hypothesis predicts that experienced listeners should do better, independent of the amount of musical training they have received.

3.3 Method

3.3.1 Participants

Invitations were sent to various mailing lists, online forums, and universities to reach a wide variety of respondents ($N = 208$).[b] Five gift certificates were raffled among those who responded. The respondents were between 12 and 63 years old ($M = 34$, $SD = 11.5$, $Mode = 26$) and had various musical backgrounds. Thirtyfour percent received little or no formal musical training, 29% could be considered musical experts (i.e., with more than 8 years of formal musical training and starting at a young age; Ericsson, Krampe, & Tesch-Romer, 1993), and the remaining 37% could be classified as "semimusician." Finally, 39% mentioned classical music as their main exposure category; 27% jazz; and 34% rock music.

3.3.2 Equipment

We processed the responses in an online version of the experiment using standard Web browser technologies (see Honing, 2006b, for details). The stimuli were excerpts of commercially available recordings and were converted to the MPEG-4 file format to guarantee optimal sound quality on different computer platforms and to minimize the download time.

3.3.3 Materials and Stimulus Preparation

For each of the three genres, 10 audio recordings were selected from commercially available CDs (see Table 3.1). Each *performance pair* (labeled A and B in the tables) consists of two recordings of the same composition. These were selected such that they differed between 20% and 30% in overall tempo. All sound excerpts were taken from the beginning of a recording and restricted to instrumental music only (see motivation below). For the classical and jazz genres it was relatively easy to find such recordings. However, for the rock genre this turned out to be quite a challenge, because it is less typical to have recordings of the same song in quite different tempi. However, using tools like iTunes (giving access to audio fragments of a large set of commercial recordings), we were able to find 10 recordings that were instrumental and had the desired tempo differences.

From each performance pair A and B, two *stimulus pairs* were derived (A/B' and A'/B, with prime indicating a tempo- transformed recording). This resulted in a total of 30 real and 30 tempo-transformed recordings. All 60 stimuli (constructed from the 30 recordings shown in Table 3.1) can be found in the supplemental materials.

Furthermore, the two stimulus pairs derived from each performance pair were presented to two groups of listeners. This was done to prevent the respondents from remembering characteristics of the stimuli in one pair and using them to make a response to the other pair. Group 1 (n = 101) was presented with 15 A/B' pairs, whereas Group 2 (n = 107) was presented with 15 A'/B pairs.

For each recording, the tempo was matched with a metronome to the first four bars and checked perceptually by playing it along with the music. The resulting tempo estimate (see Tempo column in the table) was used to calculate the tempo-scaling factor to make the stimulus pairs similar in tempo. The average tempo difference for each genre was about 24% (SD = 3.5%).

The tempo-transformed stimuli were made using state-of-the-art time-scale modification software (Bonada, 2000). This software can change the overall tempo of a recording while keeping the pitch and sound quality (e.g., attack transients and timbre) invariant. As such, this algorithm minimizes the effect of sound quality artifacts that could bias the results. This was confirmed by an earlier study (Experiment 2 in Honing, 2006b) in which audio experts were presented with original and tempo-transformed stimuli and asked to identify what they considered a manipulated recording. Over the whole set of 28 stimuli, audio experts did no better than chance. Although three stimuli attracted slightly more responses, these did not bias the overall results (in fact, these stimuli contained snippets of voice, such as audience coughs and humming, that apparently caused small phasing effects that some audio experts could spot when asked to do so).

In the current study we therefore decided to use the same stimuli for the classical genre as used in Honing (2006b), minus the pairs that could have biased the results. Furthermore, we made sure that the stimuli selected for the jazz and rock genres were instrumental and did not contain any voice.

Finally, there are two additional reasons why we think sound quality is less of an issue in this study. First, participants were explicitly instructed to base their judgment on the use of expressive timing, not on the sound quality of the recordings (see N.B. under *Procedure*). Second, we were interested in differences between listener groups: With each listener group listening to the same stimuli, it is unlikely that the occasional participant ignoring these instructions would influence the results.

The presentation of the stimuli was randomized within and between pairs for each participant, as was the assignment of participants to either Group 1 or Group 2. Participants could choose between a Dutch or English version of the instructions.

3.3.4 Procedure

Participants were invited to visit a Web page of the online experiment.[5] First, they were asked to test their computer and audio system with a short sound excerpt and to adjust the volume to a comfortable level. Second, they were asked to fill in a questionnaire to obtain information on their musical background, listening experience, and musical training. Participants were, for instance, requested to estimate the distribution of their average listening time over particular musical genres (classical, jazz, pop, rock, etc.) in percentages. This information was used for the measures of exposure and expertise (see *Analyses*). Finally, they were referred to a Web page containing the actual experiment. The following instructions were given:

> You will be presented with fifteen pairs of audio fragments in three different repertoires (classical, jazz, and rock): one being a real recording (by one artist), the other a manipulated tempo-transformed recording (by another artist). The tempo-transformed version was originally recorded at a different tempo, but it has been time-stretched (or time-compressed) to become close in tempo to the other performance of the pair. Your task is to decide which is which. This might be quite a challenge.

[5] The online experiment can be found in the supplemental materials.

> Please do the following: 1) Listen to a pair of audio fragments once and in their entirety (in a quiet environment without distractions or with headphones). 2) Focus on the use of expressive timing by the performer (such as note asynchrony, *tempo rubato* and articulation). 3) Then answer the questions listed next to the excerpts, namely: Which is the real (i.e., not tempo-transformed) recording, the top or the bottom excerpt? Are you sure? And, do you know this composition? 4) Please do this for all fifteen pairs of audio fragments presented below. N.B. All fragments are processed in some way, so please ignore sound quality as a possible cue for deciding which is which: Just focus on the timing of the performer(s).

The total experiment took, on average, 38 min to complete.[6]

3.3.5 Analyses

The response forms were automatically sent by e-mail to the authors and converted into a tabulated file for further analysis with POCO (Honing, 1990), music software for symbolic and numerical analyses, and SPSS (Version 11), for statistical analyses. To filter out the occasional nonserious responses, we included only entirely completed response forms and those responses that took more than 10 min for the listening part of the experiment. Dropout (percentage of visitors who did not finalize the experiment or did it too quickly, e.g., against instruction to listen completely through each audio fragment) was 36% of all respondents.

The information as collected in the questionnaire was used to assign expertise and exposure levels to each participant. With regard to expertise, participants were classified into three categories: (a) nonmusicians, who had received less than 3 years of training or no training at all; (b) expert musicians, with formal musical training longer than 8 years starting before the age of 9; and (c) semimusicians, participants that fall between these two extremes. We refer to these categories as *expertise.*

With regard to exposure, participants were also classified into three categories: classical, jazz, and rock listener. A participant was assigned to a certain listener category when he or she indicated preference for one particular genre (with a minimum difference to the other genres of 10%).[7] We refer to these categories as *exposure.*

[6] Although this might seem a long time, note that listeners could quit the experiment at any time. Furthermore, 81% indicated that they would like to participate in a future experiment. Both aspects suggest that the participant were highly motivated (cf. Honing & Ladinig, 2008).

[7] Participants who did not have a specific musical preference (not exceeding a threshold of 10% between the categories) were not considered in the ANOVAs (reducing this set to N = 131). For the other analyses all responses (N = 208) were used.

3.4 Results and Discussion

Overall the participants correctly identified the real performance 60.1% of the time ($SD = 9.7\%$). In the classical genre this was 65.3% ($SD = 21.0\%$); for jazz, 56.6% ($SD = 19.0\%$); and for rock, 58.2% ($SD = 20.2\%$). The average percentage correct for each of our nine participant groups (Exposure x Expertise) was found to be significantly above chance level (50%) using a t-test ($p > .05$). From this we can conclude that each participant group was capable of distinguishing a real recording from a manipulated, tempo-transformed performance (see Figure 3.1). As such, we were able to replicate the main result from (Honing, 2006b), which used the same task and partly the same stimuli.[8]

3.4.1 Effect of Exposure and Expertise

In this study, however, we were interested in seeing whether these judgments are the result of expert behavior or whether listeners without formal training, but with sufficient exposure to a certain musical genre, can do this equally well.

To analyze the effect of exposure and expertise on the amount of correct timing judgments of the participants, we calculated a 3 (exposure) x 3 (expertise) x 3 (genre) analysis of variance, with exposure and expertise as between-subject variables and genre, with the levels classical music, jazz music, and rock music, as a within-subject variable.

We found an effect for genre, $F(2,\ 244) = 8.19$, $p < .01$, $\eta_p^2 = .063$, showing that the overall performance, regardless of exposure or expertise, differed for each genre. Contrasts revealed that subjects performed better for classical music as compared with both jazz, $F(1,\ 122) = 15.77$, $p < .001$, and rock, $F(1,\ 122) = 8.41$, $p < .01$.[9] Furthermore, we did not find effects for either of the between-subject variables, or an interaction of these variables. However, we did find a significant three-way interaction of genre, exposure, and expertise, $F(8,\ 244) = 2.14$, $p < .05$, $\eta_p^2 = .065$.

The interactions are indicated in Figure 3.1. In the left panel of Figure 3.1 the results are grouped according to expertise levels; in the right panel the results are grouped according to listener type. The interactions are indicated by asterisks (with an arrow pointing from the cell that got significantly higher values to the cell with the lower values). The majority of the interactions between exposure and expertise occur in the jazz genre. The interactions in the right panel show that

[8]The current study shares 10 classical recordings (see Table 3.1) with the Honing (2006b) study. These 10 stimuli attracted 65.5% correct responses in the earlier study. In the current study this was 65.3%. As such, we replicate this earlier result.

[9]It is interesting to note that recent brain imaging research (Caldwell & Riby, 2007) suggests that exposure to one's favorite (preferred) music facilitates conscious cognitive processes, whereas unconscious cognitive processes might be facilitated by exposure to classical music in general, regardless of ones preferences.

expertise helps in making correct judgments, especially in the jazz genre. Also, the effect of exposure is visible in the jazz genre: Belonging to a certain listener group influences the performance, and this effect is emphasized for experts, less strong for semimusicians, and not visible for nonmusicians. The remaining interactions (not depicted in Figure 3.1) are for the participant groups experts exposed to rock and naive listeners exposed to jazz. Both performed worse in jazz than in the other genres ($p < .05$ for classical and $p < .05$ for rock for the rock listeners, $p < .05$ for rock and $p < .01$ for classical for the jazz listeners). Finally, the participant group experts exposed to classical music performed better in the classical genre than in the other genres ($p < .05$ for jazz, $p < .01$ for rock).

A possible cause of these interactions, mainly occurring in the jazz genre (see Figure 3.1, middle row), could be the special role of timing in jazz music, often intentionally deviating from standard patterns (Ashley, 2002). In a previous pilot study as well, timing in jazz turned out to be more difficult to judge, making even experts fail to recognize a tempo-transformed recording (Honing, 2007).

3.4.2 Effect of Exposure and Expertise on Sure Judgments

However, due to the relative difficulty of the task, blurring the results with responses the participants were unsure about, and that were likely a result of guessing, we decided also to consider only those judgments that the participants were sure about (referred to as correct/sure responses). For this we calculated a 3 (exposure) X 3 (expertise) X 3 (genre) analysis of variance, with exposure and expertise as between-subject variables; genre (with the levels classical music, jazz music, and rock music) as a within-subject variable; and correct/sure responses as a dependent variable.[c]

In this case the responses showed a significant interaction for genre and exposure, $F(4, 244) = 5.14$, $p < .001$, $\eta_p^2 = .078$, without apparent main effect of any variable or further interaction of these factors (see Figure 3.2).[10]

To view this interaction of genre and exposure in further detail, we first analyzed the differences in responses with regard to the different musical genres. For the classical genre, classical listeners showed higher scores ($p < .01$) than rock listeners. For the jazz repertoire, both classical and jazz listeners performed significantly better ($p < .05$ and $p < .01$, respectively) than rock listeners. For

[10] To make certain the reported result was not simply due to analyzing part of the data, we also analyzed the correct/not sure responses. For these data we found, however, neither a significant effect of the independent variables nor an interaction. As such, we can be sure that the results reported for the correct/sure responses are not an artifact of the selection made. In addition, we found the same effect of genre as we have in the genre-specific correct judgments, $F(2, 182) = 5.45$, $p < .01$, $\eta_p^2 = .057$. Contrasts revealed that subjects performed better for the classical genre than the jazz genre, $F(1, 91) = 10.74$, $p < .001$, and the rock genre, $F(1, 91) = 3.84$, $p < .05$.

the rock repertoire, there were no significant differences between listener groups.

Second, we analyzed how the responses differ within the listener groups. Classical listeners performed better on the classical repertoire than on the jazz or rock repertoire ($p < .05$ and $p < .001$, respectively) and better for the jazz genre than for the rock genre ($p < .05$). Rock listeners performed better on the rock repertoire than on the jazz repertoire ($p < .05$). No significant differences were found for the jazz listeners (although there was a tendency; see Figure 3.2).

In short, these results are in line with the idea that listeners perform best in the genre they listen to most, irrespective of expertise level, as was suggested by the exposure hypothesis.

3.5 Conclusion

This study addresses the influence of exposure versus expertise in making expressive timing judgments. It involved using an online listening experiment in which listeners with different musical preferences (exposure) and music education (expertise) were asked to compare two performances of the same composition (15 pairs, grouped in three musical genres), one of which was tempo-transformed (manipulating the expressive timing). An earlier study (Honing, 2006b) showed that expert listeners perform significantly above chance in such a comparison task. Surprisingly, the current study reveals that these judgments are not primarily influenced by expertise level (e.g., years of formal training) but mainly by exposure to a certain musical idiom. The interplay of familiarity with a particular genre (exposure) and the level of formal musical training (expertise) had a significant effect on discriminating a real from a manipulated performance. In addition, taking into account confidence, exposure positively influences the performance in a listeners preferred genre. In short, performance is not simply a result of formal musical training, but is enhanced, and for the confident responses even solely influenced, by listening to one's preferred music.

These results are in line with what has been found in the pitch domain (Bigand & Poulin-Charronnat, 2006; Tillman, Bharucha, & Bigand, 2000). These studies found responses of musically untrained listeners to be highly correlated with those of musically trained listeners, suggesting a musical capacity for melody and harmony judgments that is acquired through mere exposure to music, without the help of explicit training. Although not all listeners might be able to identify, label, or name explicitly what they perceive (Honing, 2006a; Schellenberg, 2006), most listeners seem to have a shared capability to distinguish between quite subtle musical nuances in a musical task (e.g., making judgments on expressive timing in the current study), a capability that is normally attributed to musical experts only.

Furthermore, these results are in line with Dalla Bella and Peretz (2005), who found that a sensitivity to Western musical styles is influenced by, but not conditional on, formal musical training, also showing an effect of both expertise and exposure.

In conclusion, the current study provides evidence in the temporal domain for the idea that some musical capabilities are acquired through exposure to music, and that these abilities are more likely enhanced by active listening (exposure) than by formal musical training (expertise).

Acknowledgements

This research was realized in the context of the EmCAP (Emergent Cognition through Active Perception) project funded by the European Commission (FP6-IST, contract 013123) and a grant of the Dutch Science Foundation (NWO) to Henkjan Honing. Thanks to Jordi Bonada (Music Technology Group, Universitat Pompeu Fabra) for time-scaling the audio fragments used; Glenn Schellenberg (Department of Psychology, University of Toronto) for his advice on the questionnaire and analyses; Rainer Alexandrowicz (University of Klagenfurt) for statistical advice; Bas de Haas, Niels Molenaar, Maria Beatriz Ramos, and Leigh M. Smith for their help in selecting and preparing the stimuli used; and all beta testers of the University of Amsterdam and the University of Utrecht for their time and their suggestions improving the Internet version of the experiment.

Notes

[a]Bonada's time-scale technique uses an advanced algorithm to change the tempo of a recording without altering the pitch, timbre or sound quality. The novel part is that the time-scale is only applied to non-transient elements of a signal. The transient parts are left intact (i.e., are not time-scaled), and are translated into new positions. *Transient* refers to a short and sudden change in the sound signal (e.g., the attack), whereas *non-transient* refers to the parts of the sound signal that have stabilised (e.g., the sustain).

[b]Some clarification about the number of participants seems to be appropriate.

We received 325 responses initially. We considered 208 responses of the 325 as valid (meaning a drop-out of 36%) after some filtering for 'non-serious' responses (described under 3.3.5 *Analyses*, first paragraph).

We used those 208 for our descriptive analyses (percentage correct for all pieces and for the pieces of the three different genres separately).

For our inferential statistics (ANOVAs) we used only those subjects from the 208 that clearly showed a genre preference (since this was part of the hypothesis we wanted to test), thus reducing the sample to 131 subjects. In each of the nine cells (Expertise X Exposure) were at least nine subjects. In the case of a truly equal distribution of subjects across the nine cells, 14.5 subjects would have been in each cell.

[c]For the analysis in the previous section (regardless if the participant was sure or not about the response) we had 1965 responses (131 participants judged 15 pieces each). After exclusion of the responses participants were not confident about, 882 responses remained in the analysis.

3.6 Figures and tables

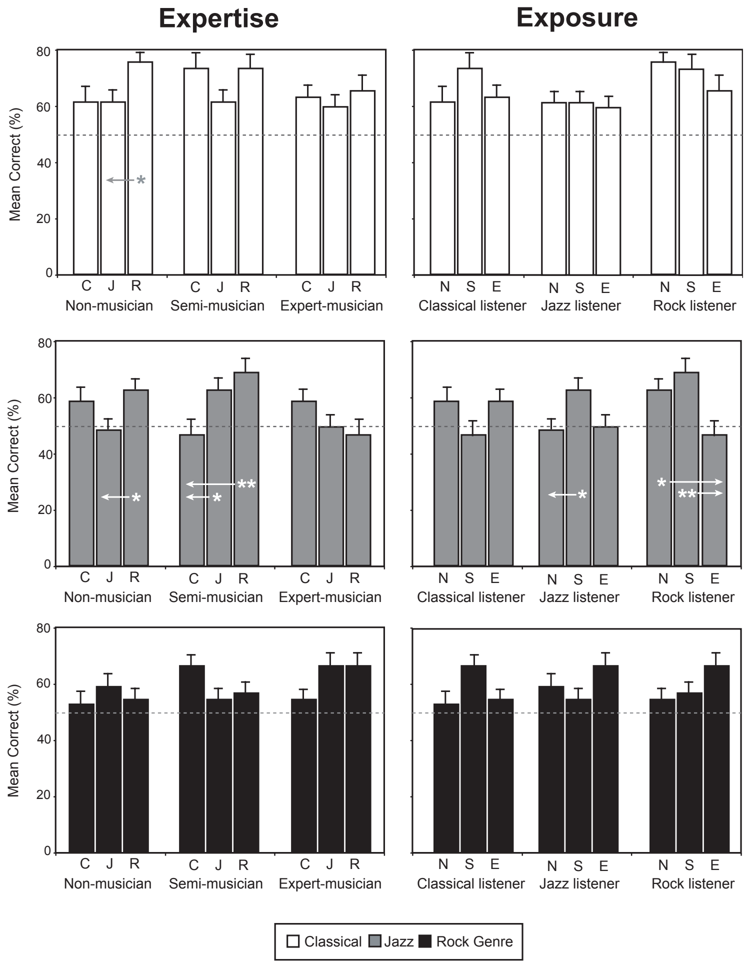

Figure 3.1: The effect of expertise and exposure on correct judgments. The panels show the mean percentage correct responses for the classical genre (top), jazz genre (middle), and rock genre (bottom). The left column shows the results grouped according to expertise levels (expertise); the right column shows the results grouped by listener type (exposure). The dotted line indicates chance level (50% correct). Asterisks mark a significant difference from the bar pointed at (*$p < .05$, **$p < .01$); error bars indicate standard error. C = classical listener, J = jazz listener, R = rock listener, N = nonmusician, S = semimusician, E = expert musician.

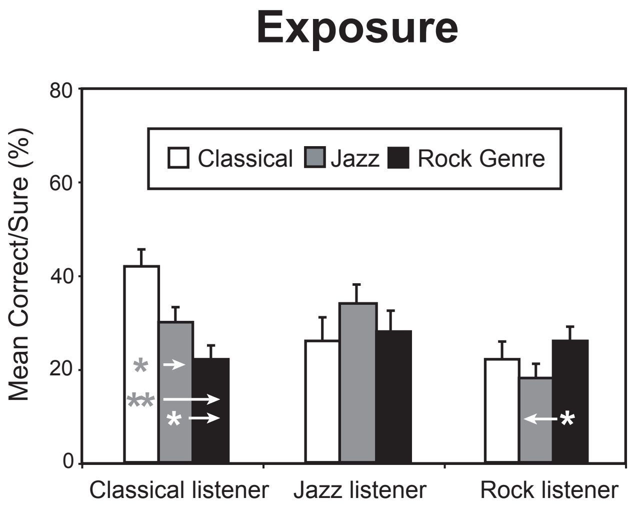

Figure 3.2: The effect of exposure on "correct/sure" judgments. Asterisks mark a significant difference from the bar pointed at (*$p < .05$, **$p < .01$); errors bars indicate standard error.

Table 3.1: Recordings used in the experiment[a]

Pair	Composition	Musician	Record label	Recording date	Tempo (BPM)
1A	J. S. Bach, English Suite No. 4, BWV 809, Allemande	Glenn Gould	Sony, SK 87766, 2001	1974/76	87
1B	J. S. Bach, English Suite No. 4, BWV 809, Allemande	Sviatoslav Richter	Delos, GH 5601, 2004	1991	70
2A	L. v. Beethoven, Piano Sonata No. 14, Op. 17, no. 2. Allegretto	Arthur Rubinstein	RCA, 09026-63056-2, 1999	1976	56
2B	L. v. Beethoven, Piano Sonata No. 14, Op. 17, no. 2. Allegretto	Vladimir Ashkenazy	Decca, 452 982-2, 1997	before 1997	75
3A	F. Chopin, Grande Valse Brillante, op. 18	Claudio Arrau	Philips, 468 391-2, 2001	1979	70
3B	F. Chopin, Grande Valse Brillante, op. 18	Vladimir Ashkenazy	Decca, 417 798-2, 1990	1983/85	88
4A	J. S. Bach, WTC II, BWV 880, Fugue 11	Glenn Gould	Sony, SX4K 60150, 1997	1969	135
4B	J. S. Bach, WTC II, BWV 880, Fugue 11	Rosalyn Tureck	BBC, BBCL 4116-2, 2002	1976	102
5A	R. Schumann, Kinderszenen, Träumerei	Vladimir Horowitz	DGG, 474 370-2, 1991	1985/89	87
5B	R. Schumann, Kinderszenen, Träumerei	Claudio Arrau	Philips, 468 391-2, 2001	1974	70
6A	Au Privave	Phil Woods	Jazz classics 1036867, 2000	1957	113
6B	Au Privave	Wes Montgomery	Riverside, 4408, 1993	195963	90
7A	Blue in Green	Bill Evans	OJC, B000000Y59, 1991	1959	67
7B	Blue in Green	Miles Davis	Sony, 64935, 1997	1959	55
8A	Dolphin Dance	Ahmad Jamal	MCA Records, IMP 12262, 1997	1970	153
8B	Dolphin Dance	Herbie Hancock	Blue Note, 7243 4 95331 2 7, 1999	1959	120
9A	Caravan	Duke Ellington	Membran Music Ltd, 222427-444, 2005	1945	114
9B	Caravan	Duke Ellington	EMI, 7243-8-29964-2-2, 1994	1962	96
10A	All the things you are	Bert van de Brink	Challenge records 70062, 1999	1999	108
10B	All the things you are	Keith Jarrett	ECM records 847135, 2000	1989	140
11A	In a Gadda da Vida	Iron Butterfly	Elektra/WEA, B0000032YA, 1993	1968	115
11B	In a Gadda da Vida	Slayer	Def Jam, B0000024K5, 1995	1989	140
12A	Killing Floor	Jimi Hendrix	Warner Bros/WEA, B000008GHU, 1990	1969	137
12B	Killing Floor	The Jimi Hendrix Experience	Rhino/WEA, B000008IKZ, 1992	1967	156
13A	Muscle Museum (Version 1)	Muse	FAAB Records, FAAB-0012-1, 2002/3	2001	161
13B	Muscle Museum (Version 3)	Muse	FAAB Records, FAAB-0012-1, 2002/3	2001	138
14A	Stairway to Heaven	Dread Zeppelin	Capitol, B000000QG4, 1991	1991	86
14B	Stairway to Heaven	Stanley Jordon	Blue Note Records, B000002UZ8, 1991	1990	66
15A	Now I Wanna Be Your Dog	The Stooges	Elektra/WEA, B0009SOFGI, 2005	1969	123
15B	Now I Wanna Be Your Dog	Iggy Pop	Other Peoples Music, B000003TWS, 1997	1979	155

[a] BPM = beats per minute; Op. = opus; BWV = Bach-Werke-Verzeichnis; WTC = Das Wohl Temperierte Clavier; BBC = British Broadcasting Corporation; RCA = Radio Corporation of America; DGG = Deutsche Grammophon Gesellschaft. SK, SX4K, GH, IMP and BBCL are parts of record label identification; OJC = Original Jazz Classics; MCA = Music Corporation of America; EMI = Electric and Musical Industries Ltd.; WEA = Warner-Elektra-Atlantic records; FAAB = Free As A Boot records.

Chapter 4

Rhythmic complexity and metric salience

Ladinig, O. & Honing, H. (under revision). Complexity judgments as a measure of event salience in musical rhythms.[1]

Abstract

This study investigates potential differences between musicians and non-musicians in their perception of meter. Listeners with a variety of musical backgrounds were asked to judge the complexity of rhythms with 4/4 time signature in a Web-based perception experiment ($N = 101$). The complexity judgments were used to derive salience values for each metric position in the rhythms, for each listener group. The judgments of the two groups were quite similar regarding the influence of the levels of metrical (hierarchical) processing. Differences between the two groups were found regarding the additional influence of the absolute position of an event in a bar (serial position effect). Listeners in both groups perceived a rhythm as more complex when syncopation occurred on an early beat of a bar than when syncopation occurred on the last beat (primacy effect). For non-musicians only, this effect could be observed on the subbeat level as well, and furthermore, a rise in salience for events at the end of a bar was found (recency effect). We propose two variants of a model of syncopation perception, one for musicians and one for non-musicians.

[1]The experiment can be found at http://cf.hum.uva.nl/mmm/exp4/

4.1 Introduction

Listener's expectations influence their perception, and in the case of musical rhythm, expectations exist about *when* an event will occur. The expectations are not the same for every event, and some events will be more and others less expected. Some events are expected very strongly, and this expectation is seen as being the basis for *beat induction* - a process in which a regular isochronous pattern (the beat) is activated internally while listening to music (for a recent overview see Patel, 2008). The *beat* (also termed *pulse* or *tactus*) is essential for time-keeping in music performance and affects the processing, coding, and appreciation of temporal patterns. Beats are positions in a rhythm that often coincide with spontaneous rhythmic behavior, like clapping hands or stomping while dancing (London, 2004; Parncutt, 1994), and there is a preference for beats to occur at intervals of about 600 msec. The induced beat underlies the perception of tempo and is the basis of temporal coding in music. Furthermore, it determines the relative importance of notes in the melodic and harmonic structure of music (Desain & Honing, 1999). Events between beats are subordinate to them, and are perceived as the weak events of a rhythm, in this paper referred to as *subbeats*. In most common Western rhythms, subbeats divide inter-beat intervals into parts whose durations form simple ratios such as 1:1, 2:1, or 3:1.

When at least two levels of metric structure are active during perception one speaks of *metric processing* (London, 2004; Yeston, 1976). The *event salience*, or sometimes termed metric salience, of a position within a pattern refers to the structural level the position is assigned to, which is an indicator of its importance relative to other positions within a certain metrical unit (e.g., bar). In general, events in positions that are perceived as salient are memorized and recalled easier, attract primary attention, are more expected to occur, and, when they are absent, lead to the impression of rhythmic complexity (Fitch & Rosenfeld, 2007; Pressing, 2002). Different theoretical models of meter perception make alternative predictions about the structure and the depth of the metric hierarchy of a rhythmic pattern.

4.1.1 Musicians vs non-musicians

There exist several conflicting theories about the influence of formal musical training on the perception of metric structure. Palmer and Krumhansl (1990) and Jongsma et al. (2004) reported differences between musicians and non-musicians. Palmer and Krumhansl analyzed goodness-of-fit judgments for single events presented in 16 positions within a 4/4 metric context; Jongsma et al. collected ERP as well as goodness-of-fit data for single events presented in seven positions within a duple and a triple metrical context. Musical training seemed to enhance depth of processing, allowing for the perception of more than two metrical levels at the same time. Palmer and Krumhansl found periodicities in the responses of non-musicians only for those positions that constitute the beat level, whereas musicians showed periodicities in their responses on lower metrical levels as well, displayed in a hierarchical structure of the positions between two beats. Results of Jongsma and colleagues are in line with those findings, but further suggest that non-musicians process temporal patterns in a more serial (as opposed to hierarchical) fashion, with a higher expectation for events to occur at the beginning of a bar.

Recently, several studies have indicated that non-musicians are more musically competent than previously thought. For example, Bigand and Poulin-Charronnat (2006), and Honing and Ladinig (2009) found evidence that if tasks and modes of responding do not require specialized training, differences between musicians and non-musicians tend to disappear.

4.1.2 Rhythmic complexity and syncopation

Several researchers have attempted to define and formalize rhythmic complexity (Essens, 1995; Pressing, 2002; Shmulevich & Povel, 2000; for an overview see Streich, 2007), but there has been little empirical validation of their models and little agreement regarding definitions of crucial concepts. In this study, perceived rhythmic complexity is thought of as being approximated by the concept of *syncopation*. Syncopation is the music-theoretical term for a moment in the music where there is a strong metric expectation that is not confirmed with a note onset. Some authors refer to this as a *loud rest* (London, 1993). A formalization of syncopation was proposed by Longuet-Higgins and Lee (1984), and is referred to as the *L-model* here. It recursively breaks down a rhythmic pattern of specific length into equal subparts, and assigns to every event a weight relating to its metrical level, assuming a metric hierarchy of maximal depth (see description of model A in the following section). For example, for a typical bar in Western music, with a 4/4 time signature and the smallest note being a 16th note, this would imply five distinct levels of event salience (e.g., Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1984). The L-model assumes that syncopation occurs if a rest or a tied note is at a higher metrical level than the immediately preceding sounding note, with the strength of the syncopation being the difference between the metrical levels of the note and the rest.

There exist two data-sets of complexity judgments in the literature. Shmulevich and Povel (2000) collected judgments of musicians, whereas Essens (1995) collected judgments of both musicians and non-musicians. The L-model accounted fairly well for the Shmulevich and Povel data, as shown by Smith and Honing (2006). For the data collected by Essens, such correlations with model predictions have not yet been reported.

The current paper first describes testable hypotheses derived from four models that differ in their assumptions regarding the salience values they assign to each metric position of a bar. Subsequently we describe test results that enabled us to derive empirically based event salience values for the average listener as well as for musicians and non-musicians separately, consisting of a metrical component and of a new component reflecting the serial position of events. We obtained these data by collecting complexity judgments about regular and syncopated rhythms.

By substituting our empirically derived salience values for the salience values assigned by the standard L-model, we generated two variants of the L-model, one for musicians and one for non-musicians. Since the new salience values are based on judgments of complexity rather than of syncopation (unlike the L-model), the resulting model variants may be suitable for specifying the complexity of a rhythm, which is here seen as superordinate to the syncopatedness of a rhythm.

4.2 Theoretical models

In this section we present four theoretical models that will enable us to construct hypotheses. The models, some of which are derived from the literature, vary in their degree of explicitness regarding the level of formalization. We also describe an empirical method for testing the relevant hypotheses. A visual representation of the four models can be found in Figure 4.1. The dotted lines represent the specified metrical levels; the dashed line indicates which events lie above or below the tactus level.

4.2.1 Model A

Model A is the L-Model, explained in the previous section. This model assumes that listeners impose as many metrical levels as possible on a rhythm. Empirical evidence for this model as representing the event salience values perceived by musicians comes from Palmer and Krumhansl (1990) and Jongsma et al. (2004). Additionally, Palmer and Krumhansl calculated frequency distributions of event onsets from a corpus of notated Western classical music, and the high correlation of those values with the event salience values perceived by musicians supports the model.

4.2.2 Model B

The same two studies (Jongsma et al., 2004; Palmer & Krumhansl, 1990), which support model A with data gathered from musicians, suggest a limited model of metric structure and event salience for non-musicians, which we call model B. As in model A, the values of event salience differ between the beat and the subbeat level, and also above the beat level (i.e., it assumes perception of a most important beat, the downbeat), but all events below the tactus belong to the same metric level and consequently have the same event salience.

4.2.3 Model C

Another model suggesting limited perception of metrical levels compared to model A is model C. This model again reflects the differentiation between beats and subbeats, but neglects the hierarchical structure of beats (no most salient beat, i.e., downbeat). Events below the beat level, however, are structured hierarchically and derived by recursive subdivision as in model A. A related formalization has recently been suggested by Gomez, Melvin, Rapaport, and Toussaint (2005), but so far has not been empirically validated.

4.2.4 Model D

A fourth model is introduced for the sake of completeness. It depicts a possible representation of the most basic metrical structure (Yeston, 1976), which contains only two different metric levels to which events are assigned: the beat and the subbeat level.

4.3 Hypotheses

We restricted this study to duple meter and a constant tempo (600 ms inter-beat interval), and kept the number of notes constant. We used rhythms commonly expressed in a 4/4 time signature, with 16 equally spaced positions of possible event onsets. We will refer to positions 1, 5, 9, and 13 as *beats*, and to all remaining positions as *subbeats*. All subbeats between two beats are considered as belonging to the same *subbeat cluster*.

To evaluate the four models of event salience, we tested the following hypotheses regarding perceived event salience:

Beat differentiation hypothesis: This hypothesis (models A and B) predicts differences in perceived event salience among the events that constitute the beat, showing a weak-strong-weak pattern following an initial downbeat. The corresponding null hypothesis (models C and D) predicts no differences in salience judgments given to beat events.

Subbeat differentiation hypothesis: This hypothesis (models A and C) predicts differences in perceived event salience among the events in each subbeat cluster, showing a weak-strong-weak pattern. The corresponding null hypothesis (models B and D) predicts no differences in the salience judgments for subbeats within a cluster.

Subbeat cluster differentiation hypothesis: This hypothesis is not based on any of the introduced models, but derives from an empirical finding by Jongsma et al. (2004), which predicts differences in salience judgments due to the position of the subbeat cluster within the bar (serial position effect). Events at the beginning of a bar may be perceived as more salient than events in the remainder of a bar. The corresponding null hypothesis (models A to D) predicts no such differences.

Beat/Subbeat relation hypothesis: This hypothesis (all models) predicts differences in perceived event salience between the beat and the subbeat level, with the beat positions receiving higher salience than the subbeat positions. The corresponding null hypothesis (not expressed in any of the introduced models) predicts no differences, and respective results would not only converse models of meter induction, but also models of beat induction.

Expertise hypothesis: Musicians are predicted to have an elaborate metrical hierarchy (model A), leading to differentiation of beats as well as subbeats. Non-musicians are predicted to show a less developed metrical structure in their salience judgments (models B, C, or D).

4.4 Methods

The purpose of the experiment was to collect relative complexity judgments about regular and syncopated rhythms. We used an online Web-based setup (see Honing & Ladinig, 2008, for a discussion of the pros and cons of this relatively novel method).

4.4.1 Participants

Invitations were sent to various mailing lists, online forums, and universities, to reach a wide variety of respondents. From the 200 initial respondents, we excluded 29% because they did not finish the experiment or did it too quickly. The remaining participants ($N = 142$) were between 17 and 63 years old (*Mode* = 20, $M = 32.7$, $SD = 11.73$) and had various musical backgrounds, ranging from no musical training up to 30 years of training. After excluding participants who could not clearly be classified as either being a musician or a non-musician (see below for the criteria), 101 participants remained in the sample, which were between 17 and 63 years old (*Modes* = 20 and 30, $M = 34.2$, $SD = 12.25$) and had a range of years of musical training from zero up to 30 years.

4.4.2 Equipment

Rhythmic stimuli were constructed using custom software and converted to MPEG-4 file format to guarantee consistent sound quality on different computer platforms and to minimize download time. The sounds were drum samples ("bongos") taken from the EZdrummer EZX Latin Percussion sample set ("Toontrack").

4.4.3 Stimuli

Sixteen rhythms, either syncopated or regular according to the definition by Longuet-Higgins and Lee (1984), were constructed (S01 - S16) and combined into seven stimulus sets, each consisting of two to four stimuli (see Figure 4.2). Small sets of stimuli were used for two reasons: First, to get clear indications of the differences in perceived complexity for a stimulus relative to certain other stimuli, as opposed to using judgments relative to the whole range of rhythms tested. And second, to employ ranking scales, as opposed to rating scales, with the former being less prone to ceiling or floor effects. The inter-onset interval (IOI) of consecutive 16th notes was 125 ms. The first position in each rhythm was marked with a louder sound to prevent listeners from perceiving it as an upbeat. Each rhythm was repeated four times without a break. Two different drum sounds (high and low congas) were used in alternation for the repetitions.

Stimulus sets 1-4 tested the structure of event salience on the subbeat level, according to the subbeat differentiation hypothesis, for subbeat clusters 1, 2, 3, and 4, respectively. Each set contained three rhythms. They had events on every beat and on two of the three subbeats within one inter-beat interval. Listeners had to compare the three stimuli within each set with regard to their perceived complexity.

Stimulus set 5 tested whether or not there are differences in event salience on the beat level, according to the beat differentiation hypothesis. Three stimuli were constructed that had events in only every other metrical position (i.e., a beat with simple subdivisions). One of the three beat events following the initial downbeat was omitted. Listeners were asked to compare the rhythms according to their perceived complexity. Stimulus set 6 was intended to shed light on whether the serial position of an invariant subbeat cluster within the rhythm affected perceived complexity, according to the subbeat cluster differentiation hypothesis. Listeners compared four stimuli, in which the same rhythmic pattern constructed of subbeats within one cluster (second and third subbeats only) occurred after beat 1, 2, 3, and 4, respectively.

Finally, stimulus set 7 provided a direct comparison of syncopation at the beat and subbeat levels. Both patterns in this set had events in the 1st, 2nd and 4th beat positions, and in the second and third subbeat position of subbeat cluster 3. The difference was that one pattern had an event on the third beat, and none on the first position of subbeat cluster 3 (subbeat syncopation), and the other pattern had no event on the third beat, but one on the first position of subbeat cluster 3 (beat syncopation).

4.4.4 Procedure

Participants were invited to visit a Web page of the experiment. They were instructed by a short screen-cast, showing examples of the experiment while the instructions were narrated, with an option to access written instructions as well. The instructions were as follows:

> In this experiment we are interested in your judgments on rhythmic complexity. We will present you seven boxes containing 2 to 4 rhythms each, and we ask you to make a judgment on the complexity of the rhythms in relation to the other rhythms within the same box (referred to as 'comparisons').
> Rhythmic complexity can be understood as a feeling of rhythmical tension, the violation of your expectation, a deviation of a regular rhythmic pattern, or non-predictability of events.
> For each of the seven sets of comparisons we ask you to listen through the whole sound samples, and, according to your perception, either 1) mark all rhythms in a box to be of equal complexity, or 2) rate their complexity on a 2 to 4 point scale (depending on the number of rhythms) where a low number indicates low complexity and a high number high complexity.
> Every rhythm is repeated four times with the percussion sounds varying for every repetition. All rhythms are played in the same tempo. You can listen to the rhythms as often as you like before making the judgment.
> N.B. There is no right or wrong answer; we are simply interested in your subjective, personal judgments.

The participants made their complexity judgments on a ranking scale that had as many increments as there where rhythms to compare in a set. The sets as well as the rhythms within each set were shown on the screen in random order. At the end of the test we asked for information about musical experience and age. We left some space for comments and feedback. The whole task typically took about 15 minutes to complete. We recorded the total time from the moment the subject started the experiment until the response form was sent, to ensure that the subject listened to all stimuli.

4.5 Data analysis

The responses were tabulated for further analysis with POCO (Honing, 1990), music software for symbolic and numerical analyses, and SPSS (Version 11) for statistical analyses.

4.5.1 Grouping by musical experience

We constructed two categories, musicians and non-musicians, and assigned participants to either of those groups. The category of musicians ($N = 57$) consisted of subjects that had between eight and 30 years of musical training (M = 15.9 years) that had started when they were between three and eight years old (M = 6.5 years). The non-musicians (N = 44) had either no formal musical training at all or had started after the age of eight (M = 17.2 years) and received training for a maximum of four years (M = 2.6). The remaining participants were excluded from the analyses.

4.6 Results

Statistical results regarding the difference between musicians and non-musicians are tested with the Wilcoxon Mann-Whitney test. Statistical results regarding the beat differentiation hypothesis, the four subbeat differentiation hypotheses, the subbeat cluster differentiation hypothesis, and the beat/subbeat relation hypothesis were tested with Friedman ANOVAs. Significant results were further analysed by obtaining pairwise comparisons using Wilcoxon tests for signed ranks with Bonferroni corrections. Mean values for all stimuli are reported in Table 4.1.

4.6.1 Differences according to musical expertise

For each stimulus, differences in responses between musicians and non-musicians were tested. Each reported Mann-Whitney U-value is based on N = 57 for musicians and N = 44 for non-musicians. Judgments differed significantly for S04 in stimulus set 2 (U = 1012.5, $p < .05$) and S10 in stimulus set 4 (U = 919, $p < .01$), and for S01 (U = 986, $p < .05$), S04 (U = 861.5, $p < .01$), and S10 (U = 854.5, $p < .001$), when presented in stimulus set 6.

4.6.2 Beat differentiation hypothesis

To test the beat differentiation hypothesis, judgments given to the stimuli of set 5 were compared. The hypothesis predicts a weak-strong-weak pattern of the three beats, with differences between the second and the third and the third and the fourth beat, but no differences between the second and the fourth beat. In other words, S14 was predicted to be judged as more complex than S13 and S15. The null hypothesis predicts no differences in judgments regarding the three stimuli. For both musicians and non-musicians, significant differences were found between the second and fourth beat, and the third and the fourth beat, but not between the second and the third beat of a bar, indicating a strong-strong-weak pattern. Thus the beat discrimination hypothesis was not confirmed, but the judgments also were not equal.

4.6.3 Subbeat differentiation hypothesis

The subbeat differentiation hypothesis suggests a weak-strong-weak pattern for each subbeat cluster individually. That is, in stimulus sets 1-4, each central stimulus was expected to be judged as more complex than the two corresponding outer stimuli. For musicians, this pattern was confirmed for all subbeat clusters. For non-musicians, this pattern was only shown for subbeat clusters one, three, and four. For subbeat cluster two, a strong-weak-weak pattern was found.

4.6.4 Subbeat cluster differentiation hypothesis

The subbeat cluster differentiation hypothesis makes predictions about a serial position effect among the subbeat clusters, with stimuli having omissions at the beginning of a bar (S01 in the context of set 6) receiving higher complexity judgments than stimuli with omissions later in the bar (S04, S07, S10 in the context of set 6). For musicians, the responses did not show any serial position effect. Complexity judgments were equally high for each subbeat cluster. For non-musicians, the responses among the subbeat clusters displayed significantly lower judgments to subbeat cluster three when compared to judgments to subbeat clusters two and four. No differences were found between subbeat cluster one and any other subbeat cluster.

4.6.5 Beat/Subbeat relation hypothesis

The beat/subbeat relation hypothesis suggests that an omission in a beat position would lead to higher complexity judgments than an omission in a subbeat position. in stimulus set 7, S16 was predicted to be judged as more complex than S07. For both musicians and non-musicians, this hypothesis was confirmed.

4.6.6 Conversion of complexity judgments into values of event salience

Participants had judged perceived complexity of rhythmic stimuli relative to one, two, or three other stimuli within the same stimulus set (see Figure 4.2). Those rankings were used for the statistical hypothesis testing reported above. To use the data as event salience values for each position in a rhythm (i.e., as values for variants of the L-model), conversions are necessary, which we will illustrate here with the non-musicians' data. The rhythms in each set can be regarded as differing in the position in which an event is omitted. Consequently, complexity judgments about a rhythmic stimulus are seen here as related to the event salience of the position of the omission. The average judgments of complexity for stimuli 01-12 in the context of stimulus sets 1-4 were taken directly as salience values for each subbeat position in a bar (see Figure 4.3, Step 1). The judgments given to stimuli in set 6 (where the same subbeat pattern occurred in different positions between beats) were added to each subbeat of the subbeat cluster represented in the stimulus (see Figure 4.3, Steps 2 and 3). They were treated as weights of each subbeat cluster within the whole measure, but leave the internal structure of each subbeat cluster intact. The resulting values were rescaled to values between 0 and 1. This was done because the judgments of stimulus set 7 indicated that participants perceived a violation of regularity on the beat level as more complex than a violation on the subbeat level. To account for this, the lowest beat position values had to be made higher than the highest subbeat position value (see Figure 4.3, Step 3). The average judgments to stimuli 13-15 were taken directly as salience values for each beat position (see Figure 4.3, Step 4). In a last step, subbeats and beats were combined (see Figure 4.3, Step 5).

4.6.7 Schematization of event salience

The derived salience values were expressed schematically as hierarchical structures, like the models shown in Figure 4.1, with each visible difference representing a significant difference in the data. None of the four proposed theoretical models of event salience could be fully validated, especially since a serial position effect occurred. Therefore we constructed two new models, one for musicians, and one for non-musicians. Figure 4.4 represents the metric processing of the different listener types, and so only the values from the judgments of stimulus sets 1-5 are considered. If significant differences exist, these are represented as differences in length of vertical lines among the subbeats belonging to one subbeat cluster, and among the beat positions. In Figure 4.5, judgments of stimulus set 6 are plotted, which represent serial processing. Figure 4.6, finally, is a combination of Figures 4.4 and 4.5. All subbeats from the subbeat-cluster that received a higher rating will be higher in salience than any subbeat from the subbeat-cluster that received a lower rating. Subbeat-cluster values that are not significantly different from either of the enclosing subbeat clusters are assigned values that overlap with the values of both enclosing subbeat clusters. In other words, if the lines of two subbeat clusters have any overlapping values, no significant difference was found between those two subbeat clusters.

4.7 Discussion and Conclusion

Contrary to what has been found in some previous studies (Jongsma et al., 2004; Palmer & Krumhansl, 1990), we found musicians and non-musicians to behave similarly in terms of hierarchical processing (see section 'Metric processing' below), but substantially different regarding the serial position of events (see section 'Serial processing' below), as assessed in terms of event salience estimates derived from complexity judgments. By using musically plausible stimulus patterns rather than probe-tones, and small sets of rhythms to compare, we gave non-musicians a chance to respond in a more natural setting. Skills that are typically very developed in musicians, like the precise subdivision of silent intervals, that can lead to good performance in temporal probe-tone tasks, were not required or in any way helpful in the current experiment.

4.7.1 Metric processing

With the exception of the second subbeat cluster in non-musicians, both musicians and non-musicians discriminated between subbeats in a hierarchical weak - strong - weak fashion. Why non-musicians differ from musicians (as well as from their own responses to the other subbeat clusters) in subbeat cluster 2 is not clear. Both listener groups showed some metric structuring on the beat level, although not in line with the predictions of any of the models we considered. While the two last beats, positions nine and thirteen, showed the expected strong-weak pattern, the second beat, position five, had a higher salience than expected, and thus has to be considered as a strong beat as well. Regardless of the differences from the proposed models, both musicians and non-musicians showed the same strong - strong - weak pattern on the beat level. We assume that a primacy effect comes into play here, which makes it more important for a rhythmic pattern to have events on earlier beats of a bar than on later beats, in order to establish a framework for meter. We consider the results for the beat level as consistent with hierarchical processing of the beat level, since the distribution shows significant differences regarding beat position four, and thus the distribution of beat saliencies is clearly not flat (excluding the first beat, which was strongest by definition).

4.7.2 Serial processing

Concerning the variation between subbeat clusters, we found effects for non-musicians, suggesting declining salience later in the bar compared to the beginning of the bar (*primacy effect*), and again a strong rise in salience at the end of the bar for non-musicians only (*recency effect*). Thus events at the beginning and the end of a pattern seemed more important than events in the middle. No serial position effect was found for musicians on the subbeat level; however, the fact that on the beat level they showed a deviation from a pure hierarchical structure at the beginning of a bar suggests a primacy effect for that listener group as well. It would be interesting to explore whether there are differences in the kinds of music musicians and non-musicians are typically exposed to. Some kinds of music, for example pop and rock, show a high tendency for upbeats, which could be a possible explanation for the high salience that the last subbeat cluster receives for non-musicians.

4.7.3 Analytic vs. heuristic listening

In order to construct a model of metrical perception that is valid for listeners from various musical backgrounds it seems to be appropriate to keep a fully metrical model for all listeners and add a serial component, which we found to be more important in listeners with less formal musical training. According to our data, the serial position effect that all listeners show is a primacy effect on the beat level. However, non-musicians in addition showed primacy and recency effects on the subbeat level as well. These results tempt us to speculate about the nature of processing of temporal information in general. Since we used rhythms that were arguably familiar to our participants, in 4/4 time-signature at a moderate tempo, probably not much cognitive effort had to be expended to relate the stimuli to known musical materials. Thereby, simple heuristic mechanisms could have come into play. Serial processing of temporal information can be seen as a quick way of grasping the structure of a rhythm, without detailed analytical, hierarchical processing. This interpretation is supported by our finding that it is non-musicians who show a stronger serial position effect than musicians.

4.8 Acknowledgments

Portions of this work were presented at the Conference of the Society for Music Perception and Cognition, July 30th - August 3rd 2007, Montreal, Canada, and at the 11th Rhythm Perception and Production Workshop, July 1st - July 5th 2007, Dublin, Ireland. Thanks to Bruno Repp for essential and thorough comments, Leigh Smith and Bas de Haas for discussions, Glenn Schellenberg for statistical advice, Andreas Witzel for technical support, and all beta testers of the Universiteit van Amsterdam and Universität Klagenfurt. This research was realized in the context of the EmCAP (Emergent Cognition through Active Perception) project funded by the European Commission (FP6-IST, contract 013123).

4.9 Figures and tables

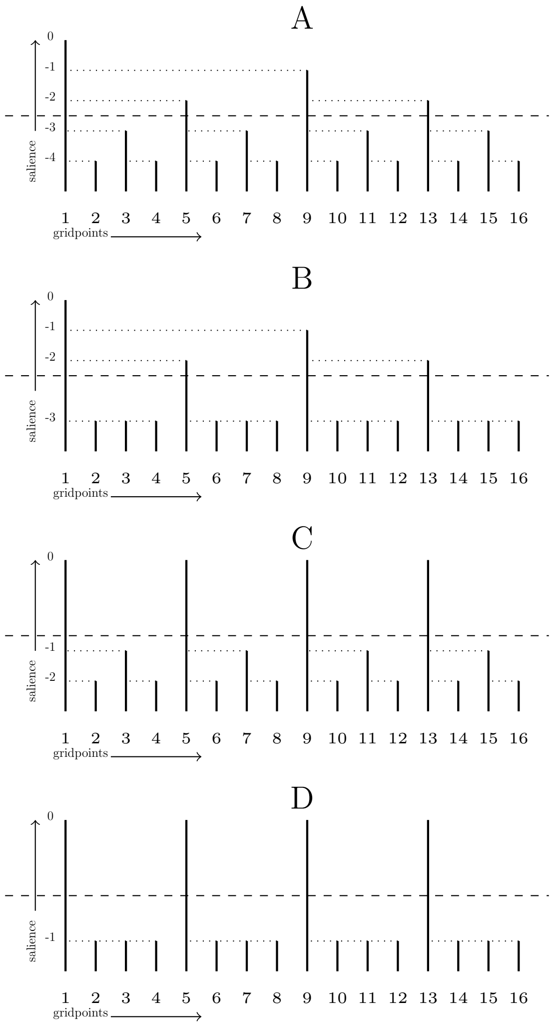

Figure 4.1: Four models of metrical structure. Predicted significant differences in event salience between grid-points within the duration of one bar (x-axis) are expressed using an ordinal scale (y-axis). The dashed lines indicate which events lie above or below the tactus

Pair	Musicians	Non-musicians
S01/S02 (SBC1)	1.28/2.12*	1.43/2.07*
S02/S03 (SBC1)	2.12/1.46*	2.07/1.68*
S01/S03 (SBC1)	1.28/1.46	1.43/1.68
S04/S05 (SBC2)	1.26/2.07*	1.52/1.84*
S05/S06 (SBC2)	2.07/1.44*	1.84/1.55
S04/S06 (SBC2)	1.26/1.44	1.52/1.55
S07/S08 (SBC3)	1.23/1.95*	1.43/1.91*
S08/S09 (SBC3)	1.95/1.39*	1.91/1.43*
S07/S09 (SBC3)	1.23/1.39	1.43/1.43
S10/S11 (SBC4)	1.30/2.26*	1.66/2.34*
S11/S12 (SBC4)	2.26/1.49*	2.34/1.43*
S10/S12 (SBC4)	1.30/1.49	1.66/1.43
S13/S14 (B)	1.72/1.65	2.05/1.75
S14/S15 (B)	1.65/1.28*	1.75/1.39*
S13/S15 (B)	1.72/1.28*	2.05/1.39*
S01/S04 (BSBC)	1.61/1.51	2.02/2.02
S01/S07 (BSBC)	1.61/1.33	2.02/1.55
S01/S10 (BSBC)	1.61/1.46	2.02/2.07
S04/S07 (BSBC)	1.51/1.33	2.02/1.55*
S04/S10 (BSBC)	1.51/1.46	2.02/1.55
S07/S10 (BSBC)	1.33/1.46	1.55/2.07*
S16/S07 (BSBR)	1.88/1.00*	1.95/1.00*

Table 4.1: Mean values for the judgments to each stimulus are given for musicians and non-musicians. Asteriks mark significant differences according to the Wilcoxon statistics, testing beat discrimination, subbeat discrimination, and subbeat cluster discrimination hypotheses. Abbreviations used for stimuli are taken from Figure 4.2.

```
Stimulus set 1
Subbeat cluster 1 (SBC1)
S01 | . | | | . . . | . . . | . . .
S02 | | . | | . . . | . . . | . . .
S03 | | | . | . . . | . . . | . . .

Stimulus set 2
Subbeat cluster 2 (SBC2)
S04 | . . . | . | | | . . . | . . .
S05 | . . . | | . | | . . . | . . .
S06 | . . . | | | . | . . . | . . .

Stimulus set 3
Subbeat cluster 3 (SBC3)
S07 | . . . | . . . | . | | | . . .
S08 | . . . | . . . | | . | | . . .
S09 | . . . | . . . | | | . | . . .

Stimulus set 4
Subbeat cluster 4 (SBC4)
S10 | . . . | . . . | . . . | . | |
S11 | . . . | . . . | . . . | | . |
S12 | . . . | . . . | . . . | | | .

Stimulus set 5
Beat (B)
S13 | . | . . . | . | . | . | . | .
S14 | . | . | . | . . . | . | . | .
S15 | . | . | . | . | . | . . . | .

Stimulus set 6
Between subbeat clusters (BSBC)
S01 | . | | | . . . | . . . | . . .
S04 | . . . | . | | | . . . | . . .
S07 | . . . | . . . | . | | | . . .
S10 | . . . | . . . | . . . | . | |

Stimulus set 7
Beat/Subbeat relation (BSBR)
S16 | . . . | . . . . | | | | . . .
S07 | . . . | . . . | . | | | . . .
```

Figure 4.2: Stimuli. The x-axis indicates the grid position, '—' marks a note/sound, '.' marks a rest/silence

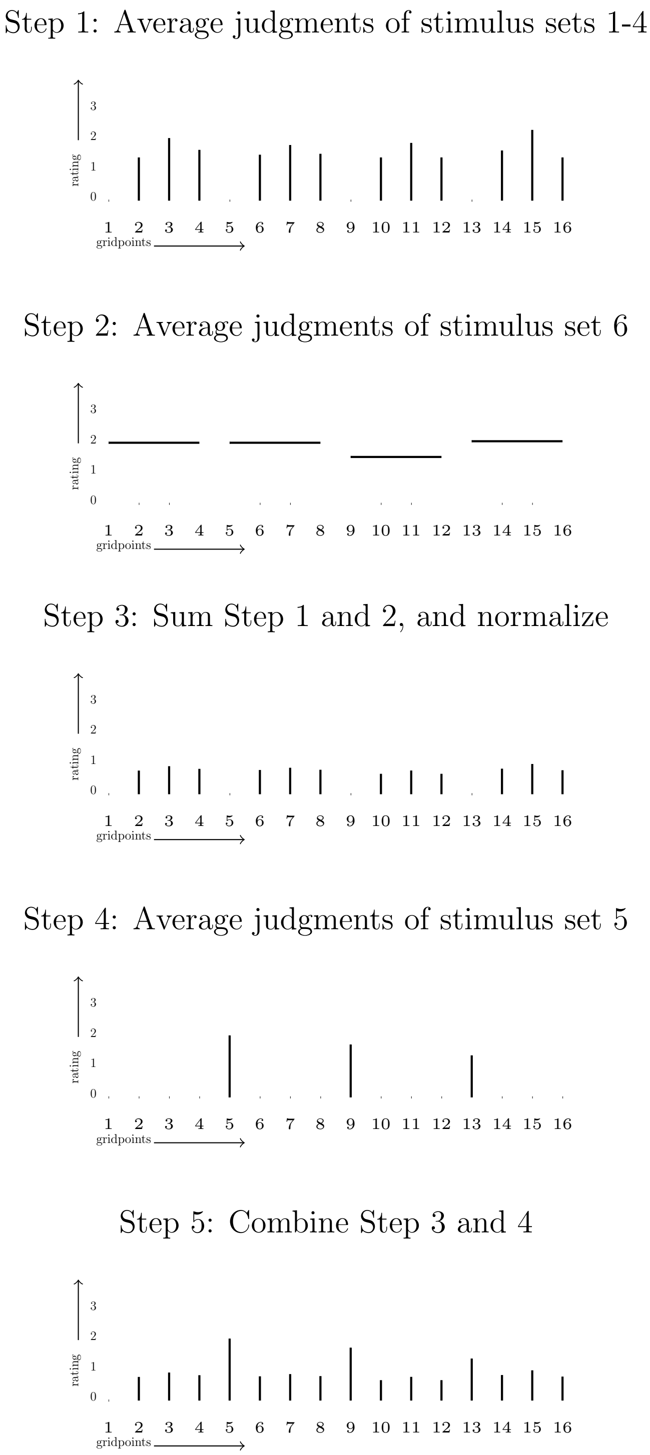

Figure 4.3: Example of conversion of complexity judgments to event salience values, using data of non-musicians

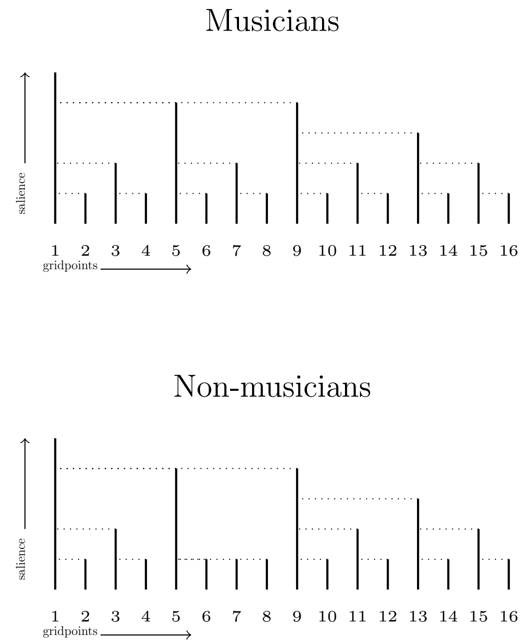

Figure 4.4: Results for beat and subbeat discrimination hypotheses for musicians and non-musicians. Significant differences in event salience between gridpoints within the duration of one bar (x-axis) are expressed using an ordinal scale (y-axis)

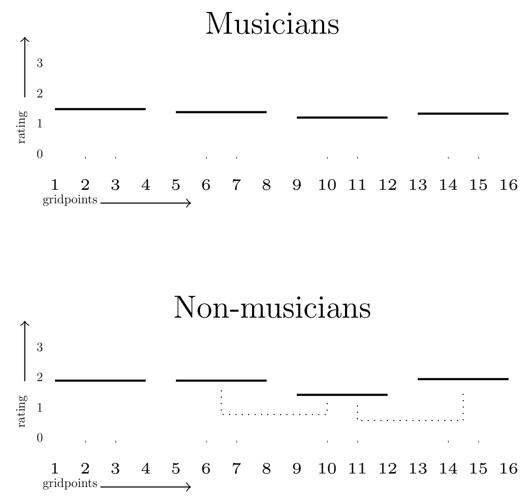

Figure 4.5: Results for serial position effect hypothesis among subbeat clusters for musicians and non-musicians. The x-axis represents gridpoints, the y-axis represents average ratings for the four subbeat-clusters (see stimulus set 6). Significant differences are indicated by dotted lines

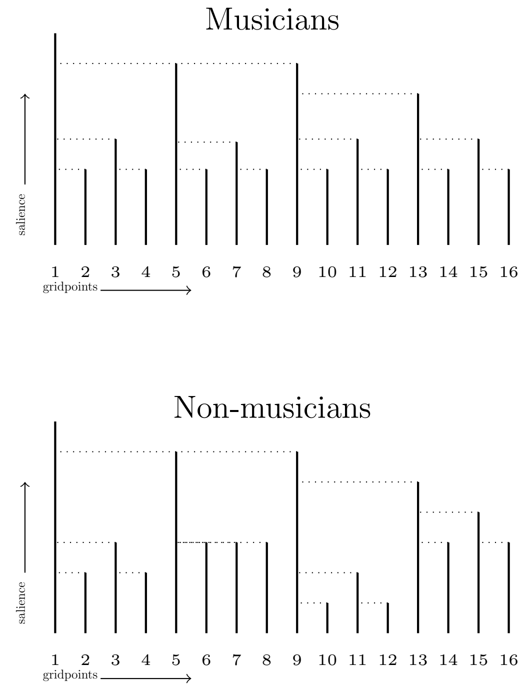

Figure 4.6: Final event salience profiles for musicians and non-musicians. The x-axis represents gridpoints, the y-axis represents event salience

Chapter 5

Meter induction in adults

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception, 26,* 377-386.

Abstract

Beat and meter induction are considered important structuring mechanisms underlying the perception of rhythm. Meter comprises two or more levels of hierarchically ordered regular beats with different periodicities. When listening to music, adult listeners weight events within a measure in a hierarchical manner. We tested if listeners without advanced music training form such hierarchical representations for a rhythmical sound sequence under different attention conditions (Attend, Unattend, and Passive). Participants detected occasional weakly and strongly syncopated rhythmic patterns within the context of a strictly metrical rhythmical sound sequence. Detection performance was better and faster when syncopation occurred in a metrically strong as compared to a metrically weaker position. Compatible electrophysiological differences (earlier and higher-amplitude MMN responses) were obtained when participants did not attend the rhythmical sound sequences. These data indicate that hierarchical representations for rhythmical sound sequences are formed preattentively in the human auditory system.

5.1 Introduction

The concepts of *beat* and *meter* are well-established terms in music production and perception (Clarke, 1999; London, 2004). Most authors agree that *beat in-*

duction, the cognitive ability that allows one to infer a regular beat (or pulse) from a musical excerpt, is universal in and unique to humans, enabling us to entrain to music, and coordinate our movements with others (Honing, 2002). Meter can be defined as being composed of at least two levels of beat with different periodicities. However, there is little agreement in the literature regarding the perceptual/cognitive reality of meter. Is meter simply a concept facilitating the structuring of written musical scores, introduced by composers and performers, or are there indeed some cognitive faculties reflected in the concept of meter? Beat induction can be considered the simplest case of meter, and refers to the subjective emphasis of certain elements of a rhythm (but also in an isochronous stream of clicks), making some elements more salient than others; the beat or *tactus* (Lerdahl & Jackendoff, 1983) is usually equally spaced in time, and is reflected in spontaneous tapping and dancing, usually with an inter-beat interval close to 600 ms (Bolton, 1894; Brochard et al., 2003; London, 2004; Yeston, 1976). Meter, seen here as a more fine-grained differentiation of the elements of a rhythm due to multiple levels of hierarchically ordered regular beats, requires the specification of a fixed entity of duration, in this case one *musical measure*. Theoretical models (Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1984) specify *metric salience*, a value assigned to each sequential position of a rhythmic sound pattern regarding to its position within that measure, by recursively breaking down a musical pattern (with an initially specified length) into subpatterns of equal length (see the top of Figure 5.1).

The number of recursive subdivisions needed to arrive at a given point (event) in a rhythmic pattern governs the salience of that point: the more subdivisions needed, the lower the salience of the point. The first position in the measure (referred to as the *downbeat*) receives the highest salience in any pattern. In other words, meter reflects the fact that different events in a musical pattern have different importance for the listener. In general, it holds that the higher the salience of an event compared to other events within the same measure, the more listeners expect it to occur. A high salience event is more important for processing the measure, as indicated for example by the fact that it gets memorized and recalled easier, and, if it is absent, the measure will be perceived as being more complex (Fitch & Rosenfeld, 2007; Pressing, 2002). Supporting this notion, Palmer and Krumhansl (1990) showed, for a corpus of Western classical music, that the average distribution of event occurrences within a measure was highly correlated with the theoretical model proposed by Lerdahl and Jackendoff (1983).

Existing theories disagree whether or not sensitivity to meter is prevalent in all listeners, and where such sensitivity, if any, would come from. Specifically, the question is, whether or not listeners form multilevel hierarchical representations for rhythmic sequences. Expectations in adult listeners with formal music training suggest that they weight events within a measure in a hierarchical manner (Jongsma et al., 2004; Palmer & Krumhansl, 1990). A study by Ladinig and Honing (under revision) shows that this holds irrespective of listener's musical expertise. Furthermore, recent evidence suggests that already at a very early age (e.g., at seven months of age), human infants are sensitive to metric violations (Hannon & Johnson, 2005). Thus it is possible that humans possess some processing predisposition to extract hierarchically structured regularities from complex patterns. Lower-level chunking processes are usually more or less automatic (i.e., they proceed even when one does not attend the given stimuli; e.g., temporal integration, see Cowan, 1984). In contrast, higher-level chunking processes typically require attention to be focused on the stimuli, because they rely on voluntary allocation of limited-capacity resources (e.g., finding sentences in continuous speech). The crucial question is whether or not the hierarchical representation characterizing meter emerges when the rhythmical sound sequence falls outside the focus of attention.

In the current study, we tested whether meter (hierarchical representation for a rhythmical sound sequence) emerges in adults with no extensive music training, and whether meter emergence is modulated by attention. To this end, reactions to meter violations were assessed using behavioral and electrophysiological measures. Reaction time (RT) and discrimination sensitivity (*d'*) measurements served to characterize active detection of meter violations, whereas event-related brain potentials (ERP) were used to assess the detection of meter violations under different task loads while the rhythmic sound sequences were not relevant to the participants' task. The mismatch negativity (MMN) ERP component (Näätänen, Gaillard, & Mäntysalo, 1978; for a recent overview see Näätänen, Paavilainen, Rinne, & Alho, 2007) can be used as a sensitive tool for determining which regular features of a sound sequence the brain has detected, because MMN is elicited by sounds violating detected auditory regularities. Furthermore, MMN is elicited even when participants perform a task that is unrelated to the test sound sequence (for a review of the effects of attention on MMN, see Sussman, 2007).

MMN has been shown to reflect violations of musical regularities and the effects of music training (for a review, see Tervaniemi & Huotilainen, 2003). For example, Trainor, McDonald, and Alain (2002) showed that participants with no formal music training detected occasional pitch interval changes within transposed melodies in the absence of focused attention. Other studies showed sensitivity to musical key (e.g., Brattico, Tervaniemi, Näätänen, & Peretz, 2006), mistuning of chords (Leino, Brattico, Tervaniemi, & Vuust, 2007), etc. Although fewer previous investigations addressed rhythm processing with the MMN method (the exceptions are Pablos Martin et al., 2007; Vuust et al., 2005), the representation of simpler temporal features has been studied in more detail. For example, it was found that occasionally shortening the inter-stimulus interval in an otherwise isochronous sequence of sounds elicits the MMN (Nordby, Roth, & Pfefferbaum, 1988). Omitting a sound from a sequence delivered at a fast presentation rate also triggers the MMN response (Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997). Regarding more complex temporal patterns, Pablos Martin et al. (2007) found faster processing of binary (e.g., 1:2) as opposed to nonbinary (e.g., 1:3) interval ratios. Finally, music training effects have been shown for both melodic (e.g., Brattico & Näätänen, 2002; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004) and rhythmic patterns (Vuust et al., 2005; Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005). The current interpretation of MMN generation suggests that this ERP component is elicited in response to deviations from expected sounds (Baldeweg, 2007; Winkler, 2007). This makes MMN especially appropriate for testing the emergence of musical meter, because it allows one to compare the strength of expectations between violations at different positions of a rhythmical pattern. The strength of expectation is a prime behavioral correlate of the hierarchical metric structure and more salient deviations trigger earlier and possibly larger-amplitude MMN responses (for a review, see Näätänen & Alho, 1997).

Based on these principles, we presented participants with sound sequences consisting of four sound patterns (Figure 5.1) having strictly metrical rhythms of the same type (Standard patterns; 90% of the patterns overall), and two patterns that were syncopated variants of the same rhythm (Deviant patterns; 10% overall).[d] One deviant violated the standard pattern at the downbeat position (strong syncopation), and the other at the second most salient position (weaker syncopation). If the brain creates a hierarchical representation for the rhythm of the sound sequences, syncopation at the downbeat is expected to elicit stronger responses from participants than syncopation at the metrically less salient position. "Stronger" response means better detection performance when syncopated patterns are designated as targets and earlier and possibly higher-amplitude MMN response when participants ignore the rhythmic sequence. If, however, the sound sequence is represented in terms of a single-level structure, then sounds in all positions are equally expected by the brain and, therefore, the responses to syncopation will not be stronger at the downbeat than in the metrically less salient position.

Effects of attention were tested at three levels: 1) meter violations are task-relevant (Behavioral Experiment); 2) meter violations are task-irrelevant: participants perform an easy concurrent task (watching a muted movie with subtitles; Electrophysiological Experiment, "Passive Condition"); and 3) meter violations are task-irrelevant: participants perform a difficult concurrent task (detecting unpredictable slight intensity changes in a noise stream; Electrophysiological Experiment, "Unattend Condition"). If forming a hierarchical representation of the rhythmical sound sequence required focused attention, then the strength of expectation should only depend on the position of the syncopation within the pattern when participants focus their attention on the sound sequence. If, however, a hierarchical representation of the rhythmical sound sequence is formed even without focused attention, then syncopation is expected to elicit a stronger response at the downbeat than in the metrically less salient position, irrespectively of the attention condition.

5.2 Method

5.2.1 Participants

Twelve healthy volunteers (seven male, $M = 22.83$, $SD = 3.93$) participated in the experiment. Participants gave informed consent after the procedures and aims of the experiments were explained to them. The study was approved by the Ethical Committee (institutional review board) of the Institute for Psychology, Hungarian Academy of Sciences. All participants had frequency thresholds not higher than 20 dB SPL in the 250-4000 Hz range and no threshold difference exceeding 10 dB between the two ears (assessed with a Mediroll, SA-5 audiometer). All participants reported to have received less than one year of music training (i.e., playing an instrument, or singing in a choir) after the obligatory music lessons in primary/secondary school in the past, and did not perform music regularly (defined as once a month) for the past two years. Each participant was tested in both experiments (behavioral and electrophysiological), which were carried out in one session on the same day. One participant's (male, age 20) data was excluded from the analyses because of measurement errors. Throughout the experiments, participants sat in a comfortable chair in the sound-attenuated experimental chamber of the Institute for Psychology, Budapest.

5.2.2 Stimuli

Six different sound patterns were constructed (see Figure 5.1), which were variants of a rhythmic rock pattern (base-pattern, S1) with eight grid points. The rhythmic patterns were presented by a typical rock-drum accompaniment using snare and bass, and with a hihat on every grid point. The base pattern and the three variants (containing omissions on the lowest metrical level) were *strictly metrical*; that is, they contained no syncopation or slurred notes throughout the pattern. Together, these four metric patterns formed the set of standard patterns (S1-S4). In order to avoid the confound of finding responses resulting from simple pattern matching, a set of sound patterns that share the characteristic of being strictly metrical and regular rhythms, instead of a single sound pattern, was employed to constitute the standard ("abstract MMN"). Two deviants were constructed by omitting events on metrically salient positions in the base-pattern, which lead to syncopated patterns: A strongly syncopated pattern was created by omitting the downbeat (D1), and a slightly weaker syncopation by omitting the second most important beat (D2). Sounds were generated using QuickTime's *drum* timbres (Apple Inc.). Sound duration was 50 ms for *hihat*, 150 ms for *snare*, and 100 ms for *bass* sounds. The interval between grid points (onset-to-onset interval) was 150 ms. Thus each pattern lasted 1200 ms, with no extra silence between patterns (i.e., they formed a continuous stream of rhythm).

5.2.3 Procedures for the Behavioral Experiment

In the behavioral experiment, we assessed the effects of different metrical positions on deviance detection by asking participants to listen to two blocks of 300 continuously presented patterns and to indicate when they felt that there was a break in the rhythm by pressing a response button placed in their dominant hand. The instructions given to participants were as follows:

> You will be presented with sequences of a continuous, regular rhythm. From time to time, the rhythm will be disrupted by some irregularity. This irregularity can be described as if the rhythm appeared to break, or stumble, or get syncopated for a moment. Please indicate by pressing the button as soon as you think such an event occurred.

Two stimulus blocks with 90% standard patterns (S1, S2, S3, and S4 with equal probabilities of 22.5% each) were presented. In one block, D1 was the deviant rhythmic pattern (10%) and in the other block, D2 was the deviant rhythmic pattern (10%). Randomization was constrained so that at least three standard patterns intervened between successive deviants and with S4 never preceding a deviant. The latter constraint was necessary to avoid concatenating two gaps, because S4 had an omission at the last grid position, whereas D1 at the first. The stimuli were presented binaurally using MATLAB via headphones (Sennheiser HD-430), 60 dB over the individual hearing threshold. The order of the two stimulus blocks (differing in the deviant pattern) was balanced across participants.

5.2.4 Data Analysis for the Behavioral Experiment

For each participant, *d'* values (a measure of discrimination sensitivity; see Macmillan & Creelman, 1991) and average reaction-times (RT) for correct responses were computed using MATLAB. The *d'* values were calculated separately from the hit rates for the D1 and D2 deviants and the overall false alarm rate. Responses given within 200-2000 ms from the target (omission) onset were regarded as hits; all other responses as false alarms. Paired two-sample t-tests were performed to compare *d'* and RT between the two deviants.

5.2.5 Procedures for the Electrophysiological Experiment

The electrophysiological experiment was conducted always before the behavioral experiment. The fixed order was necessary to avoid drawing participants' attention to the rhythmic deviations. Electrodes were removed between the two experiments, thus giving participants approximately 30 minutes of break time between the two experiments.

The rhythmic stimulus sequences were constructed from the same sound patterns as in the behavioral experiment, but they were delivered by two loudspeakers positioned 0.40 m from the side and 0.15 m behind the participants' head. Sound intensity was again 60 dB above the participant's hearing threshold. A continuous white noise with its intensity alternating between 52 and 54 dB above the participant's hearing threshold was presented concurrently with the rhythmic sound sequences. The noise stream was used to direct attention away from the rhythmic sound sequence in the Unattend condition (see below). Intensity changes occurred randomly with 1.5 - 32.0 s (M = 16.75 s) between them. The noise stream was delivered by a third loudspeaker placed directly in front of the participant at a distance of 1.35 m. During the stimulus blocks, participants also watched a self-selected muted movie with subtitles.

Two attention conditions were employed with identical auditory stimulation (rhythmic sequence and continuous noise). In the Unattend Condition, participants were asked to press a response button to the intensity changes in the noise stream. Performance in the intensity change detection task (group-average hit rate $HR = 0.78$, standard deviation $SD = 0.12$, and reaction time $RT = 1035$ ms, $SD = 77$ ms) showed that the task was difficult but possible to perform at a relatively high level. In the Passive Condition, participants were instructed to ignore all sounds (both the rhythmic sequence and the continuous noise) and to follow a muted self-selected movie. Each condition received 10 stimulus blocks of 300 continuously presented rhythmic patterns. Stimulus blocks consisted of 90% standard patterns (S1, S2, S3, and S4 with equal probabilities of 22.5%, each), 5% of the D1, and 5% of the D2 pattern. Presenting both types of deviants within the same stimulus block ensured that they appeared within exactly the same context and thus the deviance-related ERP responses could be compared directly. Randomization was constrained so that at least three standard patterns intervened between successive deviants and, for the same reasons as mentioned above, the S4 pattern never preceded a deviant pattern. Constructing 90% of the sequence from four different frequent patterns was necessary to avoid MMN being elicited by simple pattern deviation and thus to allow us to interpret the ERP responses specific to the D1 and D2 deviants as related to rhythm violations. Occasional changes of a single repeating pattern are known to elicit MMN even when rhythm is not violated (e.g., Winkler & Schröger, 1995). In the current design, the "standard" (the sequences made up of S1, S2, S3, and S4) is the rhythm, not any given sound pattern, and the deviants are the rhythmic violations caused by D1 and D2. In order to be able to directly compare the deviance-related responses elicited by D1 and D2, these responses were derived by separately subtracting the response elicited by the D1 and D2 patterns when they were regular (standard) within a sequence from when they were violating the rhythm of the sequence (deviant). Thus the pattern-specific responses were eliminated from the difference-waveforms, which could then be compared with each other. To this end, participants also were presented with two control stimulus blocks of 300 patterns presenting sequences composed of either the D1 or the D2 pattern alone. The responses recorded to the D1 and D2 patterns in the control stimulus blocks (i.e., when they are standard patterns) served to derive the MMN response (see the EEG data analysis section below). The order of the two attention conditions was balanced across participants. Stimulus blocks usually were separated by short 1-2 minutes breaks, with longer breaks allowing the participant to leave the experimental chamber inserted at need.

5.2.6 EEG Recording

The electroencephalogram (EEG) was recorded at the F3, Fz, F4, C3, Cz, and C4 scalp locations (according to the international 10-20 system) and the left and right mastoids (A1 and A2, respectively), with the common reference electrode attached to the tip of the nose. The ground electrode was placed on the forehead. Eye movements were monitored by recording the electrooculogram (EOG) between two electrodes placed above and below the left eye (vertical EOG) and between two electrodes placed lateral to the outer canthi on both sides (horizontal EOG). EEG was recorded with 32 bit resolution at a sampling rate of 250 Hz by a Neuroscan, NuAmps amplifier (Compumedics Neuroscan Inc.). The signals were on-line low-pass filtered at 40 Hz.

5.2.7 EEG Data analysis

EEG was filtered off-line between 0.1 and 20 Hz. For each D1 and D2 pattern (experimental and control stimulus blocks, separately), an epoch of 1200 ms duration was extracted from the continuous EEG record. The epoch started 600 ms before the onset of the deviation. Epochs with a voltage change below 0.1 μV or above 100 μV on any EEG or EOG channel within the -100 to 500 ms time window (relative to the deviation onset) were rejected from further analysis. Epochs were baseline-corrected by the average voltage of the whole analysis period and averaged separately for the two deviants and identical control patterns and in the two attention conditions. Using the whole analysis period as baseline balances possible slow shifts that may appear in the long analysis period. The mean number of artifact-free deviant trials per participant was 130.

MMN peak latencies were established as the central (Cz) negative maximum of the average deviant-minus-control difference waveform in the 100-250 ms *post deviance* time-range, separately for each participant, deviant, and condition. Peak latencies were established automatically in the target latency range. In cases where two or more negative peaks fell within the 100-250 post-deviance time-window and the amplitude difference between the peaks was small (< 0.5 μV), selection of the latency was aided by visual inspection of waveforms recorded by the C and F electrodes. The effects of attention and deviance position were analyzed by a repeated-measure analysis of variance (ANOVA) with the structure

Attention (Unattend vs. Passive) X Position (Strong vs. Weak).

MMN mean amplitudes were averaged from 60 ms time windows centered on the central (Cz) negative MMN peaks observed from the group-averaged deviant-minus-control difference waveforms, separately for the two deviants and two attention conditions. Thus MMN was derived by subtracting between responses elicited by identical sound patterns presented in different sequences (i.e., when D1 and D2 are deviants among standards and when D1 and D2 form homogenous control sequences). Responses elicited by the standard patterns were not used in the MMN measurements. This derivation of MMN prevents the emergence of confounding differences stemming from pattern-specific ERP responses. The group averaged central MMN peak latencies were: 160, 140, 196, and 176 ms from *deviation (omission) onset* for the Unattend-Strong, Passive-Strong, Unattend-Weak, and Passive-Weak deviant responses, respectively. The effects of attention, deviance position, and the scalp distribution of the MMN amplitudes were analyzed with a repeated-measure ANOVA of the structure Attention (Unattend vs. Passive) X Position (Strong vs. Weak) X Frontality (Frontal vs. Central electrode line) X Laterality (Left vs. Middle vs. Right). All significant effects and interactions are reported below. Greenhouse-Geisser correction of the degrees of freedom was applied where appropriate and the ϵ correction factor as well as η^2 effect size are reported.

5.3 Behavioral Data

Discrimination sensitivity was significantly higher for Strong than for Weak deviants, $t(10) = 2.80$, $p < .05$; *d'*(Strong) = 2.77, *d'*(Weak) = 2.13. There was also a tendency for faster RTs for Strong than for Weak deviants, $t(10) = 1.85$, $p < .10$; RT(Strong) = 536.69 ms, RT(Weak) = 585.68 ms.

5.3.1 Discussion of the Behavioral Data

Higher sensitivity and shorter RT's for Strong as compared to Weak deviants suggest that theoretical metrical salience affected the processing of rhythmic patterns in our participants when they attended the stimulus sequence.

5.4 Electrophysiological Data

The D1 and D2 patterns elicited a fronto-centrally more negative response between 100 and 250 ms from the onset of the omissions when the patterns violated the rhythmic context set up by the frequent standard patterns (S1-S4) than when the same patterns were presented alone in the homogeneous control stimulus blocks (Figure 5.2). The difference between the ERP responses elicited by the deviant and the identical control stimuli can be identified as an MMN response (cf. below). Significantly shorter MMN peak latencies (measured from the onset of deviation; see Figures 5.2 and 5.3) were obtained for Strong as compared to Weak deviants, $F(1, 10) = 20.69$, $p < .01, \eta^2 = 0.67$ (average peak latencies: Passive[Strong] = 145.45 ms, Passive[Weak] = 165.45 ms, Unattend[Strong] = 149.09 ms, and Unattend[Weak] = 190.18 ms). The ANOVA of MMN amplitudes (see Figures 5.2 and 5.3, and Table 5.1 for mean MMN amplitudes) yielded main effects of Position, $F(1, 10) = 5.62$, $p < .05$, $\eta^2 = 0.36$, Frontality, $F(1, 10) = 10.56$, $p < .01, \eta^2 = 0.51$, and Laterality, $F(2, 20) = 13.86$, $p < .001$, $\epsilon = 0.83$, $\eta^2 = 0.58$. Strong deviants elicited higher-amplitude MMN responses as compared to Weak deviants. MMN was larger over central than frontal electrodes and over midline than lateral electrodes. There was also a significant interaction between Attention and Frontality, $F(1, 10) = 35.24$, $p < .001$, $\eta^2 = 0.78$, stemming from lower frontal MMN amplitudes in the Passive condition than in any other combination of these two factors (Tukey HSD posthoc test with $df = 10$, $p < .001$ for all of the referred comparisons). This result rules out the possibility that the deviant-minus-control difference waveform would contain significant contribution from the N2b ERP component. This is because N2b is elicited only when participants actively detect a stimulus (Novak, Ritter, Vaughan, & Wiznitzer, 1990). Furthermore, the ERP difference cannot reflect difference between two N1 components, because it is elicited by sound omissions, which do not elicit the N1 component. Very importantly, the Attention factor did not significantly interact with the Position factor for either peak latencies or MMN amplitudes. This means that Strong deviants elicited earlier and higher-amplitude MMN responses than Weak deviants irrespective of the attention conditions.[e]

5.4.1 Discussion of the Electrophysiological Data

MMN responses were elicited by deviations in both metrical positions and in both attention conditions. This suggests that rhythmic violations are detected even when attention is not focused on the sound sequence. Furthermore, Strong deviants elicited a stronger (earlier and higher-amplitude) response than Weak ones. This result corroborates the behavioral data in suggesting that metric salience affected the detection of rhythm violations. Stronger MMN responses are usually recorded to perceptually more salient deviations (Näätänen & Alho, 1997). Since the amount of raw acoustic deviation did not differ between the two deviant positions, larger perceived deviations suggest sharper (more precise) memory representations for metrically salient elements of rhythmic patterns (a similar effect on the sharpness of the memory representations underlying MMN has been demonstrated by masking studies; see Winkler, Reinikainen, & Näätänen, 1993). Modulation of the memory representations by metric salience strongly argues for the conclusion that the brain formed hierarchical representations for the rhythmic stimulus sequences.

The only effect of attention was lower frontal MMN amplitudes in the Passive compared with the Unattend condition. This effect was not significantly different between MMNs elicited by Strong and Weak deviants. Rather, it probably reflects differences in the general activity of the frontal cortex in the two attention conditions (e.g., difference in the arousal level or between processing simple sound change as opposed to following a movie). Thus it appears that the processing of meter (forming hierarchical representations for rhythmical sound sequences) does not require significant amounts of limited higher-level capacities, a sign that meter may be processed at lower levels of auditory perception. The picture emerging from the electrophysiological data is that meter is extracted more or less automatically from rhythmic sequences, suggesting that it is an "intelligent" low level auditory processing capability, of which more and more are discovered by recent research (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001).

5.5 General Discussion and Conclusion

The behavioral detection of syncopated rhythms data as well as the ERPs recorded under two different attention conditions concerning meter induction were consistent in our group of participants. They were able to detect syncopated rhythms in an active behavioral task (indicated by the accuracy and speed of detection), as well as passively in the ERP experiment when they focused their attention on a task unrelated to the rhythmic sound sequences. Not only did participants distinguish syncopated patterns from strictly metrical ones, but they also showed sensitivity to the position (metric salience) or in other words, to the strength of the syncopation. This result is in full accordance with the Longuet-Higgins and Lee (1984) model, which predicts that the most salient position elicits a significantly stronger response than syncopation on any lower salient position of the rhythm. Furthermore, this result suggests that meter is not only a concept facilitating the structuring of written musical scores, but it corresponds to the structure of memory representations in the human brain.

These results suggest that beat induction, which according to Povel (1981) is an essential first step in the perception of temporal sequences, is functional both in active and passive listening situations. Furthermore, our participants clearly were sensitive to the hierarchical ordering in beat perception (as revealed by the difference in responses between D1 and D2; cf. Figure 5.3). This provides further evidence for the general perceptual/cognitive capability based interpretation of meter. While earlier research showed only a marginal sensitivity to meter in listeners with little or no formal music training (e.g., Jongsma et al., 2004; Palmer & Krumhansl, 1990), the current study demonstrated that meter is a mental representation that does not require advanced formal music training. This conclusion does not rule out the possibility that, similarly to other music-related processing capabilities, the representation of rhythmic structures can be improved by music training (see for example Zuijen et al., 2005). It remains a question for future research whether basic sensitivity for meter is a result of learning by exposure to one's musical environment (Huron, 2006), or, as hinted by the current results as well as studies showing sensitivity to meter at a very early age (Hannon & Johnson, 2005), whether it stems from a general cognitive predisposition of the human brain for breaking down complex patterns recursively into equal-sized subpatterns (Martin, 1972).

Notes

[d]The term 'strictly metrical' might be misleading, since all our stimuli (standards as well as deviants) have the same unambiguous metric framework. It would have been better to use 'non-syncopated' instead of 'strictly metrical'.

[e]An error was discovered after the manuscript was published. The figures as well as the statistics done on latency report deviant-minus-control differences, while the statistics on amplitude report deviant-minus-standard differences. To account for this, the correct statistics for amplitude on the deviant-minus-control differences are reported, and implications on the general interpretation of the results are discussed.

The ANOVA of MMN amplitudes yielded main effects of Attention, $F(1{,}10) = 5.41$, $p < .05$, $\eta^2 = 0.35$, and Frontality, $F(1,\ 10) = 12{,}48$, $p < .01$, $\eta^2 = 0.55$. MMN amplitude was larger in the Unattend than in the Passive condition, and higher in central than in frontal areas. Further, there was a significant interaction between Attention and Frontality, $F(1{,}10) = 13.29$, $p < .01$, $\eta^2 = 0.57$, stemming from higher central MMN amplitudes in the Unattend condition than in any other combination of these factors (Tukey HSD posthoc test with $df = 10$, $p < .001$ for all of the referred comparisons). We did not find a significant effect of Position for the statistics on the MMN amplitudes.

We do not think that the lack of a strong-weak amplitude effect modifies our conclusions, since the latency measurements support our hypothesis that strong deviants elicit a stronger deviation than weak deviants, and this goes in line with the results in the behavioural experiment. As Schröger and Winkler (1995) report, the amplitude is less reliable than the latency in reflecting deviations. The addition of the attention effect also does not affect our conclusions, because the effect appears on a general level and not specific for either the strong or weak deviant (i.e., we did not find an interaction between Position and Attention). Previously, MMN amplitude was found to increase with task-load (Zhang, Chen, Yuan, Zhang, & He, 2006), and it was suggested that increased task-load decreases the capacity to suppress the detection of irrelevant deviance. We suggest that our results show a similar effect of task-load, thus leading to higher MMN amplitudes in the Unattend as compared to the Passive condition.

5.6 Figures and tables

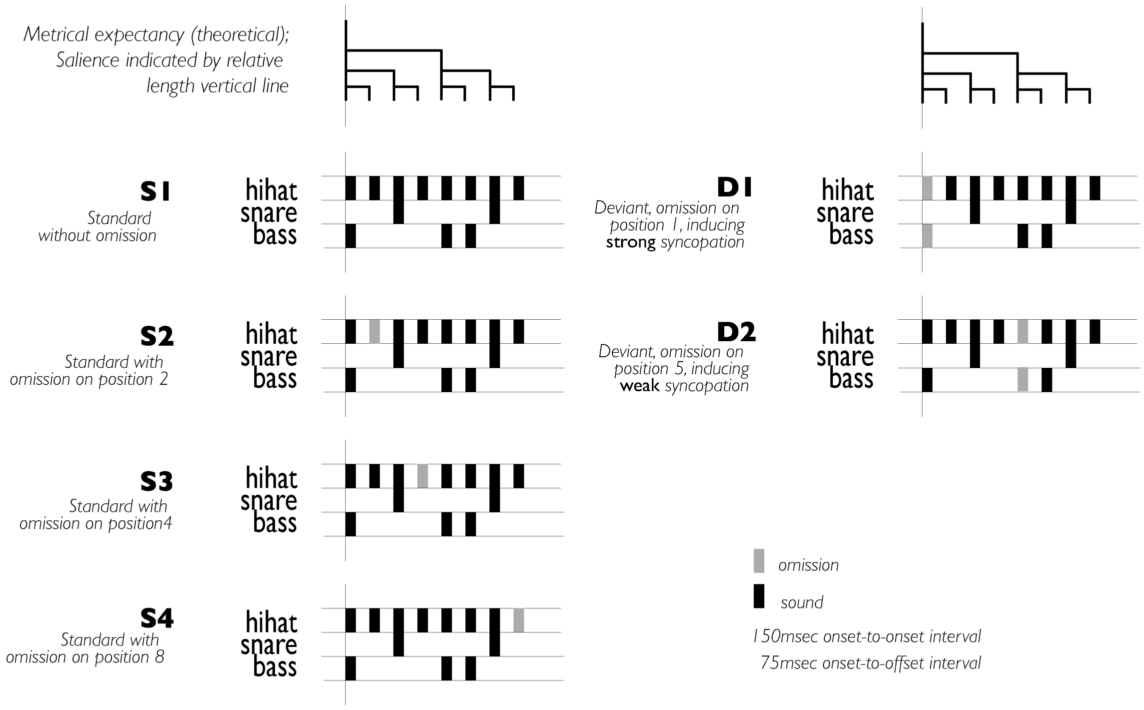

Figure 5.1: Schematic illustration of the stimuli used in the experiment. The top of the figure represents the recursive subdivision of a rhythmic pattern with eight equidistant grid points. The horizontal dimension represents the subdivisions of one musical measure; the vertical dimension represents event salience (i.e., increasing salience with longer lines)

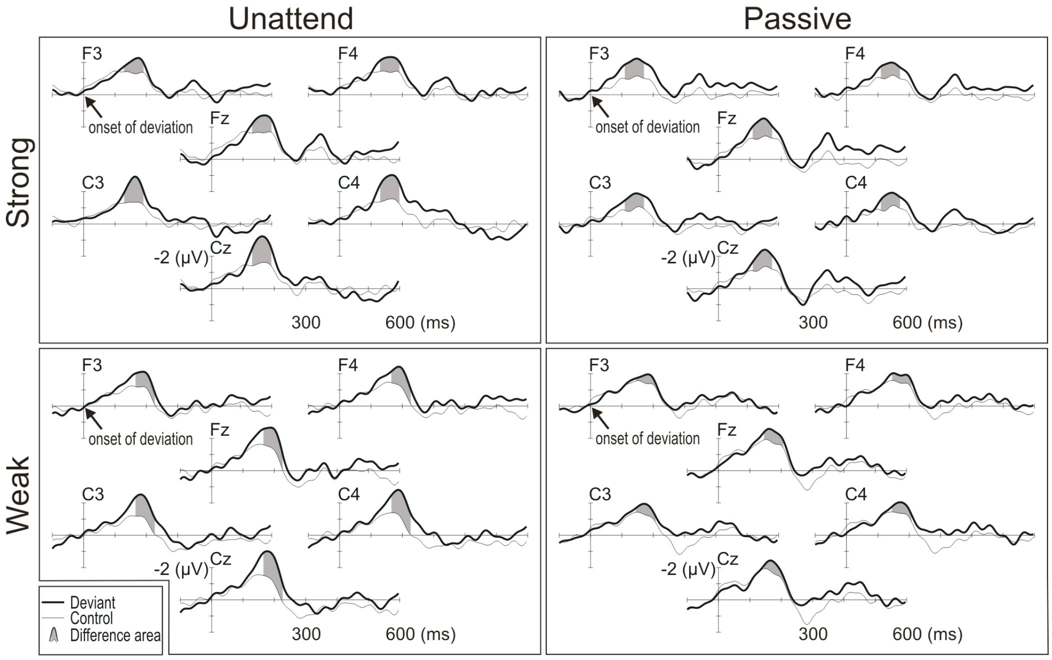

Figure 5.2: Group-averaged (n=11) ERP responses elicited by deviant patterns (experimental stimulus blocks; thick lines) and identical control patterns (control stimulus blocks; thin lines). Left: Unattend condition; right: Passive condition. Upper panels show the responses to Strong, lower panels to Weak metrical position deviants. The area between deviant and control responses within the measurement window is marked by grey shading. Responses are aligned at the onset of deviation (the time point at which the omitted sound appears in the S1 pattern).

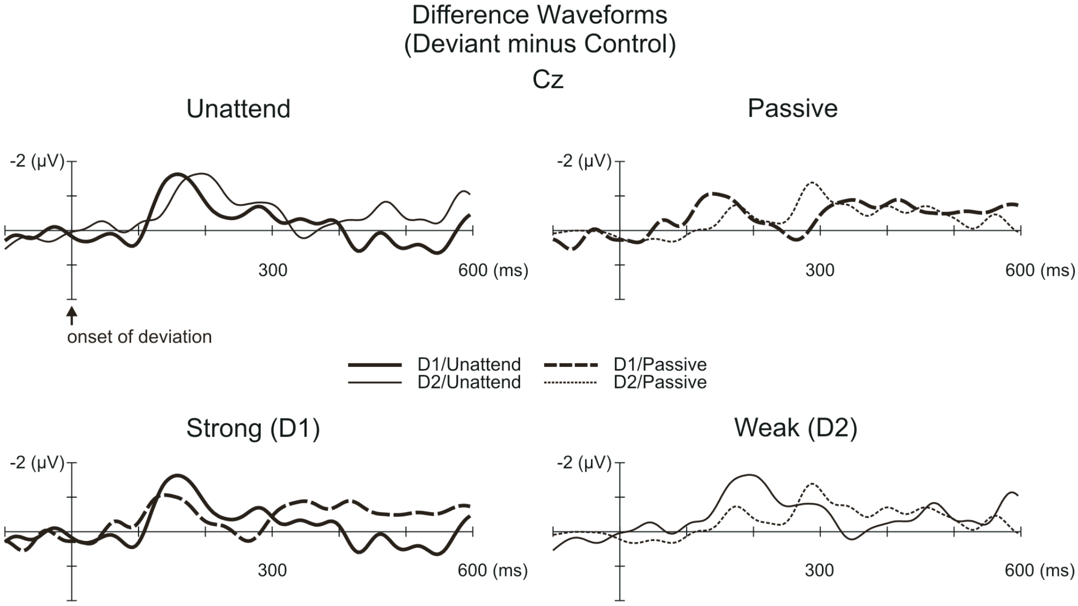

Figure 5.3: Group-averaged (n=11) deviant-minus-control difference waveforms (thick lines for Strong, thin lines for Weak deviants; continuous lines for the Unattend, dashed lines for the Passive condition). Top panels: Comparison between responses elicited by Strong and Weak deviants, separately for the Unattend (left) and Passive (right) conditions. Bottom panels: Comparison between the two attention conditions, separately for Strong (left) and Weak (right) deviants. Responses are aligned at the onset of the deviation.

Attention	Passive		Unattend	
Electrode/Position	Strong	Weak	Strong	Weak
F3	-2.23 (0.40)	-1.20 (0.29)	-2.00 (0.19)	-1.53 (0.40)
Fz	-2.62 (0.47)	-1.70 (0.38)	-2.58 (0.28)	-1.99 (0.47)
F4	-1.93 (0.41)	-1.27 (0.45)	-2.10 (0.31)	-1.68 (0.41)
C3	-2.03 (0.37)	-1.42 (0.35)	-2.72 (0.34)	-2.15 (0.37)
Cz	-2.57 (0.47)	-1.71 (0.41)	-3.29 (0.30)	-2.49 (0.47)
C4	-2.08 (0.41)	-1.48 (0.40)	-2.99 (0.35)	-2.38 (0.41)

Table 5.1: Group-Averaged MMN Amplitudes in μV with Standard Errors of the Mean (SEM) in Parentheses

Chapter 6

Beat induction in newborn infants

Winkler, I., Háden, G., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences, 106,* 2468-2471.[1]

Abstract

To shed light on how humans can learn to understand music, we need to discover what the perceptual capabilities with which infants are born. Beat induction, the detection of a regular pulse in an auditory signal, is considered a fundamental human trait that, arguably, played a decisive role in the origin of music. Theorists are divided on the issue whether this ability is innate or learned. We show that newborn infants develop expectation for the onset of rhythmic cycles (the downbeat), even when it is not marked by stress or other distinguishing spectral features. Omitting the downbeat elicits brain activity associated with violating sensory expectations. Thus, our results strongly support the view that beat perception is innate.

6.1 Introduction

Music is present in some form in all human cultures. Sensitivity to various elements of music appears quite early on in infancy (Trehub, 2003; Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005; Patel, 2008), with understanding and appreciation of music emerging later through interaction between developing perceptual capabilities and cultural influence. Whereas there is already some

[1] Supplementary material can be found at
http://www.pnas.org/content/early/2009/01/26/0809035106/suppl/DCSupplemental

information regarding spectral processing abilities of newborn infants (Novitski, Huotilainen, Tervaniemi, Näätänen, & Fellman, 2007; Winkler et al., 2003), little is known about how they process rhythm. The ability to sense beat (a regular pulse in an auditory signal; termed "tactus" in music theory; Lerdahl & Jackendoff, 1983; Honing, 2002) helps individuals to synchronize their movements with each other, such as necessary for dancing or producing music together. Although beat induction would be very difficult to assess in newborns using behavioral techniques, it is possible to measure electrical brain responses to sounds (auditory event related brain potentials, ERP), even in sleeping babies. In adults, infrequently violating some regular feature of a sound sequence evokes a discriminative brain response termed the mismatch negativity (MMN) (Kujala, Tervaniemi, & Schröger, 2007; Näätänen et al., 2001). Similar responses are elicited in newborns (Alho, Saino, Sajaniemi, Reinikainen, & Näätänen, 1990) by changes in primary sound features (e.g., the pitch of a repeating tone) and by violations of higher-order properties of the sequence, such as the direction of pitch change within tone pairs (ascending or descending) that are varying in the starting pitch (Carral et al., 2005). Newborns may even form crude sound categories while listening to a sound sequence (Kushnerenko et al., 2007): an additional discriminative ERP response is elicited when a harmonic tone is occasionally presented among noise segments or vice versa, suggesting a distinction between harmonic and complex sounds.

Neonates are also sensitive to temporal stimulus parameters (e.g., sound duration, Kushnerenko, Ceponienė, Fellman, Huotilainen, & Winkler, 2001) and to the higher-order temporal structure of a sound sequence (such as detecting periodical repetition of a sound pattern, Stefanics et al., 2007). Because the MMN is elicited by deviations from expectations (Winkler, 2007), it is especially appropriate for testing beat induction. One of the most salient perceptual effects of beat induction is a strong expectation of an event at the first position of a musical unit, i.e., the "downbeat" (Fitch & Rosenfeld, 2007). Therefore, occasionally omitting the downbeat in a sound sequence composed predominantly of strictly metrical (regular or "nonsyncopated") variants of the same rhythm should elicit discriminative ERP responses if the infants extracted the beat of the sequence.

6.2 Results and Discussion of the Neonate Experiment

We presented 14 healthy sleeping neonates with sound sequences based on a typical 2-measure rock drum accompaniment pattern (S1) composed of snare, bass and hi-hat spanning 8 equally spaced (isochronous) positions (Fig. 6.1 A and B). Four further variants of the S1 pattern (S2-S4 and D) (Fig. 6.1 C-F) were created by omitting sounds in different positions. The omissions in S2, S3, and S4 do not break the rhythm when presented in random sequences of S1-S4 linked together, because the omitted sounds are at the lowest level of the metrical hierarchy of this rhythm (Fig. 6.1 A) and, therefore, perceptually less salient (Lerdahl & Jackendoff, 1983). The 4 strictly metrical sound patterns (S1-S4; standard) made up the majority of the patterns in the sequences. Occasionally, the D pattern (Fig. 6.1 F, deviant) was delivered in which the downbeat was omitted. Adults perceive the D pattern within the context of a sequence composed of S1-S4 as if the rhythm was broken, stumbled, or became strongly syncopated for a moment (Ladinig, Honing, Háden, & Winkler, 2009). A control sequence repeating the D pattern 100% of the time was also delivered ("deviant-control").

Fig. 6.2 shows that the electrical brain responses elicited by the standard (only S2-S4; see *Methods*) and deviant-control patterns are very similar to each other, whereas the deviant stimulus response obtained in the main test sequence differs from them. The deviant minus deviant-control difference waveform has two negative waves peaking at 200 and 316 ms followed by a positive wave peaking at 428 ms. The difference between the deviant and the other two responses was significant in 40-ms-long latency ranges centered on the early negative and the late positive difference peaks (see Table 6.1 for the mean amplitudes) as shown by dependent measures ANOVAs with the factors of Stimulus (Standard vs. Deviant control vs. Deviant) X Electrode (C3 vs. Cz vs. C4). The Stimulus factor had a significant effect on both peaks (for the early negative waveform: $F[2,26] = 3.77$, $p < 0.05$, with the Greenhouse-Geisser correction factor $\epsilon = 0.85$ and the effect size $\eta^2 = 0.22$; for the positive waveform: $F[2,26] = 8.26$, $p < 0.01$, $\epsilon = 0.97$, $\eta^2 = 0.39$). No other main effects or interaction reached significance. Posthoc Tukey HSD pairwise comparisons showed significant differences between the deviant and the deviant-control responses in both latency ranges (with $df = 26$, $p < 0.05$ and 0.01 for the early negative and the late positive waveforms, respectively) and for the positive waveform, between the deviant and the standard response ($df = 26$, $p < 0.01$). No significant differences were found between the standard and the deviant control responses.

Results showed that newborn infants detected occasional omissions at the first (downbeat) position of the rhythmic pattern, but, whereas the S2S4 patterns omitted only a single sound (the hi-hat), the D pattern omitted 2 sounds (hi-hat and bass). The double omission could have been more salient than the single sound omissions, thus eliciting a response irrespective of beat induction. However, the omission in D could only be identified as a double omission if the neonate auditory system expected both the bass and the hi-hat sound at the given moment of time. Because bass and hi-hat cooccurs at 3 points in the base pattern (see Fig. 6.1 B), knowing when they should be encountered together requires the formation of a sufficiently detailed representation of the whole base pattern in the neonate brain. In contrast, beat detection requires only that the length of the full cycle and its onset are represented in the brain. It is possible that neonates form a detailed representation of the base pattern. This would allow them not only to sense the beat, but also to build a hierarchically ordered representation of the rhythm (meter induction), as was found for adults (Ladinig et al., 2009). This exciting possibility is an issue for further research.

Another alternative interpretation of the results suggests that newborn infants track the probabilities of the succession of sound events (e.g., the probability that the hi-hat and bass sound event is followed by a hi-hat sound alone). However, in this case, some of the standard patterns (e.g., S2) should also elicit a discriminative response, because the omission has a low conditional probability (e.g., the probability that the hi-hat and bass sound event is followed by an omission, as it occurs in S2, is 0.078 within the whole sequence).

Finally, it is also possible that newborn infants segregated the sounds delivered by the 3 instruments, creating separate expectations for each of them. This explanation receives support from our previous results showing that newborn infants segregate tones of widely differing pitches into separate sound streams (Winkler et al., 2003). If this was the case, omission of the bass sound could have resulted in the observed ERP differences without beat being induced. To test this alternative, we presented the test and the control sequences of the neonate experiment to adults, silencing the hi-hat and snare sounds. All stimulation parameters, including the timing of the bass sounds and the probability of omissions (separately for the test and the control sequences) were identical to the neonate experiment.

6.3 Results and Discussion of the Adult Control Experiment

Fig. 6.3 shows that the ERP responses elicited by the deviant and the control patterns are highly similar to each other. Taking the peaks where the central (C3, Cz, and C4 electrodes) deviant-minus-control difference was largest (132 and 296 ms) within the latency range in which discriminative ERP responses are found in adults, we conducted ANOVAs of similar structure as was done for the neonate measurements [dependent measures factors of Stimulus (Deviant control vs. Deviant) X Electrode (C3 vs. Cz vs. C4); standard patterns could not be used, because they contained no omission in the bass sequence]. We found no significant main effect of Stimulus or interaction between the Stimulus and the Electrode factor ($p>$ 0.2 for all tests). The only significant effect found was that of Electrode for the later latency range ($F[2,24] = 7.30$, $p< 0.01$,$\epsilon = 0.75$, $\eta^2 = 0.38$) However, because this effect does not include the Stimulus factor, it is not the sign of a response distinguishing the deviant from the control response.

Thus, in adults, omission of the position-1 bass sound does not result in the elicitation of discriminative ERP responses in the absence of the rhythmic context. This result is compatible with those of previous studies showing that stimulus omissions (without a rhythmic structure) only elicit deviance-related responses at very fast presentation rates (<170-ms onset-to-onset intervals; see Yabe et al., 1997). In our stimulus sequences, the omitted bass sound was separated by longer intervals from its neighbors. It should be noted that adult participants elicited the MMN discriminative ERP response, when they received the full stimulus sequence (all 3 instruments) as presented to newborn babies in the neonate experiment (Ladinig et al., 2009).

6.4 Discussion

These results demonstrate that violating the beat of a rhythmic sound sequence is detected by the brain of newborn infants. In support of this conclusion we showed that the sound pattern with omission at the downbeat position elicited discriminative electrical brain responses when it was delivered infrequently within the context of a strictly metrical rhythmic sequence. These responses were not elicited by the D pattern per se: When the D pattern was delivered in a repetitive sequence of its own, the brain response to it did not differ from that elicited by the standards. Neither were discriminative responses simply the result of detecting omissions in the rhythmic pattern. Omissions occurring in non-salient positions elicited no discriminative responses (see the response to the standards in Fig. 6.2). Furthermore, the discriminative ERP response elicited by the D pattern was not caused by separate representations formed for the 3 instruments: only omissions of the downbeat within the rhythmic context elicit this response.

So it appears that the capability of detecting beat in rhythmic sound sequences is already functional at birth. Several authors consider beat perception to be acquired during the first year of life (Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005; Patel, 2008), suggesting that being rocked to music by their parents is the most important factor. At the age of 7 months, infants have been shown to discriminate different rhythms (Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005). These results were attributed to sensitivity to rhythmic variability, rather than to perceptual judgments making use of induced beat. Our results show that although learning by movement is probably important, the newborn auditory system is apparently sensitive to periodicities and develops expectations about when a new cycle should start (i.e., when the downbeat should occur). Therefore, although auditory perceptual learning starts already in the womb (Visser, Mulder, & Prechtl, 1992; Huotilainen et al., 2005), our results are fully compatible with the notion that the perception of beat is innate. In the current experiment, the beat was extracted from a sequence comprised of 4 different variants of the same rhythmic structure. This shows that newborns detect regular features in the acoustic environment despite variance (Carral et al., 2005) and they possess both spectral and temporal processing prerequisites of music perception.

Many questions arise as a result of this work. Does neonate sensitivity to important musical features mean that music carries some evolutionary advantage? If so, are the processing algorithms necessary for music perception part of our genetic heritage? One should note that the auditory processing capabilities found in newborn babies are also useful in auditory communication. The ability to extract melodic contours at different levels of absolute pitch is necessary to process prosody. Sensing higher-order periodicities of sound sequences is similarly needed for adapting to different speech rhythms e.g., finding the right time to reply or interject in a conversation (Jaffe & Beebe, 2001). Temporal coordination is essential for effective communication. When it breaks down, understanding and cooperation between partners is seriously hampered. Therefore, even if beat induction is an innate capability, the origin and evolutionary role of music remains an issue for further research.

6.5 Methods

6.5.1 Neonate Experiment

Sound sequences were delivered to 14 healthy full term newborn infants of 37-40 weeks gestational age (3 female, birth weight 2650-3960 g, APGAR score 9/10) on day 2 or 3 post partum while the electroencephalogram was recorded from scalp electrodes. The study was approved by the Ethics Committee of the Semmelweis University, Budapest, Hungary. Informed consent was obtained from one or both parents. The mother of the infant was present during the recording.

The experimental session included 5 test sequences, each comprising 276 standard (S1-S4; Fig. 6.1) and 30 deviant (D) patterns and a control sequence in which the D pattern (termed deviant-control) was repeated 306 times. The control stimulus block was presented at a randomly chosen position among the test stimulus blocks. In the test sequences, S1-S4 appeared with equal probability, 22.5%, each, with the D pattern making up the remaining 10%. This is a prerequisite of the deviance detection method, which requires deviations to be infrequent within the sequence (Kujala et al., 2007). The order of the 5 patterns was pseudorandomized, enforcing at least 3 standard patterns between successive D patterns. The onset-to-onset interval between successive sounds was 150 ms with 75-ms onset-to-offset interval (75-ms sound duration). Patterns in the sequence were delivered without breaks. Loudness of the sounds was normalized so that all stimuli (including the downbeat) had the same loudness.

EEG was recorded with Ag-AgCl electrodes at locations C3, Cz, and C4 of the international 10-20 system with the common reference electrode attached to the tip of the nose. Signals were off-line filtered between 1 and 16 Hz. Epochs starting 600 ms before and ending 600 ms after the time of the omission in the sound patterns (compared with the S1 pattern) were extracted from the continuous EEG record. Epochs with the highest voltage change outside the 0.1-100

μV range on any EEG channel or on the electrooculogram (measured between electrodes placed below the left and above right eye) were discarded from the analysis. Epochs were baseline corrected by the average voltage during the entire epoch and averaged across different sleep stages, whose distribution did not differ between the test and the control stimulus blocks. Responses to the S2-S4 patterns were averaged together, aligned at the point of omission (termed "Standard"). Responses were averaged for the D pattern separately for the ones recorded in the main test and those in the control sequences (Deviant and Deviant-control responses, respectively). For assessing the elicitation of differential ERP responses, peaks observed on the group-average difference waveforms between the deviant and deviant-control responses were selected. For robust measurements, 40-ms-long windows were centered on the selected peaks. Amplitude measurements were submitted to ANOVA tests (see the structure in *Results and Discussion* of the *Neonate Experiment*). Greenhouse-Geisser correction was applied. The correction factor and the effect size (partial eta-square) is reported. Tukey HSD pairwise posthoc comparisons were used. All significant results are discussed.

6.5.2 Adult Experiment

Fourteen healthy young adults (7 female, 18-26 years of age, mean: 21.07) participated in the experiment for modest financial compensation. Participants gave informed consent after the procedures and aims of the experiments were explained to them. The study was approved by the Ethical Committee of the Institute for Psychology, Hungarian Academy of Sciences. All participants had frequency thresholds not > 20 dB SPL in the 250-4000 Hz range and no threshold difference exceeding 10 dB between the 2 ears (assessed with a Mediroll, SA-5 audiometer). Participants watched a silenced subtitled movie during the EEG recordings. One participant's data were rejected from the analyses due to excessive electrical artifacts.

All parameters of the stimulation, EEG recording, and data analysis were identical to the neonate study except that hi-hat and snare sounds were removed from the stimulus patterns without changing the timing of the remaining sounds.

Acknowledgements

We thank János Rigó for supporting the study of neonates at the First Department of Obstetrics and Gynecology, Semmelweis University, Budapest and Erika Józsa Váradiné and Zsuzsanna D'Albini for collecting the neonate and adult ERP data, respectively and to Martin Coath and Glenn Schellenberg for their comments on earlier versions of the manuscript. This work was supported by European Commission 6th Framework Program for "Information Society Technologies" (project title: EmCAP Emergent Cognition through Active Perception).

6.6 Figures and tables

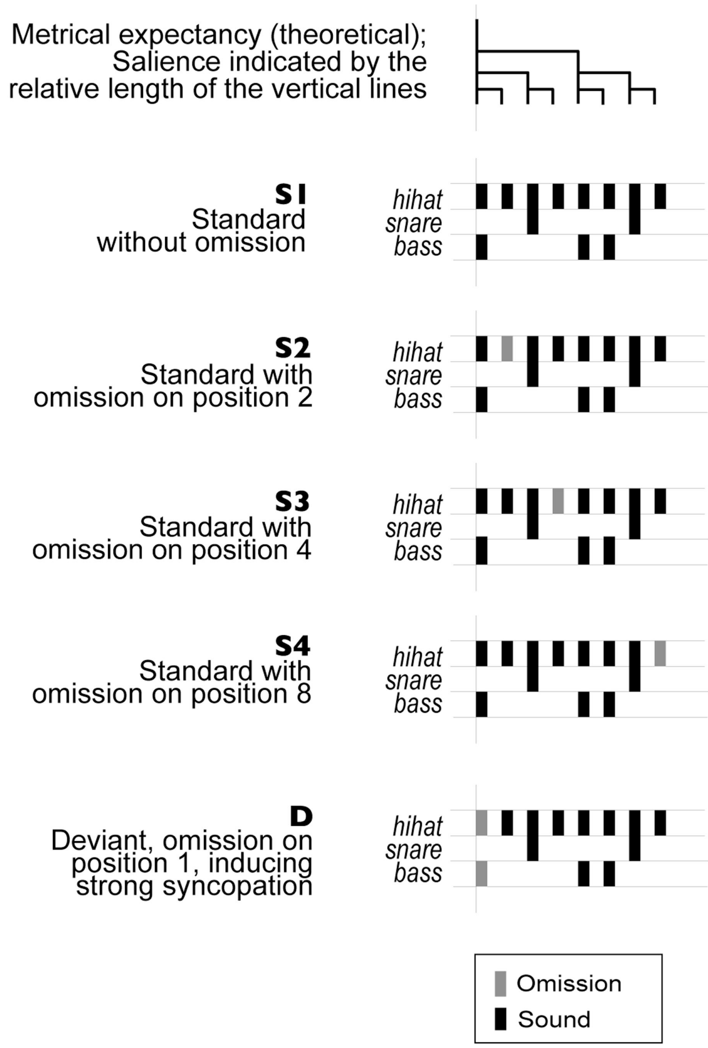

Figure 6.1: Schematic diagram of the rhythmic stimulus pattern

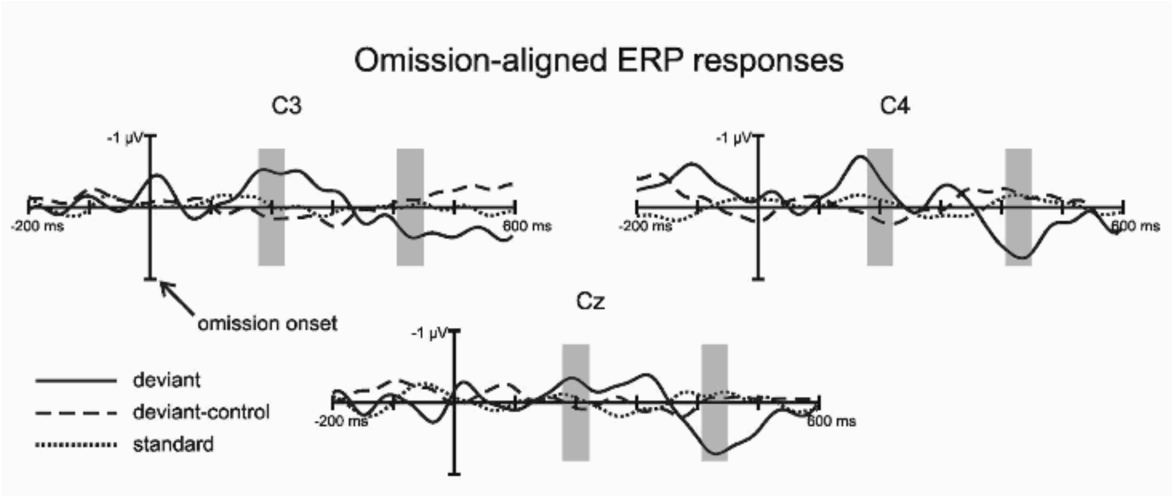

Figure 6.2: Group-averaged ($n = 14$) electrical brain responses elicited by rhythmic sound patterns in neonates. Responses to standard (average of S2, S3, and S4; dotted line), deviant (D; solid line), and deviant-control patterns (D patterns appearing in the repetitive control stimulus block; dashed line) are aligned at the onset of the omitted sound (compared with the full pattern: S1) and shown from 200 ms before to 600 ms after the omission. Gray-shaded areas mark the time ranges with signicant differences between the deviant and the other ERP responses.

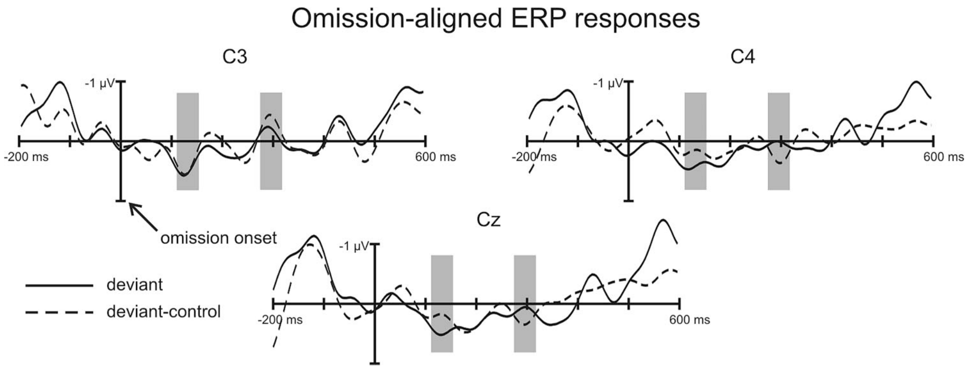

Figure 6.3: Group-averaged ($n = 13$) electrical brain responses elicited by the bass sound patterns in adults. Responses to deviant (D; solid line), and deviant-control patterns (D patterns appearing in the repetitive control stimulus block; dashed line) are aligned at the onset of the omitted bass sound (compared with the standard patterns: S1-S4) and shown from 200 ms before to 600 ms after the omission. Gray-shaded areas mark the time ranges in which amplitudes were measured.

	Amplitude μV					
	180-220 ms interval			408-448 ms interval		
Group	C3	Cz	C4	C3	Cz	C4
Deviant	-0.50 (0.19)	-0.30 (0.22)	-0.41 (0.27)	0.38 (0.17)	0.67 (0.13)	0.67 (0.27)
Deviant-control	0.14 (0.13)	0.06 (0.16)	0.18 (0.25)	-0.10 (0.13)	-0.06 (0.16)	-0.18 (0.15)
Standard	-0.03 (0.09)	-0.06 (0.12)	-0.11 (0.09)	-0.02 (0.09)	-0.12 (0.12)	-0.16 (0.13)

Table 6.1: Group-averaged (n = 14) mean ERP amplitudes

[2]SEM values are shown in parentheses

Chapter 7

Outlook

This thesis contains a collection of papers that investigate temporal expectations, by looking at their violations and considering their relations to musical expertise (rule-based learning through formal music training), exposure (implicit learning of statistical regularities), as well as innate cognitive mechanisms. The term 'expectation' implies an active role of the listener, who constantly predicts what events will happen at what time in the future. The more confident the predictions are, the more will an outcome that is different to what was predicted lead to a feeling of surprise or violation. This makes the responses to violations of expectations informative about underlying cognitive schemes that generated the expectations.

In the presented work, two kinds of expectation determined by different salience of events in rhythmic patterns were shown to be active (although this is not to say that there are not many other expectations). The first one is based on hierarchical structuring of event salience. In this regard, it could be shown that meter is induced in all listeners, regardless of the level of formal musical training. Hierarchical structuring could be found on all levels of a musical measure.

Since the present study dealt only with the most common meter in Western music, namely 4/4, further investigations will be made with other available simple meters like 2/3, 3/4, and 6/8, and the study eventually will be extended to compound or additive meters.

Furthermore, it was shown that the most fundamental instance of meter induction, namely the discrimination of the downbeat from other positions in a rhythm, was active in newborn infants, which gives rise to exciting speculations about the origins of music and, in particular, rhythm, as well as about a possible fundamental predisposition of the cognitive system to structure novel incoming information in a hierarchical way. As a logical next step, the design is currently being extended to allow probing other metric levels in newborns as well, as such levels could be found in adults with and without formal music training. To explore the nature of beat and meter even further, non-human animals could serve as potential subjects in the future.

The second type of expectation based on event salience that was considered in this thesis was the serial position effect, known mainly from memory research. It could be shown that for non-musicians, as well as for musicians (but less so), a mechanism was active that can be seen as complementary to hierarchical processing. Primacy and recency effects appeared, which led to an increased salience of events located at the beginning and at the end of a rhythmic pattern. Known from memory experiments showing that items at the beginning and at the end of sequences are recalled faster and memorized longer, it can be argued that these mechanisms can facilitate the perception and processing of rhythmic sequences as well. Since this effect has received little consideration in the music literature, it will require more systematic studies to understand the importance of serial position, and its interplay with hierarchically structured processing. The serial position effect appeared to be more pronounced in non-musicians than musicians, although there were few differences in the hierarchical processing between those groups. However, both mechanisms seem to be active even in subjects located at the ends of the expertise continuum. The serial position effect was only studied by means of listener's judgments in the course of this thesis. Further information could be gained by investigating this mechanism with listeners being in an inattentive state, by employing electrophysiological methods.

On a much smaller time-scale, support was found for the hypothesis that listeners are sensitive to deviations on a temporal micro-level, being able to distinguish tempo-transformed from non-transformed performances, by only focusing on expressive timing. This is supporting previous evidence that timing does not scale proportionally with tempo, with the new finding that also non-musicians are sensitive to distortions. A more surprising finding was that not only the level of formal music training was responsible for this sensitivity, but that exposure to a certain musical genre was giving the listener an advantage in spotting the deviations. For sensitivity to violations on a minute scale, formal musical training was even less an explanatory factor than for expectations based on event salience, and the crucial parameter was primarily the familiarity with the respective musical style.

Another issue that has received attention throughout all chapters of this thesis, and that deserves consideration in future work, is listening competence. Formal musical training alone was shown to be too crude a measure to classify listeners as competent or not regarding skills that do not involve musical performance. A first step in this direction was taken by considering exposure to certain musical styles in classifying a listener as being potentially sensitive to that particular style. Another measure that was developed, but did not become part of this thesis, is based on the breadth of knowledge a listener has regarding various musical styles. This measure asks for familiarity with (not preference for) a variety of genres and styles, intended to measure the degree of musical eclecticism.

To conclude, this thesis contributes to the growing evidence that the perceiver shapes the percept, by emphasizing the active role of the listener in rhythm perception and processing.

References

Acheson, D., & MacDonald, M. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*, 50-68.

Alho, K., Saino, K., Sajaniemi, N., Reinikainen, K., & Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology*, *77*, 151–155.

Ashley, R. (2002). Do[n't] change a hair for me: The art of jazz rubato. *Music Perception*, *19*, 311-332.

Auditory. (2007). *Online listening tests and psychoacoustic experiments with large n.* (Discussion on Auditory mailing list, see http://www.auditory.org)

Baldeweg, T. (2007). Erp repetition effects and mismatch negativity generation: A predictive coding perspective. *Journal of Psychophysiology*, *211*, 204-213.

Berlyne, D. (1970). Novelty, complexity and hedonic value. *Perception and Psycophysics*, *8*, 279-286.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? a review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*, 100-130.

Bigand, E., Tillman, B., Poulin, B., D'Adamo, D., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring. *Cognition*, *81*, B11-B20.

Birnbaum, M. (2004). Human research and data collection via the internet. *Annual Review of Psychology*, *55*, 803-832.

Blacking, J. (1974). *How musical is man?* Washington: University of Washington Press.

Bolton, T. (1894). Rhythm. *The American Journal of Psychology*, *6*, 145-238.

Bonada, J. (2000). Automatic technique in frequency domain for near-lossless time-scale modification of audio. In *Proceedings of the international computer music conference* (p. 396-399). San Francisco: Computer Music Association.

Brattico, E., & Näätänen, R. (2002). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception*, *19*, 199–222.

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*, *1117*, 162–174.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The 'ticktock' of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*, 362-366.

Caldwell, G. N., & Riby, L. M. (2007). The effects of music exposure and own genre preference on conscious and unconscious cognitive processes: A pilot ERP study. *Consciousness and Cognition*, *16*, 992–996.

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Randy McNally.

Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., Näätänen, R., & Escera, C. (2005). A kind of auditory "primitive intelligence" already present at birth. *European Journal of Neuroscience*, *21*, 3201–3204.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (second ed., pp. 473–500). San Diego, CA: Academic Press.

Dalla Bella, S., & Peretz, I. (2005). Differentiation of classical music requires little learning but rhythm. *Cognition*, *96*, B65-B78.

Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research*, *28*, 29–42.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*, 341–65.

Dienes, Z., & Longuet-Higgins, H. C. (2004). Can musical transformations be implicitly learned? *Cognitive Science*, *28*, 531-558.

Drake, C., Jones, M., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, *77*, 251-288.

Ebbinghaus, H. (1885). *Über das Gedächtnis*. Leipzig: Dunker.

Ericsson, K., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.

Essens, P. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception and Psychophysics*, *57*, 519–532.

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception*, *25*, 43-58.

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

Francès, R., Zenatti, A., & Imberty, M. (1979). Psychologie de l'art et de l'esthétique. In R. Francès (Ed.), (p. 139-193). Paris: Presses Universitaires de France.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*, 1010-1021.

Gomez, F., Melvin, A., Rapaport, D., & Toussaint, G. T. (2005). Mathematical measures of syncopation. In *Proceedings of bridges: Mathematical connections in art, music, and science* (p. 47-56). Alberta, Canada: Bridges.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, *50*, 354–377.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, *1*, 48–55.

Honing, H. (1990). Poco: An environment for analysing, modifying, and generating expression in music. In *Proceedings of the 1990 international computer music conference* (p. 364-368). San Francisco: Computer Music Association.

Honing, H. (2002). Structure and interpretation of rhythm and timing. *Dutch Journal of Music Theory*, *7*, 227-232.

Honing, H. (2006a). De analfabetische luisteraar (kan ook groot luisteren) [The iliterate listener]. *NRC Handelsblad, Opinie & Debat*, 19.

Honing, H. (2006b). Evidence for tempo-specific timing in music using a web-based experimental setup. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 780-786.

Honing, H. (2007). Is expressive timing relationally invariant under tempo transformation? *Psychology of Music*, *35*, 276-285.

Honing, H., & Ladinig, O. (2006). The effect of exposure and expertise on timing judgments: Preliminary results. In *Proceedings of the International Conference on Music Perception and Cognition* (p. 80-85). Bologna: Italy: Bononia University Press.

Honing, H., & Ladinig, O. (2008). The potential of the internet for music perception research: A comment on lab-based versus web-based studies. *Empirical Musicology Research*, *3*, 4-7.

Honing, H., & Ladinig, O. (2009). Exposure influences timing judgments in music. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 281-288.

Hudson, R. (1994). *Stolen time: The history of tempo rubato*. Oxford: Clarendon Press.

Huotilainen, M., Kujala, A., Hotakainen, M., Parkkonen, L., Taulu, S., Simola, J., et al. (2005). Short-term memory functions of the human fetus recorded with magnetoencephalography. *NeuroReport*, *16*, 81–84.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge: MIT Press.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*, 33-72.

Jaffe, J., & Beebe, B. (2001). *Rhythms of dialogue in infancy* (Vol. 66). Boston: Blackwell.

Johnson-Laird, P., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, *71*, 191-229.

Jones, M. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*, 323–335.

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

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biological Psychology*, *66*, 129–152.

Klauer, K., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, *107*, 852-884.

Krantz, J. (2000). Tell me, what did you see? the stimulus on computers. *Behavior Research Methods, Instruments, and Computers*, *32*, 221-229.

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology*, *74*, 1–19.

Kushnerenko, E., Ceponienė, R. ; Fellman, V., Huotilainen, M., & Winkler, I. (2001). Event-related potential correlates of sound duration: Similar pattern from birth to adulthood. *NeuroReport*, *12*, 3777–3781.

Kushnerenko, E., Winkler, I., Horváth, J., Näätänen, R., Pavlov, I., Fellman, V., et al. (2007). Processing acoustic change and novelty in newborn infants. *European Journal of Neuroscience*, *26*, 265–274.

Ladinig, O., & Honing, H. (under revision). Complexity judgments as a measure of event salience: Are there indeed differences between musicians and nonmusicians? *Music Perception*, *x*, x-x.

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception*, *26*, 377–386.

Large, E., & Jones, M. (1999). The dynamics of attending: how people track time-varying events. *Psychological Review*, *106*, 119–159.

Large, E., Velasco, M., & Gray, P. (2008). *Rhythmic analysis of musical interactions between bonobo and human.* Presentation given at the International Conference on Music Perception and Cognition. Sapporo, Japan.

Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, *1142*, 169-177.

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

Levitin, D. (2006). *This is your brain on music: The science of a human obsession.* New York: Penguin Books.

London, J. (1993). Loud rests and other strange metric phenomena (or, meter as heard). *Music Theory Online*, *0*.

London, J. (2004). *Hearing in time: psychological aspects of musical meter.* New York: Oxford University Press.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, *1*, 424–441.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University press.

Martin, J. (1972). Rhyhtmic (hierarchic) versus serial structure in speech and other behaviour. *Psychological Review*, *79*, 487-509.

McGraw, K., Tew, M., & Williams, J. (2000). The integrity of web-delivered experiments: Can you trust the data? *Psychological Science*, *11*, 502-506.

Mehler, J. (1999). Editorial. *Cognition*, *71*, 187–189.

Meyer, L. (1965). *Emotion and meaning in music.* Chicago, Il.: University of Chicago Press.

Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind and body.* London: Weidenfeld Nicolson.

Musch, J., & Reips, U. D. (2000). Psychological experiments on the internet. In M. Birnbaum (Ed.), (p. 61-87). San Diego, CA: Academic Press.

Näätänen, R., & Alho, K. (1997). Mismatch negativity (mmn) - the measure for central sound representation accuracy. *Auditory Neuro Ontology*, *2*, 341–353.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (mmn) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence' in the auditory cortex. *Trends in Neurosciences*, *24*, 283–288.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model.* Chicago, Il.: University of Chicago Press.

Nordby, H., Roth, W. T., & Pfefferbaum, A. (1988). Event related potentials to time deviant and pitch deviant tones. *Psychophysiology*, *25*, 249-261.

Novak, G. P., Ritter, W., Vaughan, H. G., Jr., & Wiznitzer, M. L. (1990). Differentiation of negative event related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*, *75*, 255-275.

Novitski, N., Huotilainen, M., Tervaniemi, M., Näätänen, R., & Fellman, V. (2007). Neonatal frequency discrimination in 250 4000 hz frequency range: Electrophysiological evidence. *CLinical Neurophysiology*, *118*, 412–419.

Ollen, J. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings*. Unpublished doctoral dissertation, Ohio State University.

Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (mmn). *Psychophysiology*, *38*, 359-365.

Pablos Martin, X., Deltenre, P., Hoonhorst, I., Markessis, E., Rossion, B., & Colin, C. (2007). Perceptual biases for rhythm: The mismatch negativity latency indexes the privileged status of binary vs non-binary interval ratios. *Clinical Neurophysiology*, *118*, 2709-2715.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115-138.

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

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464.

Patel, A. (2008). *Music, language, and the brain.* New York: Oxford University Press.

Patel, A., Iversen, J., Bregman, M., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*, 827-830.

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

Phillips-Silver, J., & Trainor, L. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*, 1430.

Pinker, S. (1997). *How the mind works.* New York: Norton.

Povel, D. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 3–18.

Pressing, J. (2002). *Cognitive complexity and the structure of musical patterns.* (http://psy.uq.edu.au/CogPsych/Noetica/OpenForumIssue8/Pressing.html)

Reips, U. (2002). Standards for internet-based experimenting. *Experimental Psychology*, *49*, 243-256.

Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *Journal of the Acoustical Society of America*, *88*, 622–41.

Repp, B. H. (1995). Expressive timing in Schumann's "Träumerei": An analysis of performances by graduate student pianists. *Journal of the Acoustical Society of America*, *98*, 2413–9.

Repp, B. H. (2007). Hearing a melody in different ways: Multistability of metrical interpretation, reflected in rate limits of sensorimotor synchronization. *Cognition*, *102*, 434–454.

Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, M., & Näätänen, R.

(1992). Representation of abstract stimulus attributes in human brain. *NeuroReport*, *3*, 1149–1151.

Schachner, A., Brady, T., Pepperberg, I., & Hauser, M. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, *19*, 831-836.

Schellenberg, E. (2006). The child as musician: A handbook of musical development. In G. McPherson (Ed.), (p. 111-134). Oxford: Oxford University Press.

Schröger, E., & Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on pre-attentive change detection. *Neuroscience Letters*, *193*, 185–188.

Shmulevich, I., & Povel, D. J. (2000). Measures of temporal pattern complexity. *Journal of New Music Research*, *29*, 61–69.

Sloboda, J. (1994). What makes a musician? *EGTA Guitar Journal*, *5*, 18-22.

Sloboda, J. (2000). Individual differences in musical performance. *Trends in Cognitive Sciences*, *4*, 397-403.

Smith, J. (2002). Standards for internet-based experimenting. *Experimental Psychology*, *49*, 243-256.

Smith, L., & Honing, H. (2006). *Evaluating and extending computational models of rhythmic syncopation in music.* New Orleans,LA: International Computer Music Association.

Stefanics, G., Háden, G., Huotilainen, M., Balázs, L., Sziller, I., Beke, A., et al. (2007). Auditory temporal grouping in newborn infants. *Psychophysiology*, *44*, 697–702.

Streich, S. (2007). *Music complexity: a multi-faceted description of audio content.* Unpublished doctoral dissertation.

Sussman, E. (2007). A new view on the MMN and attention debate: The role of context in processing auditory events. *Journal of Psychophysiology*, *21*, 164–175.

Tervaniemi, M., & Huotilainen, M. (2003). The promises of change-related brain potentials in cognitive neuroscience of music. *Neurosciences and Music*, *999*, 29–39.

Tillman, B., Bharucha, J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *89*, 334-368.

Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, *14*, 1–13.

Trehub, S. (2003). The developmental origins of musicality. *Nature Neuroscience*, *6*, 669-673.

Visser, G., Mulder, E., & Prechtl, H. (1992). Studies on developmental neurology in the human fetus. *Developmental Pharmacology and Therapeutics*, *18*, 175–183.

Vuust, P., Pallesen, K. J., Bailey, C., Zuijen, T. L. van, Gjedde, A., Roepstorff, A., et al. (2005). To musicians, the message is in the meter. pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, *24*, 560–564.

Winkler, I. (2007). Interpreting the mismatch negativity (MMN). *Journal of Psychophysiology*, *21*, 147–163.

Winkler, I., Kushnerenko, E., Horváth, J., Ceponienė, R. ; Fellman, V., Huotilainen, M., et al. (2003). Newborn infants can organize the auditory world. *Proceedings of the National Academy of Sciences USA*, *100*, 1182–1185.

Winkler, I., Reinikainen, K., & Näätänen, R. (1993). Event related brain potentials reflect traces of the echoic memory in humans. *Perception and Psychophysics*, *53*, 443–449.

Winkler, I., & Schröger, E. (1995). Neural representation for the temporal structure of sound patterns. *Neuroreport*, *6*, 690–694.

Wolpert, R. (2000). Attention to key in a nondirected music listening task: musicians vs. nonmusicians. *Music Perception*, *18*, 225-230.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by mmn to sound omission. *Perception and Psychophysics*, *8*, 1971–1974.

Yeston, M. (1976). *The stratification of musical rhythm.* New Haven, Conn.: Yale University Press.

Zanto, T., Snyder, J., & Large, E. (2006). Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology*, *2-3*, 221–231.

Zatorre, R. (2005). Music: The food of neuroscience? *Nature*, *434* (1312-1315).

Zhang, P., Chen, X., Yuan, P., Zhang, D., & He, S. (2006). The effect of visuospatial attentional load on the processing of irrelevant acoustic distractors. *NeuroImage*, *33*, 715–724.

Zuijen, T. L. van, Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity: A mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research*, *23*, 270–276.

Samenvatting

Dit proefschrift bevat een collectie artikelen die temporele verwachtingen bij het luisteren naar muziek onderzoeken. De term 'verwachting' (*expectation*) impliceert een actieve rol van de luisteraar: hij of zij voorspelt voortdurend welke gebeurtenissen, wanneer in de nabije toekomst zullen gebeuren. Hoe sterker deze verwachtingen zijn, hoe groter een gevoel van schending van deze verwachtingen als deze niet bevestigd worden. Als een verwachting geschonden wordt geeft dit informatie over de onderliggende cognitieve functies die deze verwachtingen genereren.

Het proefschrift onderzoekt met name wat de rol is van muzikale deskundigheid (*expertise*: expliciet leren d.m.v formele muzieklessen) en blootstelling (*exposure*: impliciet leren van statistische regelmatigheden door luisteren naar muziek) in de vorming van verwachtingen in het luisteren, maar ook in hoeverre aangeboren cognitieve mechanismen een rol spelen.

Het onderzoek laat zien dat twee typen verwachting actief zijn in het luisteren naar muzikale ritmes. De eerste is gebaseerd op een hiërarchisch gestructureerde verwachting die een accentstructuur legt op ritmische gebeurtenissen (metrische verachting). In dit verband werd aangetoond dat een *metrum* (of maatsoort) wordt geïnduceerd (opgewekt) in alle luisteraars, ongeacht het niveau van formele muzikale opleiding. Bovendien kon worden aangetoond dat het meest saillante niveau van een metrum - de 'beat' of *tactus* -, actief is in pasgeboren baby's.

Het tweede aspect van verwachting dat in dit proefschrift beschreven wordt is het *serial position effect*. Zowel musici and niet-musici hebben een sterkere verwachting voor gebeurtenissen aan het begin (*primacy* effect) en aan het einde (*recency* effect) van een metrische eenheid.

Daarnaast werd op een kleinere tijdschaal steun gevonden voor de hypothese dat luisteraars gevoelig zijn voor afwijkingen op een temporeel micro-niveau, en onderscheid kunnen maken tussen temporeel getransformeerde en niet getransformeerde muzikale uitvoeringen. Leken luisteraars blijken deze oordelen even goed te kunnen maken als muzikale experts. Niet muzikale *expertise*, maar *expo-*

sure - blootstelling aan een bepaald muzikaal genre - blijkt een rol te spelen in dit type luisteroordelen.

Abstract

This thesis contains a collection of papers that investigate temporal expectations. Responses that indicate that a stimulus is perceived as unnatural, complex, or surprising, as opposed to a stimulus being perceived as natural, simple, or unsurprising, are taken as indicators that expectations were not fulfilled but rather violated. The term 'expectation' implies an active role of the listener, who constantly predicts what events will happen at what time in the future. The more confident the predictions are, the more will an outcome that is different to what was predicted lead to the violation of an expectation. This makes the responses to violations of expectations informative about underlying cognitive schemes that generated the expectations.

Special consideration is given to musical expertise (rule-based learning through formal music training), exposure (implicit learning of statistical regularities), as well as innate cognitive mechanisms.

Two kinds of expectation determined by different salience of events in rhythmic patterns were shown to be active. The first one is based on hierarchical structuring of event salience. In this regard, it could be shown that meter is induced in all listeners, regardless of the level of formal musical training. Hierarchical structuring could be found on all levels of a musical measure. Furthermore, it was shown that the most fundamental instance of meter induction, namely the discrimination of the downbeat from other positions in a rhythm, was active in newborn infants. The second type of expectation based on event salience that was considered in this thesis was the serial position effect. It could be shown that for non-musicians, as well as for musicians (but less so), a mechanism was active that can be seen as complementary to hierarchical processing. Primacy and recency effects appeared, which led to an increased salience of events located at the beginning and at the end of a rhythmic pattern.

On a much smaller time-scale, support was found for the hypothesis that listeners are sensitive to deviations on a temporal micro-level, being able to distinguish tempo-transformed from non-transformed performances, by only focusing on expressive timing. This is supporting previous evidence that timing does not scale proportionally with tempo, with the new finding that also non-musicians are sensitive to distortions. A more surprising finding was that not only the level of formal music training was responsible for this sensitivity, but that exposure to a certain musical genre was giving the listener an advantage in detecting the deviations.

Titles in the ILLC Dissertation Series:

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability

ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations

ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs

ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software

ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance

ILLC DS-2002-02: **Marco Aiello**
Spatial Reasoning: Theory and Practice

ILLC DS-2002-03: **Yuri Engelhardt**
The Language of Graphics

ILLC DS-2002-04: **Willem Klaas van Dam**
On Quantum Computation Theory

ILLC DS-2002-05: **Rosella Gennari**
Mapping Inferences: Constraint Propagation and Diamond Satisfaction

ILLC DS-2002-06: **Ivar Vermeulen**
A Logical Approach to Competition in Industries

ILLC DS-2003-01: **Barteld Kooi**
Knowledge, chance, and change

ILLC DS-2003-02: **Elisabeth Catherine Brouwer**
Imagining Metaphors: Cognitive Representation in Interpretation and Understanding

ILLC DS-2003-03: **Juan Heguiabehere**
Building Logic Toolboxes

ILLC DS-2003-04: **Christof Monz**
From Document Retrieval to Question Answering

ILLC DS-2004-01: **Hein Philipp Röhrig**
Quantum Query Complexity and Distributed Computing

ILLC DS-2004-02: **Sebastian Brand**
Rule-based Constraint Propagation: Theory and Applications

ILLC DS-2004-03: **Boudewijn de Bruin**
Explaining Games. On the Logic of Game Theoretic Explanations

ILLC DS-2005-01: **Balder David ten Cate**
Model theory for extended modal languages

ILLC DS-2005-02: **Willem-Jan van Hoeve**
Operations Research Techniques in Constraint Programming

ILLC DS-2005-03: **Rosja Mastop**
What can you do? Imperative mood in Semantic Theory

ILLC DS-2005-04: **Anna Pilatova**
A User's Guide to Proper names: Their Pragmatics and Semanics

ILLC DS-2005-05: **Sieuwert van Otterloo**
A Strategic Analysis of Multi-agent Protocols

ILLC DS-2006-01: **Troy Lee**
Kolmogorov complexity and formula size lower bounds

ILLC DS-2006-02: **Nick Bezhanishvili**
Lattices of intermediate and cylindric modal logics

ILLC DS-2006-03: **Clemens Kupke**
Finitary coalgebraic logics

ILLC DS-2006-04: **Robert Špalek**
Quantum Algorithms, Lower Bounds, and Time-Space Tradeoffs

ILLC DS-2006-05: **Aline Honingh**
The Origin and Well-Formedness of Tonal Pitch Structures

ILLC DS-2006-06: **Merlijn Sevenster**
Branches of imperfect information: logic, games, and computation

ILLC DS-2006-07: **Marie Nilsenova**
Rises and Falls. Studies in the Semantics and Pragmatics of Intonation

ILLC DS-2006-08: **Darko Sarenac**
Products of Topological Modal Logics

ILLC DS-2007-01: **Rudi Cilibrasi**
Statistical Inference Through Data Compression

ILLC DS-2007-02: **Neta Spiro**
What contributes to the perception of musical phrases in western classical music?

ILLC DS-2007-03: **Darrin Hindsill**
It's a Process and an Event: Perspectives in Event Semantics

ILLC DS-2007-04: **Katrin Schulz**
Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity, and Conditionals

ILLC DS-2007-05: **Yoav Seginer**
Learning Syntactic Structure

ILLC DS-2008-01: **Stephanie Wehner**
Cryptography in a Quantum World

ILLC DS-2008-02: **Fenrong Liu**
Changing for the Better: Preference Dynamics and Agent Diversity

ILLC DS-2008-03: **Olivier Roy**
Thinking before Acting: Intentions, Logic, Rational Choice

ILLC DS-2008-04: **Patrick Girard**
Modal Logic for Belief and Preference Change

ILLC DS-2008-05: **Erik Rietveld**
Unreflective Action: A Philosophical Contribution to Integrative Neuroscience

ILLC DS-2008-06: **Falk Unger**
Noise in Quantum and Classical Computation and Non-locality

ILLC DS-2008-07: **Steven de Rooij**
Minimum Description Length Model Selection: Problems and Extensions

ILLC DS-2008-08: **Fabrice Nauze**
Modality in Typological Perspective

ILLC DS-2008-09: **Floris Roelofsen**
Anaphora Resolved

ILLC DS-2008-10: **Marian Counihan**
Looking for logic in all the wrong places: an investigation of language, literacy and logic in reasoning

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Simon Andreas Witzel**
Knowledge and Games: Theory and Implementation