WHAT DO WE NEED TO HEAR A BEAT?

The influence of attention, musical abilities, and accents on the perception of metrical rhythm

FLEUR L. BOUWER-ODIJK

What do we need to hear a beat?

The influence of attention, musical abilities, and accents on the perception of metrical rhythm

Fleur Leonie Bouwer

What do we need to hear a beat?

The influence of attention, musical abilities, and accents on the perception of metrical rhythm

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 Aula der Universiteit
op woensdag 8 juni 2016, te 11.00 uur

door

Fleur Leonie Bouwer

geboren te Haarlem

Contents

Acknowledgements

Dit proefschrift is het product van vijf jaar nieuwe dingen leren, vijf jaar nieuwsgierig zijn, en vooral vijf jaar hard werken. Heel veel mensen om mij heen hebben me hierbij geholpen, soms in praktische zin, maar veel vaker gewoon door er te zijn. Ik ben al deze mensen zeer dankbaar voor hun steun en vertrouwen. Zonder hen zou dit proefschrift er niet geweest zijn!

In de eerste plaats wil ik mijn promotor bedanken. Henkjan, jouw tomeloos en aanstekelijk enthousiasme voor de wetenschap in het algemeen en de muziekcognitie in het bijzonder zijn en blijven een inspiratie voor mij als wetenschapper. Jouw inzet om ons vakgebied op de kaart te zetten en te houden is bewonderenswaardig en hartverwarmend. Je hebt me de afgelopen jaren voortdurend gesteund in het onderzoek en mij de kans gegeven me te ontwikkelen als wetenschapper, zelfs als je daarvoor FGw-bergen moest verzetten. Bovenal heb je ondanks alle dagelijkse beslommeringen nooit de menselijke kant van het werk uit het oog verloren en ben je daarom altijd een hele fijne begeleider geweest. Bedankt voor vijf prachtige jaren!

Secondly, I would like to thank my co-supervisor. Jessica, you gave me guidance, not only on fMRI techniques, research-related issues, and proper writing in English (notice the comma!), but also on life as a (female) scientist, which I am very grateful for. When I came to LondON I got a warm welcome, and I was included in the Grahnlab family without any reservations. You made me feel at home away from home. I loved having the opportunity to learn from you!

Members of my committee, Sonja, Julia, Paul, Victor, Erik, and István, thank you for taking the time to read and evaluate this dissertation and for taking the time to be present for my defense ceremony.

Makiko, Ashley, Paula, Berit, Carlos, Joey, Bastiaan, Gábor, Titia, Conor, and Aline, dear fellow MCG members, music cognition buffs, and co-authors, thank you for all our endless conversations over lunch and tea, whether on research, life as a researcher, or simply life. I couldn't have done it without your support! Dan, Aaron, Tram, Dirk, Li-Ann, Roxy, Sarah, and Sarah, I will always feel a little bit part of the Grahnlab as well. Thank you for welcoming me into your group and for making my time in LondON unforgettable.

Lieve studenten, Carola, Myrthe, en Tom, wat ben ik trots geweest om jullie begeleider te mogen zijn. Ik heb jullie hopelijk iets kunnen leren, maar ik heb vooral zelf ook heel veel geleerd van het begeleiden van jullie projecten. Daarnaast was het ontzettend gezellig om het lab met jullie te delen! Dank voor jullie inzet en voor jullie geduld met mij als begeleider.

Dirk, je had voor iedere kapotte elektrode en crashende computer wel een oplossing. Als het juiste kabeltje of apparaat er niet was maakte jij het gewoon, altijd met engelengeduld en een goed humeur. Dank voor al je steun in en rond het lab!

Grote dank aan al die mensen, familie, vrienden, collega's, en onbekenden, die aan mijn experimenten hebben meegedaan en die soms urenlang de meest saaie geluidjes hebben aangehoord. Het spreekt voor zich dat ik dit proefschrift nooit had kunnen afronden zonder jullie! Daarbij was het werk in het lab een stuk leuker door jullie gezelschap.

Ik wil ook graag de docenten bedanken die me hebben opgeleid voordat ik begon met promoveren. Annette, bedankt voor het vertrouwen en de begeleiding tijdens mijn master project, waarin je me zoveel leerde over onderzoek doen en mij mijn eerste stapjes als wetenschapper liet zetten. Herman, jij bracht me de liefde voor de muziek en de klarinet bij. Mijn tijd aan het conservatorium heeft me gevormd als mens en geleerd om door te zetten en hard te werken. Dank voor je steun en begeleiding, en de goede gesprekken die we ook nu nog hebben.

Zonder mijn familie had ik het niet volgehouden de afgelopen jaren. In de eerste plaats natuurlijk mijn ouders. Lieve papa en mama, jullie hebben me altijd gesteund met woorden, daden, en simpelweg door er te zijn. Jullie vertrouwen in mijn kunnen en jullie trots, maar ook jullie aanmoediging om het beste uit mijzelf te halen hebben mij gemaakt wie ik ben. Zonder dit alles was dit proefschrift er zeker niet geweest! Lieve bonusouders, Bo en Rebecca, dank voor jullie eindeloze geduld tijdens het tot stand komen van dit proefschrift en voor de goede keukentafelgesprekken die we hebben gehad over werk en wetenschap. Jullie hebben in jullie steun voor mij altijd pal naast mijn ouders gestaan en daar ben ik jullie heel erg dankbaar voor!

Lieve David, dank je voor je niet aflatende geduld met je wispelturige kleine zusje en je lieve woorden, aanwezigheid, en begrip op al die momenten dat het nodig was. Danielle, Eef, Roos, en Gideon, dank voor alle prachtige momenten die we samen hebben beleefd en gevierd. Dat het er nog maar veel meer mogen worden!

Ook wil ik mijn schoonfamilie bedanken. Nardi en Marja, Jasper en Marleine, en Pieter en Steph, jullie hebben me met heel veel liefde en warmte opgenomen binnen de Odijken en waren een geweldige steun in de afgelopen jaren!

Daan, je weet waarom. Ik hou van je.

Contributions

1. Introduction

Fleur L. Bouwer wrote and revised the introduction, with contributions from Henkjan Honing and Jessica A. Grahn.

2. Perceiving temporal regularity in music: The role of auditory event-related potentials (ERPs) in probing beat perception

Chapter 2 is adapted from a chapter published in *Neurobiology of Interval Timing* (Henkjan Honing, Fleur L. Bouwer, & Gábor Háden, 2014). The chapter has been adapted to reflect primarily the sections on which FLB was the lead author. For the majority of the sections included in this dissertation, FLB wrote a first version and revised the text, with contributions from HH and GH. Smaller sections are included that were written in a first version by HH and GH, and that were revised by all three authors.

3. What makes a rhythm complex? The influence of musical training and accent type on beat perception

Chapter 3 has been submitted for publication (Fleur L. Bouwer, J. Ashley Burgoyne, Daan Odijk, Henkjan Honing, & Jessica A. Grahn, 2016). FLB, HH, and JAG conceived of the research questions for this chapter. FLB designed the stimuli and the experiment. DO created the web-based application to run the experiment. JAB ran the ordinal regression analysis. FLB wrote and revised the paper, with contributions to the text from all other authors.

Natural Sciences and Engineering Research Council (NSERC) and the James S. McDonnell Foundation.

4. Temporal attending and prediction influence the perception of metrical rhythm: evidence from reaction times and ERPs

Chapter 4 has been published in *Frontiers in Psychology* (Fleur L. Bouwer & Henkjan Honing, 2015). FLB and HH posed the research question. FLB designed the experiment, ran the experiment, conducted the analyses, and wrote and revised the paper, with contributions from HH.

5. Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study

Chapter 5 has been published in *PLoS ONE* (Fleur L. Bouwer, Titia L. Van Zuijen & Henkjan Honing, 2014). FLB, TLvZ, and HH conceived and designed the experiment. FLB created the stimuli, ran the experiment, conducted the analyses, and wrote and revised the text, with contributions from TLvZ and HH.

6. Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm

Chapter 6 has been published in *Neuropsychologia* (Fleur L. Bouwer, Carola M. Werner, Myrthe Knetemann, & Henkjan Honing, 2016). FLB and HH proposed the research question and designed the experiment. FLB, CMW, and MK ran the experiment. FLB conducted the statistical analyses and wrote and revised the text, with contributions from HH, CMW and MK.

7. Discussion

Fleur L. Bouwer wrote and revised the discussion, with contributions from Henkjan Honing and Jessica A. Grahn.

Chapter 1

<div align="right">

Introduction

</div>

Humans all over the world engage in musical behavior. We use music to celebrate and to mourn, in rituals and in play, and almost always we perform music together, in a social setting (Trehub, Becker, & Morley, 2015). Making and experiencing music in a group requires us to synchronize our behavior to each other and to the music. To synchronize precisely to musical events and with fellow performers, we need to initiate movements before a musical event has occurred, which requires prediction of exactly when the next tone will be played. The process that allows us to make these predictions, and that is therefore crucial to successful synchronization, is *beat perception*.

Beat perception seems a trivial task. When music is played during half time at a football match, thousands of people may synchronize to the music by swaying their bodies or clapping along. The ability to perceive a beat in music does not require formal training (Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015). Indeed, even very inexperienced listeners – infants – are sensitive to the beat in music (Hannon & Johnson, 2005; Phillips-Silver & Trainor, 2005; Zentner & Eerola, 2010). Moreover, we may be capable of extracting a beat from a rhythm even without attention directed at a rhythm (Ladinig, Honing, Háden, & Winkler, 2009).

From a comparative and computational standpoint, the ease with which humans sense the beat in music is puzzling. Monkeys have only a very limited capacity for beat perception (Merchant & Honing, 2014) and while various attempts have been made to create a computational model that is capable of finding a beat, most of these models cannot deal with the complex input of real music (Honing, 2013; Temperley, 2013). Recently, both the ubiquity of beat perception abilities in humans and the validity of the stimuli that are used to test beat perception have been questioned (Tranchant & Vuvan, 2015), raising the question of whether we really are so apt at extracting a beat from music. Do we perceive a beat even when we are not paying attention to a rhythm? Do we need training to do so? And does it matter how the beat is indicated in the music?

In this dissertation, I address these questions in an exploration of what the necessary ingredients are that allow us to perceive a beat. I examine the characteristics of the context, the listener, and the music, and I examine the processes underlying beat perception. In addition, throughout the dissertation, several issues concerning how beat perception may be measured in a more controlled way are discussed. In this introduction, I will briefly clarify the terminology pertaining to beat perception and give a short

description of the mechanisms thought to contribute to beat perception. Subsequently, the methods used to test beat perception are briefly discussed and the ambiguity of the term *attention* is addressed. Finally, I will give a short overview of the remaining chapters of this dissertation.

1.1   Beat perception terminology

In everyday language, the term *rhythm* is often associated with dancing and moving to music. In this dissertation however, a more strict definition of rhythm will be used, as is common in the literature. A rhythm is defined as a succession of events in time (Honing, 2013; London, 2012). In auditory rhythm, the temporal organization of events is defined by the inter-onset intervals (IOIs) between event onsets, and not by their duration (London, 2012). Rhythm can be found in many domains, including both music and language. What makes musical rhythm special is that we often perceive regularity in a rhythm in the form of "regularly recurring, precisely equivalent" psychological events (Cooper & Meyer, 1960, p.3; Grahn, 2012; Large, 2008), known as the *beat*, the pulse, or the tactus (Lerdahl & Jackendoff, 1983a). In addition to a beat, we can perceive higher-order regularities in the form of recurrent strong and weak beats (referred to as *meter*) and lower-order regularities (termed *subdivisions*). Together, these regularities at multiple levels constitute a hierarchical framework, which will be referred to as the *metrical structure*. Note that sometimes the term *meter* is used to denote the entire metrical hierarchy and not just the highest level of regularity. The beat is the most salient level of regularity in the metrical hierarchy and is the level of regularity at which people usually tap along with the music. People in general prefer a beat rate of around 100 beats per minute (London, 2012), though the exact preferred rate varies across individuals and can change with age and musical training (Drake, Jones, & Baruch, 2000). The rates of meter and subdivisions are usually related to the beat rate at integer ratios, with the rate of meter being two or three times slower and the rate of subdivisions being two or three times faster than the beat rate (London, 2012).

A beat is inferred from musical rhythm through *accents*. Lerdahl and Jackendoff (1983a, p. 17) describe phenomenal accents as "any event at the musical surface that gives emphasis or stress to a moment in the musical flow". An accent may be any event that is salient to the listener (Ellis & Jones, 2009), like a sudden increase in intensity, a change in timbre, a pitch leap or a harmonic change. In addition, accents may arise from the temporal grouping structure of events in a rhythm (Povel & Okkerman, 1981; Povel & Essens, 1985). If accents are organized in a structural way, with regular time intervals separating them, we can use them to infer a metrical structure from a rhythm (Lerdahl & Jackendoff, 1983a; Povel & Essens, 1985).

One important thing to note is that the beat is a psychological construct, which only loosely relates to the structure of a rhythm (Honing, 2013; Large, 2008; London, 2012). Although we infer a beat from the sensory input, once established, the percept of a metrical structure can remain stable even if a rhythm does not fit the perceived metrical structure. The perceived metrical structure will only be changed if the rhythm provides strong enough evidence for an alternative structure (Lerdahl & Jackendoff, 1983a).

When the evidence is not strong enough to change the perceived meter, accents can occur at metrically weak moments and events can be left out at metrically strong moments, as occurs during syncopation (Honing, 2013).

1.2   Mechanisms of beat perception

A widely accepted theory of beat perception is the Dynamic Attending Theory (DAT, see Jones & Boltz, 1989; Jones, 2009; Large & Jones, 1999). According to DAT, attention is not constant over time, but constantly fluctuates (Henry & Herrmann, 2014). The fluctuations in attentional energy exhibit regularity and can be described using nonlinear oscillator models, which have been linked to neural oscillations (Henry & Herrmann, 2014; Large, Herrera, & Velasco, 2015; Large & Jones, 1999; Large, 2008). The phase and period of the regular fluctuations in attentional energy can adapt, or *entrain*, to an external regularity. When music exhibits regularity in time, as when accents are regularly spaced, internal rhythms of attentional energy entrain to the music and, if this occurs, we perceive a beat. Peaks in attentional energy then coincide with metrically strong moments. Several properties of oscillator models can theoretically be linked to the phenomenological properties of perceiving a beat in music. For example, oscillators have the capacity for entrainment to an external stimulus, they can spontaneously occur, they can remain stable if the external stimulus is removed or changed (as in a syncopation), and oscillator models predict higher-order resonance, which may account for hearing regularity at multiple different levels of a metrical hierarchy (Large, 2008).

DAT predicts that attentional energy is heightened at metrically salient points in a rhythm, which is thought to lead to a processing advantage for events that coincide with attentional peaks (Large & Jones, 1999). Behavioral studies have provided support for this notion, by showing enhanced performance at temporally expected times on time judgment tasks (Barnes & Jones, 2000; McAuley & Fromboluti, 2014), pitch judgment tasks (Jones, Moynihan, MacKenzie, & Puente, 2002), phoneme monitoring (Quené & Port, 2005), and even visual tasks (Escoffier, Herrmann, & Schirmer, 2015; Escoffier, Sheng, & Schirmer, 2010). Recently however, in a more controlled experiment, the findings for processing of pitch could not be replicated (Bauer, Jaeger, Thorne, Bendixen, & Debener, 2015), questioning the ubiquity of dynamic attending.

An alternative theory explaining the perception of metrical structure comes from the framework of predictive coding (cf. Vuust & Witek, 2014). The theory of predictive coding views the brain from a Bayesian perspective (Friston, 2005) and essentially considers the brain as a prediction machine (Clark, 2013). A mental representation of the world is used to predict incoming information, and this representation is continuously updated based on the discrepancy between the prediction and the actual sensory input (the prediction error). Such a model of the brain is reminiscent of ideas about beat perception, which also rely on a mental model (the metrical structure), which is used to predict incoming sensory information (the rhythm) and is updated by that very same information. Based on these parallels, Vuust and Witek (2014) have suggested that the perception of metrical structure can be understood within the framework of predictive coding.

Both DAT and predictive coding describe an interplay between top-down and bottom-up processes, with a perceived metrical structure both being inferred from a rhythm and influencing processing of that same rhythm (Vuust & Witek, 2014). However, the way that metrical structure affects processing of a rhythm is described subtly differently, with DAT emphasizing the role of attention, and predictive coding emphasizing the role of prediction. These processes are closely related and often it is unclear whether attention or prediction is tested, as both can be affected by probabilistic information and temporal regularity (Lange, 2013; Schröger, Kotz, & SanMiguel, 2015). Moreover, prediction may guide the focus of attention (Schröger, Kotz, et al., 2015), while attentional focus, as described by DAT, may lead to predictions (Large & Jones, 1999). At the same time, the effects of attention and prediction on processing of sensory information are in fact opposite, with attention enhancing early sensory responses and prediction attenuating them (Lange, 2013). While it is unclear how DAT and predictive coding relate to each other, importantly, both theories of beat perception assume that processing of rhythmic events is influenced by the perceived metrical structure and that a rhythm is expected to adhere to the internal metrical representation.

1.3   Research methods

Behavioral studies have often studied beat perception by looking at overt motor responses to a beat (e.g., tapping to a beat; for a review see Repp, 2005). However, beat perception is thought to exist from a very young age, before precise entrainment of actions to a beat is possible (Phillips-Silver & Trainor, 2005; Winkler, Háden, Ladinig, Sziller, & Honing, 2009). Also, while entrainment to music is a universal human behavior, many people find precise entrainment to a beat somewhat difficult, and this motoric skill may require specific training to perfect (Schaefer & Overy, 2015). To measure beat perception, especially in untrained subjects, purely perceptual methods like discrimination tasks (cf. Grahn & Brett, 2007) or ratings tasks (cf. Grube & Griffiths, 2009) may be a more useful approach. In some cases, a behavioral response may not be possible at all (e.g., when a rhythm is not attended). Neuroimaging techniques may then be used to show whether subjects perceived a beat and additionally shed some light on the neural mechanisms underlying beat perception.

Several studies have used fMRI to examine activity elicited by rhythms that vary in metrical complexity. By comparing the response to rhythms that contain a strong beat with the response to rhythms that contain a weak beat or no beat at all, it has been repeatedly shown that a network of motor areas, specifically the supplementary motor area (SMA) and basal ganglia, is involved in the perception of metrical structure (Bengtsson et al., 2009; Chen, Penhune, & Zatorre, 2008a; Grahn & Brett, 2007; Grahn & Rowe, 2009, 2013; Kung, Chen, Zatorre, & Penhune, 2013; Teki, Grube, Kumar, & Griffiths, 2011). Using EEG, several studies have focused on the role of oscillatory activity in beat perception, as neural resonance has been suggested to relate to DAT (Large, 2008). The results however have been mixed, with studies showing a role for beta synchronization (Cirelli et al., 2014; Fujioka, Trainor, Large, & Ross, 2012), beta desynchronization following the beat (Fujioka, Ross, & Trainor, 2015), beta desynchronization preceding the beat (Te Woerd, Oostenveld, de Lange, & Praamstra, 2014), gamma synchronization (Fujioka, Trainor, Large, & Ross, 2009; Zanto, Snyder,

& Large, 2006) and phase alignment of delta oscillations (Nozaradan, Peretz, & Mouraux, 2012; Nozaradan, Peretz, Missal, & Mouraux, 2011). Moreover, in most of these studies, very sparse and often isochronous stimuli without clear accents indicating a metrical structure were used, which may not be optimal to induce a beat (Tierney & Kraus, 2013).

In this dissertation, a different approach has been used to study beat perception, which utilizes the proposed influence a perceived metrical structure has on processing of an incoming rhythm. If a beat is perceived, the responses to rhythmic events that differ in metrical salience should also differ. Using EEG to measure event-related potentials (ERPs), it is possible to measure several well-studied responses to sound and thus probe beat perception by comparing the responses to events in metrically strong and metrically weak positions.

First, one can examine the obligatory ERP responses to sound. Both attention and prediction can affect the auditory P1 and N1 responses, which are part of the auditory evoked potentials, a series of well-studied ERP components that are elicited by sound (Näätänen & Picton, 1987; Picton, Hillyard, Krausz, & Galambos, 1974). While attention enhances these components (Picton & Hillyard, 1974; Woldorff et al., 1993), prediction attenuates them (Lange, 2009; Schwartze, Farrugia, & Kotz, 2013). As the amount of attentional energy directed toward rhythmic events and their predictability may depend on their metrical salience, the ERP responses to rhythmic events should be affected by a perceived metrical structure, even if the sounds themselves are identical.

A second approach to probing beat perception by looking at its influence on processing of rhythmic events is to violate the predictions that are generated by a perceived metrical structure. Several ERP components, including the mismatch negativity (MMN), the N2b, the P3a and the P3b, are elicited by unexpected auditory events and are sensitive to the magnitude of a regularity violation (Näätänen, Paavilainen, Rinne, & Alho, 2007; Polich, 2007). A perceived metrical structure influences how unexpected a violation of regularity is, as the metrical expectations differ between metrical positions. In addition, peaks in attentional energy may enhance the detection of a regularity violation in metrically strong but not weak positions. Thus, like the obligatory responses to sound, the ERP components elicited by regularity violations should be affected by metrical position if a beat is perceived.

The approach to measuring beat perception in this dissertation is thus to probe whether a listener perceives a beat not by measuring beat perception itself, but rather by looking at the result of the perceived beat: the way beat perception affects processing of rhythmic events. The advantage of this approach is that we measure ERP components that have been studied extensively. However, it is important to note that these components are highly susceptible to the physical and probabilistic properties of auditory events (Luck, 2005; Woodman, 2010). Thus, using this approach to studying beat perception, the biggest challenge lies in sufficient experimental control to ensure that any differences between responses to rhythmic events in different metrical positions can be attributed to beat perception.

1.4 Attention

In this dissertation, the necessary ingredients for beat perception to occur are examined. One of the ingredients that is considered is the direction of attention. The term attention, especially in the context of beat perception research, can be somewhat ambiguous (Henry & Herrmann, 2014). It is both used to indicate the fluctuating attentional resources as described by DAT (Large & Jones, 1999), and to describe the more general cognitive notion of enhancement of task-relevant information (i.e., whether we selectively attend to a stimulus or not). However, these two usages of the term attention are not necessarily mutually exclusive. Selective attention to task-relevant information is related to the notion of limited general processing resources (Gazzaley & Nobre, 2012; Kiyonaga & Egner, 2013), with selective attention indicating the direction of the general processing resources when aimed at some external event. While processing resources may fluctuate with the metrical structure, as proposed by DAT, processing resources directed at a rhythm may generally be enhanced or attenuated by the direction of selective attention. In addition, while attention is often described as a top-down process, external events can involuntarily capture and guide our attention (Rinne, Särkkä, Degerman, Schröger, & Alho, 2006). Thus, even if selective attention is not directed at a rhythm, it is still possible for attentional resources to fluctuate according to the external input.

1.5 Outline

In this dissertation, I examine what is needed for a listener to perceive a beat in music. Properties of the context, the listener, and the music are examined with behavioral and neuroimaging methods. In addition, I examine the processes underlying beat perception, and I address several questions related to stimulus design. First, in Chapter 2, I provide a more thorough theoretical overview of beat perception, its importance, and how to study it using ERPs.

In Chapter 3, we look at three components that may influence whether we are able to perceive a beat in a rhythm: the complexity of the rhythm (i.e., in how much the structure of accents in the rhythm fits a metrical structure), the type of accents that indicate the metrical structure (e.g., intensity increases or phenomenal accents created by the grouping structure of the rhythm), and the musical experience of the listener. Using a web-based experiment, we show that existing models of the relationship between the structure of temporal accents and a perceived beat are not necessarily applicable to rhythms with intensity accents. In addition, the results suggest that musical training enhances sensitivity to the structure of accents in a rhythm that indicates the beat.

In Chapter 4, the contributions of attention and prediction to beat perception are examined. The results from a speeded detection task suggest that both temporal attending and temporal prediction influence the responses to metrical rhythm, and that these two processes interact. Using EEG, the presence of temporal attending and prediction is probed while attention is not directed at the rhythm. In a group of highly trained musicians, we provide tentative evidence that both processes are active even with lower processing resources available.

While in Chapter 4 rhythms with very sparse cues indicating the metrical structure are used, in Chapter 5 beat perception is probed using rhythms that have very clear accents, with intensity, timbre, and duration indicating the beat. ERP responses to unexpected silences are recorded while attention is directed away from the rhythm. Responses of professional musicians are compared to responses of non-musicians to examine the effects of musical training. We show that even in untrained participants, beat perception can influence the responses to rhythmic events when attention is not directed at the rhythm.

In Chapter 6, the influence of attention and musical abilities on beat perception is further examined. Here, we specifically aim to disentangle beat perception from confounding processes that may also influence responses to rhythmic events. ERP responses to intensity decrements in different metrical positions are recorded both with and without attention directed at the rhythm, We compare responses to decrements in regular sequences, in which both beat perception and sequential learning can occur, to responses to decrements in irregular sequences, in which the statistical properties of the sequence are preserved, allowing for sequential learning, but in which no beat can be perceived. We show that beat perception, independently of sequential learning, affects processing of rhythmic events without attention directed at a rhythm. In addition, we show that sequential learning influences responses to rhythmic events, even when a rhythmic sequence is temporally irregular. Interestingly, musical abilities affect responses to metrical rhythm only when attention is directed at the rhythm.

In Chapter 7, I summarize the main findings of this dissertation and offer some concluding remarks about the mechanisms underlying beat perception.

Chapter 2

Perceiving temporal regularity in music: The role of auditory event-related potentials (ERPs) in probing beat perception[*]

The aim of this chapter is to give an overview of how the perception of a regular beat in music can be studied in human adults using event-related brain potentials (ERPs). Next to a review of the recent literature on the perception of temporal regularity in music, we will discuss in how far ERPs, and especially the component called mismatch negativity (MMN), can be instrumental in probing beat perception. We conclude with a discussion on the pitfalls and prospects of using ERPs to probe the perception of a regular beat, in which we present possible constraints on stimulus design and discuss future perspectives.

2.1 Introduction

In music, as in several other domains, events occur over time. The way events are ordered in time is commonly referred to as *rhythm*. In musical rhythm, unlike in other domains, we often perceive an underlying regularity in time, which is known as the pulse or the *beat*. The beat is a regularly recurring salient moment in time (Cooper & Meyer, 1960). The beat often coincides with an event, but a beat can also coincide with plain silence (Longuet-Higgins & Lee, 1984, see Figure 2.1). At a higher level, we can hear regularity in the form of regular stronger and weaker beats and at a lower level, we can perceive regular subdivisions of the beat. We thus can perceive multiple levels of regularity in a musical rhythm, which together create a hierarchical pattern of saliency known as *metrical structure* or simply, *meter*. In this chapter, we will mainly focus on the processes underlying the perception of the most salient level of regularity in this perceived metrical structure: the beat.

The sensory and cognitive mechanisms of beat perception have quite a history as a research topic (Clarke, 1999; Fraisse, 1982; Honing, 2013; Large & Jones, 1999; London, 2012; Povel & Essens, 1985). These mechanisms have been examined in many music perception studies, mostly from a theoretical and psychological point of view (Desain & Honing, 1999; Large & Jones, 1999; Parncutt, 1994; Povel & Essens, 1985). More recently, beat perception has attracted the interest of developmental psychologists (Hannon & Trehub, 2005a), cognitive biologists (Fitch, 2006), evolutionary psychologists (Honing & Ploeger, 2012), and neuroscientists (Grahn & Brett, 2007; Grube, Cooper, Chinnery, & Griffiths, 2010). In addition, in the last decades a change can be observed from studying beat perception from a psychophysical perspective (studying the relation between stimulus and sensation) using relatively simple stimulus materials (Handel, 1989), to studying beat perception with more ecologically valid materials that take the task and the effect of musical context into account (Clarke &

Figure 2.1 A rhythm notated in common music notation (labeled Score) and as dashes (sound) and dots (silence) on a grid (labeled Rhythm). The perceived beat is marked with bullets; one possible metrical interpretation is marked with a metrical tree, with the length of the branches representing the theoretical metric salience and bullets marking the regularities at each metrical level. The rest (labeled R) marks a 'loud rest' or syncopation: a missing event on a perceived beat.

Cook, 2004; Honing, 2013). In its entirety this has resulted in a substantial body of work using a variety of methods. In this chapter we will focus on studying the perception of the beat using electrophysiological methods.

2.2   Beat perception as a fundamental cognitive mechanism

It seems a trivial skill: children that clap along with a song, musicians that tap their foot to the music, or a stage full of line dancers that dance in synchrony. And in a way it is indeed trivial. Most people can easily pick up a regular pulse from the music or can judge whether the music speeds up or slows down. However, the realization that perceiving this regularity in music allows us to dance and make music together makes it a less trivial phenomenon. Beat perception might well be conditional to music (Honing, 2012), and as such it can be considered a fundamental human trait that, arguably, has played a decisive role in the origins of music (Honing & Ploeger, 2012). Three properties of the ability to perceive a beat can be looked at when considering its role in the origins of music: whether it is an innate (or spontaneously developing) ability, whether it is specific to the domain of music, and whether it is a species-specific ability.

2.3   Innateness, domain-, and species-specificity

Scientists are still divided whether beat perception develops spontaneously (emphasizing a biological basis) or whether it is learned (emphasizing a cultural basis). Some authors consider a sensitivity to the beat to be acquired during the first years of life, suggesting that the ways in which babies are rocked and bounced in time to music by their parents is the most important factor in developing a sense for metrical structure (Phillips-Silver & Trainor, 2005). By contrast, more recent studies emphasize a biological basis, suggesting that beat perception is already functional in young infants (Zentner & Eerola, 2010) and possibly even in 2–3 day old newborns (Winkler et al., 2009). These recent empirical findings can be taken as support for a genetic predisposition for beat perception, rather than it primarily being a result of learning.

Furthermore, developmental studies suggest that infants are not only sensitive to a regular pulse, but also to regularity at a higher level (two or more levels of pulse; Hannon & Johnson, 2005). Thus it is possible that humans possess some processing predisposition to extract hierarchically structured regularities from music (Ladinig et al., 2009; Ladinig, Honing, Háden, & Winkler, 2011). To understand more about these capacities to hear regularity in music and to examine whether they are indeed (partly) innate, research with newborns provides a suitable context (Honing, 2012; Winkler et al., 2009).

With regard to the domain-specificity of beat perception convincing evidence is still lacking, although it was recently argued that beat perception does not play a role (or is even avoided) in spoken language (A. D. Patel, 2008). Furthermore, the perception of a beat occurs more easily with auditory than visual temporal stimuli (Repp & Penel, 2002), with audition priming vision (Bolger, Trost, & Schön, 2013), but not vice versa (Grahn, Henry, & McAuley, 2011).

With regard to the species specificity of beat perception, it is still unclear which species have this ability. It was recently shown that rhythmic entrainment, long considered a human-specific mechanism, can be demonstrated in a select group of bird species (Hasegawa, Okanoya, Hasegawa, & Seki, 2011; A. D. Patel, Iversen, Bregman, & Schulz, 2009; Schachner, Brady, Pepperberg, & Hauser, 2009), and not in more closely related species such as nonhuman primates (Honing, Merchant, Háden, Prado, & Bartolo, 2012; Zarco, Merchant, Prado, & Mendez, 2009). This is surprising when one assumes a close mapping between a genetic predisposition (specific genotypes) and specific cognitive traits. However, more and more studies show that genetically distantly related species can show similar cognitive skills; skills that more genetically closely related species fail to show (De Waal & Ferrari, 2010). The observations regarding beat perception in animals support the *vocal learning hypothesis* (A. D. Patel, 2006) that suggests that rhythmic entrainment is a by-product of the vocal learning mechanisms that are shared by several bird and mammal species, including humans, but that are only weakly developed, or missing entirely, in nonhuman primates (Fitch, 2009). Nevertheless it has to be noted that, since no evidence of rhythmic entrainment was found in many vocal learners (including dolphins and songbirds; Schachner et al., 2009), vocal learning may be necessary, but clearly is not sufficient for beat perception and rhythmic entrainment. Furthermore, vocal learning itself may lie over a continuum rather than being a discrete ability, as for example sea lions (*Zalophus californianus*) seem capable of rhythmic entrainment (Cook, Rouse, Wilson, & Reichmuth, 2013) while there is little or no evidence of vocal learning (Arnason et al., 2006). Whereas research in human newborns can answer questions about the innateness of beat perception, research in various animals can answer questions about the species-specificity of beat perception.

2.4 Beat induction

Sometimes, the term *beat induction* is used for the cognitive mechanism that supports the detection of a regular pulse from the varying surface structure of musical sound. This term stresses that the perception of a beat is not a passive process but an active one in which a listener induces a particular regular pattern from a rhythm. It emphasizes that a beat does not always need to be physically present in order to be perceived. This is, for example, the case when we hear a *syncopation* (or 'loud rest'; see Figure 2.1), in which the beat does not coincide with an event in the musical surface, but with a silence (Honing, 2012).

As we have seen, beat perception and beat induction can be considered fundamental to music perception and production. Questions of innateness, domain-specificity and species-specificity need to be addressed to further reveal the relationship between beat perception and the origins of music. Before we turn to a possible method to answer questions about beat perception, first, the possible mechanisms that constitute beat perception will be discussed.

2.5   Possible mechanisms of beat perception

2.5.1   The perception of a beat

The perception of a beat is a bi-directional process: not only can a varying musical rhythm induce a regular beat; a regular beat can also influence the perception of the very same rhythm that induces it. Hence beat perception can be seen as an interaction between bottom-up and top down sensory and cognitive processes (Desain & Honing, 1999). Initially, we induce a beat from various cues in the music. Once a context of regularity is established, we use the inferred beat to interpret the music within this context and to predict future events (London, 2012). A perceived pulse is stable and resistant to change (Large, 2008). However, if the sensory input provides clear evidence for a different metrical structure, our perception of the beat can change. The relation between the events in the music and the perceived temporal regularity thus is a flexible one, in which the perceived metrical structure is both inferred from the music and has an influence on how we perceive the music (Desain & Honing, 2003; Grube & Griffiths, 2009).

2.5.2   Boundaries on beat perception

We can perceive regularity in music at different metrical levels and thus at different timescales. It should be noted that the perception of temporal regularity is restricted by several perceptual boundaries. We can perceive temporal regularity with a period roughly in the timescale of 200 to 2000 ms (London, 2002). Within this range, we have a clear preference for beats with a period around 600 ms or 100 beats per minute. This rate is referred to as *preferred tempo* (Fraisse, 1982). A beat at this tempo is usually very salient. Most empirical studies looking at beat perception use a rate of stimulus presentation that makes it possible to hear a beat at preferred tempo level.

2.5.3   Beat perception through accent structure

To infer a metrical structure from music we make use of accents. In a sequence of events, an accent is a more salient event because it differs from other, non-accented events along some auditory dimension (Ellis & Jones, 2009). When accents exhibit regularity in time, we can induce a regular beat from them. Accented tones are then usually perceived as on the beat or, on a higher level, as coinciding with a strong rather than a weak beat (Lerdahl & Jackendoff, 1983b).

A sequence of events in time, such as a musical rhythm, also contains purely *temporal accents* that arise from the structure of event onsets rather than from acoustic changes in the sound. Events are perceived as more or less salient depending on their length and position in a rhythm. Povel and Essens (1985) describe three ways in which a temporal accent can occur. First, when an onset is isolated relative to other onsets, it sounds like an accent. Second, when two onsets are grouped together, the second onset sounds accented. Finally, for groups of three or more onsets, the first and/or last tone of the group will be perceived as an accent.

While it has been suggested that beat perception is mainly guided by these temporal accents (Snyder & Krumhansl, 2001), recently it has been shown that pitch accents

also play a role in perceiving the beat (Ellis & Jones, 2009; Hannon, Snyder, Eerola, & Krumhansl, 2004). It is very likely that in natural music, many features of tones can contribute to an accent structure and our perception of the beat, including not only pitch, but also timbre and intensity. In line with this, Bolger et al. (2013) and Tierney and Kraus (2013) showed that the use of ecologically valid stimuli can actually enhance the perception of a beat. However, to date, melodic, timbre and intensity accents have been largely ignored in many studies examining beat perception.

2.5.4 Beyond accents

While accents explain a large part of how we infer a beat and a metrical structure from music, several other processes must be taken into account. First, it must be noted that we sometimes perceive temporal structure without any accents present. Rather, we actually imagine accents where they are not physically present. This phenomenon has been termed *subjective rhythmization* and is very apparent when listening to a clock. Whereas every tick of a clock is equal, we often hear every other tick as an accent (i.e., "tick-tock" instead of "tick-tick"). Direct evidence for the presence of subjective rhythmization in isochronous sequences comes from studies comparing the brain response to tones in odd positions (which are subjectively accented) with the response to tones in even positions (which are not subjectively accented). It was found that slightly softer tones were perceived as more salient in odd than in even positions (Brochard, Abecasis, Potter, Ragot, & Drake, 2003). While this shows the presence of the effect, the mechanism underlying subjective rhythmization is still unclear (Potter, Fenwick, Abecasis, & Brochard, 2009).

A second influence on beat perception is our previous experience. (Hannon & Trehub, 2005b) showed how cultural background and exposure to music can affect how well we can discern a metrical structure. In their study, participants listened to folk melodies with either a simple or a complex metrical structure. They were subsequently presented with two alterations of the melody, one in which the metrical structure was preserved, and one in which the metrical structure was violated. Participants then rated the similarity of the altered melodies to the original melody. Adults of Bulgarian and Macedonian origin, who are accustomed to complex metrical structures (i.e., compound meters like 5/8 or 7/8), differentiated between structure-preserving and structure-violating alterations in both complex and simple metrical structures. However, participants with a Western background did so only in the melodies with a simple meter. This was most likely due to the fact that Western listeners are not familiar with complex meters. Interestingly, 6 month-old infants responded differentially to structure-preserving and structure-violating alterations regardless of whether they occurred in a simple or complex metrical structure. This implies that the difference between the adults from Western and Balkan cultures is due to enculturation, which takes place sometime after the age of 6 months. It shows that the culture with which we are familiar influences how we perceive the metrical structure (for more evidence regarding the effect of culture on beat and meter perception, see Gerry, Faux, & Trainor, 2010). In addition to the familiarity with different metrical structures, our culture can also provide us with template of certain patterns that specify a certain metrical structure. For example, snare drum accents in rock music often indicate the *offbeat* rather than the beat (London, 2012).

Finally, in addition to the influence of an accent structure, subjective rhythmization, and our previous experience, the perception of a beat can also be guided by conscious effort. By consciously adjusting the phase or period of the regularity we perceive, we can influence which tones we hear on the beat. For example, when we listen to an isochronous series of tones, without any instruction, we will hear every other tone as accented (Potter et al., 2009). However, by conscious effort, we can project a beat on every third tone, thus adjusting the period of the beat to our will. This ability has been very useful in examining beat and meter perception, because it can allow us to hear a physically identical stimulus as on the beat or not, depending on the instructions (for examples, see Iversen, Repp, & Patel, 2009; Nozaradan et al., 2011). Any change in neural activity found can then reliably be attributed to beat perception, without having to control for physical differences between tones that are on or off the beat.

To summarize, beat perception is guided by the temporal and acoustic structure of events. It is constrained by our perceptual system and can be influenced by our earlier exposure to music, subjective rhythmization, and conscious effort. When we listen to music, we induce a beat from the sensory input and then use that information to predict future events within a metrical framework. One way of understanding the mechanisms of beat perception is in the framework of the *predictive coding theory* (see Vuust, Gebauer, & Witek, 2014). Another prominent theory explaining the interaction between the varying sensory input and beat perception is the Dynamic Attending Theory (Jones, 2009).

2.5.5 Dynamic Attending Theory

Dynamic Attending Theory (DAT) explains the perception of metrical structure as regular fluctuations in attention. It proposes that internal fluctuations in attentional energy, termed *attending rhythms*, generate expectancies about when future events occur. When attentional energy is heightened an event is expected. Such a peak in attentional energy is perceived as a metrically strong position (i.e., on the beat). The internal fluctuations in attentional energy can entrain to the rhythm of external events, by adapting their phase and period, which corresponds to how we infer a metrical structure from events in the music. The attending rhythms are self-sustaining and can occur at multiple levels, tracking events with different periods simultaneously (Drake, Jones, et al., 2000; Large & Jones, 1999). These features correspond respectively to the stability of our metrical percept and the perception of multiple hierarchical levels of regularity (Large, 2008). As such, DAT can explain many aspects of beat and meter perception. Behavioral support for DAT comes from studies showing a processing advantage in metrically strong positions for temporal intervals (Large & Jones, 1999), pitch (Jones et al., 2002) and phonemes (Quené & Port, 2005). This is thought to be the result of the peaks in attentional energy associated with metrically salient positions.

At a neural level, beat and meter perception have been hypothesized to originate from neural oscillations that resonate to external events (*neural resonance*, see Large, 2008). This view on the perception of metrical structure can be seen as an extension of DAT and makes largely the same predictions. Like the attending rhythms in DAT, neural oscillations are suggested to be self-sustaining and are suggested to adapt their phase and period to an external rhythm. In addition to these features, neural oscillations may

arise at frequencies that are not in the stimulus, which may be an explanation for the phenomenon of subjective rhythmization (Large, 2008).

Snyder and Large (2005) provided some empirical evidence for the neural resonance theory, by showing that high frequency neural oscillations reflect rhythmic expectancy. They presented participants with a rhythm consisting of alternating loud and soft tones, while measuring their brain activity using electroencephalography (EEG). With this method it is possible to measure the electric activity of the brain with high temporal precision and thus, it is possible to show high frequency neuronal oscillations. The results showed that a peak in induced gamma oscillations (20–80 Hz) coincided with the sounds. When a loud sound was omitted, this peak was still present, which was interpreted as evidence that the induced activity represented the regular underlying beat, which continued even without physical input. Additional evidence in this line was provided by Fujioka et al. (2012), Iversen et al. (2009), and Zanto, Large, Fuchs, and Kelso (2005). In each of these studies, induced oscillatory activity was shown to relate to metrical expectations. The question remains, however, whether neural resonance is actively influencing rhythm perception or whether it is an emergent attribute of the EEG response induced by the rhythmic structure of the stimulus itself (Smith & Honing, 2008). Also, to date, support for neural resonance as an explanation for beat perception only comes from studies using isochronous stimuli. Whether neural resonance also explains phenomena such as subjective rhythmization and beat perception with more complex stimuli remains to be tested.

2.5.6 Metrical structure is perceived in motor areas of the brain

EEG provides excellent temporal resolution. However, to localize the networks involved in beat perception, the superior spatial resolution of functional magnetic resonance imaging (fMRI) is needed. The overall picture emerging from fMRI studies looking at beat perception is that of large involvement of the motor areas in the brain. Grahn and Brett (2007) examined beat perception using different rhythmic sequences, containing temporal accents (i.e. accents that arise from the structure of event onsets; cf. Povel & Essens, 1985). In some rhythms these accents were spaced evenly, while in other rhythms they were irregular. Rhythms with regular accents were considered to be metrical rhythms and rhythms with irregular accents non-metrical. Only metrical rhythms induced a beat, as was confirmed by a behavioral test. Using fMRI it could be shown that during listening to metrical rhythms the basal ganglia and the supplementary motor area (SMA) were more active than during listening to non-metrical rhythms, implicating these areas in beat perception. The findings of Grahn and Brett (2007) were confirmed by several subsequent studies showing activations not only in the basal ganglia and SMA, but also in the cerebellum and pre-motor areas (Bengtsson et al., 2009; Chen et al., 2008a; Grahn & Rowe, 2009). Importantly, activity in a network of motor areas was consistently observed, even when participants were asked not to make overt movements. This shows that these areas are involved when people just listen to a metrical rhythm (for a review on the neural correlates of beat and meter perception, see Grahn, 2009a, 2012).

Motor areas have been implicated in time perception in general. However, recently it was shown that specific networks are dedicated to perceiving absolute and relative

durations respectively. While a network comprising the cerebellum and the inferior olive is involved in absolute *duration-based* timing, a different network, including the basal ganglia and the SMA, is active for relative or *beat-based* timing (Teki et al., 2011). The perception of a beat, which requires the perception of temporal regularity, thus appears to be a distinct process from the general perception of temporal intervals. We will refer to this as the auditory timing *dissociation hypothesis* (see also Honing, Ladinig, Háden, & Winkler, 2009; Merchant & Honing, 2014).

To summarize, regular fluctuations in attentional energy and neural resonance have been suggested to explain the perception of metrical structure. Also, a role for a network of motor areas in the brain, including the basal ganglia and the SMA, has been implicated. Finally, a dissociation between rhythm perception and beat perception has been suggested.

Electroencephalography (EEG) is a neuroimaging method that is non-invasive and does not require an overt response from the participant. In addition, EEG has the temporal resolution to track the perception of a beat over time. Previously, beat perception has been examined using EEG with the traditional and well-studied approach of looking at *event-related potentials* (ERPs). In the remainder of this chapter we will focus on using auditory ERPs in probing beat perception.

2.6 Measuring beat perception with event-related potentials (ERPs)

2.6.1 Using ERPs to probe beat perception

ERPs are hypothesized to reflect the sensory and cognitive processing in the central nervous system associated with particular (auditory) events (Luck, 2005). ERPs are isolated from the EEG signal by averaging the signal in response to many trials containing the event of interest. Through this averaging procedure, any activity that is not time-locked to the event is averaged out, leaving the response specific to the event of interest: the ERP. While ERPs do not provide a direct functional association with the underlying neural processes, there are several advantages to the technique, such as the ability to record temporally fine-grained and covert responses not observable in behavior. Also, several ERP components have been well studied and documented. Some of these components, used in testing beat perception, are elicited with an *oddball paradigm*.

An auditory oddball paradigm consists of a regular sequence of stimuli (standards), in which infrequently a stimulus is changed (deviant) in some feature (e.g., pitch, intensity, etc.). The deviant stimulus thus violates a regularity that is established by the standard stimuli. Depending on the task of the subject a deviant stimulus elicits a series of ERP components reflecting different stages and mechanisms of processing. The *mismatch negativity* (MMN), which is a negative ERP component elicited between 100 and 200 ms after the deviant stimulus, reflects automatic deviance detection through a memory-template matching process (see Figure 2.2). The N2b is a component similar to the MMN in latency, polarity and function, but it is only elicited when

the deviant is attended and relevant to the task. At around 300 ms after the deviant stimulus, a positive component can occur, known as the P3a, which reflects attention switching and orientation towards the deviant stimulus. For task relevant deviants, this component can overlap with the slightly later P3b, reflecting match/mismatch with a working memory representation (S. H. Patel & Azzam, 2005; Polich, 2007). Finally, the reorientation negativity (RON; 400–600 ms) reflects switching back attention to the original task (Horváth, Winkler, & Bendixen, 2008). Several of these ERP components are known to index the magnitude of a regularity violation. A larger deviation from regularity yields a MMN, N2b, P3a and P3b with earlier latency and larger amplitude (Comerchero & Polich, 1999; Fitzgerald & Picton, 1983; Rinne et al., 2006; Schröger & Winkler, 1995). This property is exploited when probing beat perception with ERPs.

The general idea of using ERPs to probe beat perception is that an event on the beat is perceived differently from an event occurring not on the beat and thus that two physically identical events in different metrical positions should yield different brain responses. Moreover, because we perceive events on the beat as different from events off the beat, we also perceive deviants on the beat as different from deviants off the beat. An effect of metrical position on the ERP response to a deviant event is therefore interpreted as evidence for the presence of beat perception. In general, it is thought that deviant events on the beat are detected better than deviant events off the beat and thus that the former elicit earlier and larger amplitude ERP responses than the latter (Schwartze, Rothermich, Schmidt-Kassow, & Kotz, 2011).
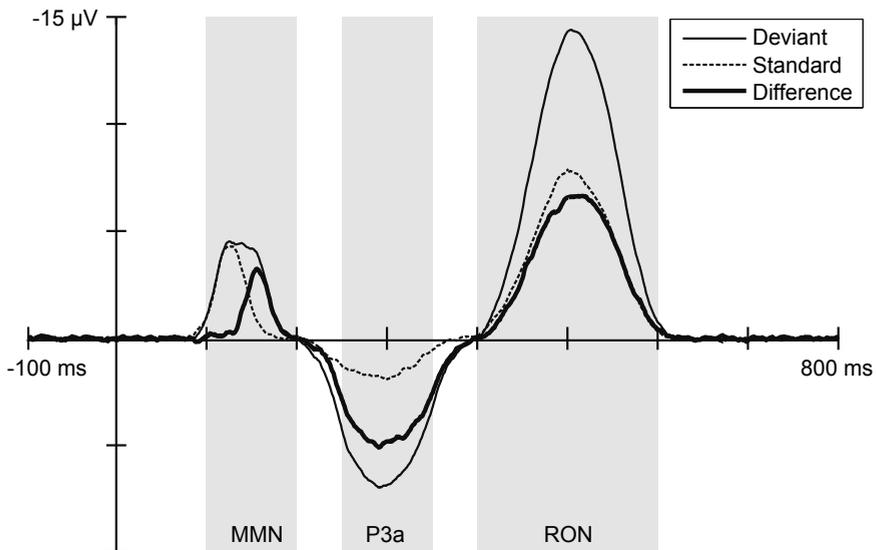

Figure 2.2 Idealized event-related potential (ERP) responses to unattended stimuli in an oddball paradigm, showing the standard (dotted line), deviant (solid line) and deviant minus standard difference waveform (bold line). The mismatch negativity (MMN), P3a and reorientation negativity (RON) components are highlighted with grey shading indicating standard latency windows.

An example of how deviant detection can show the presence of beat perception comes from studies examining subjective rhythmization (Brochard et al., 2003; Potter et al., 2009). In these studies, participants were presented with an isochronous series of tones. They were hypothesized to perceive the tones in odd positions as stronger than tones in even positions. Infrequently, a softer tone was introduced, either in odd or in even positions. These deviants elicited an N2b and a P3b. The P3b to deviants in odd positions had a larger amplitude than the P3b to deviants in even positions, showing that the deviants were indeed detected better – or perceived as more violating – on the beat. Other studies have shown that the P3b component to deviants is larger when the deviants occur in a regular sequence than when they occur in a sequence with random inter-onset intervals (Schmidt-Kassow & Kotz, 2009; Schwartze et al., 2011).

While the elicitation of an N2b and a P3b requires attention and a conscious effort towards detecting deviant stimuli, the MMN is automatic and mostly independent of attention. This makes the MMN an ideal ERP component to test beat perception when attention is directed away from a rhythm, provided that the MMN response is indeed sensitive to metrical structure and that beat perception can be shown to be *pre-attentive* in human adults. In the following sections, the MMN component and its relation to beat perception is discussed.

2.6.2 The Mismatch Negativity (MMN)

In general, the MMN is elicited when incoming sounds mismatch the neural representations of regularities extracted from the acoustic environment. Violations of the regularity in sound features such as pitch, duration or timbre can elicit an MMN (Winkler & Czigler, 2012; Winkler, 2007). Also violations of abstract rules (i.e. one auditory feature predicting another; Paavilainen, Arajärvi, & Takegata, 2007) or stimulus omissions (Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997) can cause an MMN. The MMN is regarded as a predictive process (Bendixen, SanMiguel, & Schröger, 2012) reflecting the detection of regularity-violations (for reviews see Kujala, Tervaniemi, & Schröger, 2007; Näätänen et al., 2007).

The processes underlying the MMN are thought to be automatic, however, the MMN can be modulated by attention (Haroush, Hochstein, & Deouell, 2010) and even be completely eliminated when deviations in attended and unattended auditory streams vie for feature specific processing resources (Sussman, 2007). The fact that MMN can be elicited even in comatose patients (Näätänen et al., 2007), sleeping newborns (Alho, Woods, Algazi, & Näätänen, 1992) and anesthetized animals (Csépe, Karmos, & Molnár, 1987) illustrates the relative independence from attention. The latency and amplitude of the MMN are sensitive to the relative magnitude of the regularity violation (Rinne et al., 2006; Schröger & Winkler, 1995) and correspond to discrimination performance in behavioral tasks (Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004). These properties can be exploited when, for example, beats on metrically strong and weak positions are compared or the relation between attention and beat perception is tested.

2.6.3 Using MMN to probe beat perception in human adults

To date there has been only a handful of studies that used MMN to study beat perception. The different methods in these studies have two common design goals: First, all studies present subjects with stimuli that induce a metrical structure and compare the responses to regularity violations occurring on different metrical positions (e.g. on the beat and off the beat). Second, all studies try to control attention to test whether the processes involved in differentiating between different metrical positions are automatic or dependent on attention (i.e., to study whether beat perception is pre-attentive; Bouwer, Van Zuijen, & Honing, 2014). The existing literature, however, contains inconsistent results (for a related review, see Grahn, 2009a).

Geiser, Ziegler, Jäncke, and Meyer (2009) presented subjects with rhythmic patterns containing temporal accents consistent with a regular 3/4 bar (e.g., the metrical structure of a waltz). In these metrically regular sequences infrequently a pitch deviant, a violation of the metrical structure, or a violation of the temporal surface structure of the rhythm was introduced. The meter violations consisted of the addition or removal of an eight note to the regular 3/4 bar. To create the rhythm violations, one or two eight notes were substituted by two or four sixteenth notes, leaving the metrical structure intact. Subjects had to either ignore the changes in the temporal domain and detect the pitch changes (unattended condition) or ignore the pitch changes and detect the temporal changes (attended condition). Regardless of subjects' musical training, rhythm violations elicited an MMN-like component in both attended and unattended conditions. Meter violations however only elicited an MMN-like component in the attended condition, implying that attention is required to induce a beat. In two experiments with similar attentional control, Vuust, Ostergaard, Pallesen, Bailey, and Roepstorff (2009) and Vuust et al. (2005) did find MMN responses to large temporal violations of the metrical structure regardless of musical training and attention. Unfortunately the large changes violated not only the meter but also other parameters, like the underlying temporal grid. As this in itself would lead to a MMN response, it is not clear from these results whether the MMN system is indeed sensitive to metrical structure.

A converse result comes from the experiment of Geiser, Sandmann, Jäncke, and Meyer (2010) who used identical regular 3/4 bar sequences as in their earlier study (Geiser et al., 2009). However, in this study deviants in the form of intensity accents were introduced at meter-congruous and meter-incongruous positions. The attention control was achieved in this experiment by asking subjects to attend to a silenced movie, a common procedure in many MMN experiments (Kujala et al., 2007). Geiser et al. (2010) found an enhanced MMN to accents in meter-incongruous positions for musicians and, to a lesser extent, for non-musicians, providing evidence in support of beat perception being pre-attentive. The conclusions drawn by this and the previous (Geiser et al., 2009) study are radically different, while identical beat inducing stimuli were used. As such, these studies very clearly show how large the influence of different attentional controls and experimental design on the results can be.
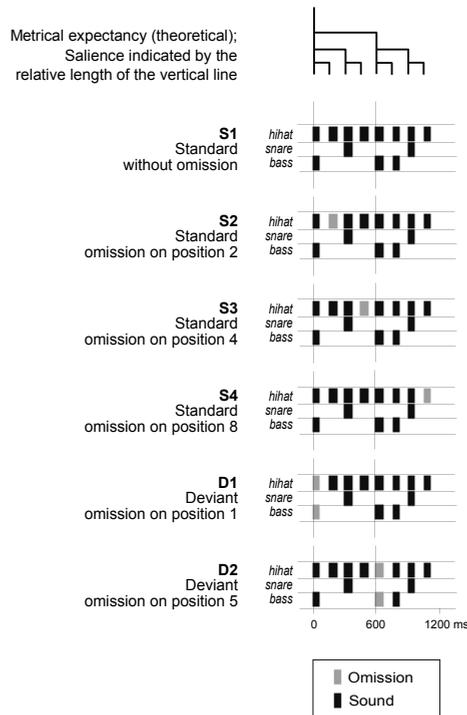
Figure 2.3 Stimuli as used in several studies on beat and meter perception (Honing et al., 2012; Ladinig et al., 2009; Winkler et al., 2009). S1–S4 are the standards and D1 and D2 the deviants used in an oddball paradigm. The different percussion sounds are marked as hi-hat, snare and bass.

Ladinig et al. (2009, 2011) took a somewhat different approach to meter perception in a study where they compared the responses of musically untrained subjects to omissions of tones with two different levels of metrical salience in a rock drum pattern (see Figure 2.3). Two different levels of attention control were employed. In the *passive* condition subjects were attending to a silent movie, as in Geiser et al. (2010). In the *unattended* condition subjects were attending to intensity changes in a continuous stream of white noise. The latter condition was designed to be a strict control for attention as it required attention in the same modality, but for a different auditory stream. Results showed that the MMN responses elicited by infrequent omissions on the first beat (deviant D1; large violation of the metrical structure) and the second beat (deviant D2; smaller violation of the metrical structure) differed in latency but not in amplitude. The latency difference indicates faster processing for the larger metric violation, suggesting that the metrical structure was picked up without attention.

Studying pre-attentive beat perception using the MMN is not as straightforward as one might like. Most notably, the use of acoustically rich stimuli (with potential differences between sounds in different metrical positions) may interfere in unforeseen ways with the ERP results (cf. Bouwer et al., 2014). One possible future direction is to strive for

even more minimalistic paradigms and to test whether the auditory system automatically imposes structure to incoming unattended stimuli that have no apparent structure (e.g., isochronous sequences of the same sounds; subjective rhythmization). Alternatively, priming paradigms could be used that test how long externally imposed structure persists when the input is no longer structured. As the MMN responds not only to temporal but also to pitch and timbre deviants, it does allow studying more complex accent structures, a topic mostly ignored so far.

In summary, while the automatic nature of beat perception is not yet fully understood, MMN seems to be a promising candidate for measuring beat perception.

2.7   Discussion and conclusion

In this chapter we have seen that the perception of metrical structure seems specific to the domain of music and is shared with only a limited number of non-human animals. Nonetheless, this ability seems very basic to humans. People readily synchronize to a beat in a wide variety of settings, like concerts, demonstrations, when marching and when singing a song together. This apparent contradiction between the ease with which we are capable of hearing a beat and the uniqueness of this skill raises several questions about how fundamental the perception of metrical structure really is.

We have shown how ERPs can be used to answer fundamental questions about beat perception. Measuring ERPs is relatively straightforward, it can be realized when a listener does not attend to a rhythm, and it is a well-researched method. However, several issues remain.

One of the challenges in examining beat perception is to balance the need for highly controlled stimuli with the aim to use stimuli that are ecologically valid. On the one hand, future research must address the role of different acoustic features in beat perception. Most research in this area has focused on temporal accents and has used either very simple or even isochronous sequences. While this is useful in controlling acoustic factors, it is not a very natural way of testing beat perception. In natural music, different types of accents often work together in shaping our metrical expectancies. The role of intensity accents, melodic accents and our previous experience has only been looked at very sparsely. However, using more natural stimuli can create problems in interpreting the results.

In natural music, a beat is induced by creating accents on the beat. Because accented sounds by definition need to stand out from non-accented sounds, this often means that tones on the beat have a different sound than tones that are not on the beat. When comparing the response to events on the beat and events off the beat, these sound differences need to be taken into account. An example of this problem can be found in the work of Winkler et al. (2009), who showed that newborn infants respond to the omission of a beat, but not to the omission of a sound off the beat. While these results showed that the newborns differentiated between sounds in different metrical positions, it cannot be completely ruled out that they did so on the basis of differences in sound rather than position. The sounds that were on the beat were composed of a bass drum and a hi-hat sound, while the sounds that were off the beat were composed of a

single hi-hat sound. This means it is possible that the newborns responded differently to the omission of different sounds. To exclude alternative explanations like these, stimuli must be designed in which physical differences between the sounds in different metrical positions cannot influence the results. Thus, balancing the design of ecologically valid stimuli with the experimental control needed to draw firm conclusions continues to be a challenge.

Another issue to be addressed in future research is the apparent gap between the sometimes contradicting results obtained with the different methods used in probing beat perception. Some consensus is emerging on which brain networks are involved in the perception of beat and meter and how brain dynamics might be accountable for our metrical expectations. However, the connection between these findings remains unclear. Also, studies to date have all used slightly different stimuli and tasks, which in some cases results in radically different or even contradicting conclusions (Geiser et al., 2009; Grahn, 2012; Ladinig et al., 2009). Once the different methods are used with similar paradigms, tasks and stimuli, it will be possible to directly compare the results and this will hopefully allow us to get a more coherent picture of the perception of beat and meter, and address its apparent innateness, domain- and species-specificity. All in all, this research will contribute to a better understanding of the fundamental role that beat and meter perception play in music.

Chapter 3

What makes a rhythm complex? The influence of musical training and accent type on beat perception[*]

Perception of a regular beat in music is inferred from different types of accents. For example, increases in loudness cause intensity accents, and the grouping of time intervals in a rhythm creates temporal accents. Accents are expected to occur on the beat: when accents are "missing" on the beat, the beat is more difficult to perceive. However, it is unclear whether the accents occurring *off* the beat alter beat perception similarly to missing accents on the beat. Moreover, no one has examined whether intensity accents influence beat perception in the same way as temporal accents, nor how musical expertise affects sensitivity to each type of accent. In two experiments, we measured complexity ratings for rhythms, with both temporal and intensity accents, that varied in the number of missing accents on the beat and the number of accents off the beat. In both experiments, musical expertise increased sensitivity to missing accents on the beat. In addition, listeners were more sensitive to missing accents on the beat for temporally-accented rhythms than intensity-accented rhythms. The effect of accents off the beat was weak and depended on both musical expertise and missing accents on the beat: lots of missing accents on the beat made beat perception very difficult, thus adding accents off the beat did not reduce beat perception further. Overall, the different types of accents were processed qualitatively differently, depending on musical expertise. These findings indicate the importance of designing ecologically valid stimuli for musical novices when testing beat perception.

3.1   Introduction

In musical rhythm, we often perceive a regular beat. The beat is what we tap our feet to, and the perception of a beat in music makes some musical events sound more prominent than others. To perceive a beat in a rhythm, we rely on various types of *accents* (Honing, Bouwer, & Háden, 2014; Lerdahl & Jackendoff, 1983a). An accent is an acoustic event that is more salient than its surrounding context. Salience can be caused by differences in pitch, intensity or timbre (Ellis & Jones, 2009), but it can also arise from variation in the temporal structure of a rhythm (Povel & Okkerman, 1981). When accents occur at regularly spaced points in time, a listener can perceive a beat in a rhythm (Lerdahl & Jackendoff, 1983a), and the beat generally coincides with accented events (Honing et al., 2014). Once a beat has been inferred from a rhythm, its perception remains stable (Large, 2008), and thereafter the beat can coincide with silence, or an accent can even occur off the beat, as in a syncopation (Lerdahl & Jackendoff, 1983a), without beat perception being too disrupted. The relationship between the structure of accents in music and the perceived beat is thus flexible, and as such, the perception of a beat is regarded as a psychological construct (Grahn, 2012; Honing, 2013; Large, 2008; Merchant et al., 2015). A beat is often embedded in a hierarchical organization with several nested levels of perceived regularity, the metrical structure. Within the metrical structure, the beat is the most salient level of regularity. The faster regularity at a hierarchically lower level than the beat is termed a *subdivision* of the beat, and the slower, higher-order regularity of more and less salient beats is referred to as *meter*.

A variety of stimuli have been used to study beat perception, ranging from isochronous sequences (Fujioka et al., 2012; Nozaradan et al., 2011; Potter et al., 2009; Schwartze et al., 2013, 2011) to rhythms with varying inter-onset intervals but identical sounds (Grahn & Brett, 2007; Grube & Griffiths, 2009; Kung et al., 2013), rhythms with varying acoustic properties but with identical inter-onset intervals (Bouwer et al., 2014; Bouwer, Werner, Knetemann, & Honing, 2016; Ellis & Jones, 2009; Repp, 2010; Vuust et al., 2005, 2009), and real music (Bolger et al., 2013; Tierney & Kraus, 2013, 2014). Stimuli may contain various types of accents that indicate the beat to a listener. In an isochronous sequence of tones, listeners may spontaneously hear a binary beat (e.g., 'tick-tock-tick-tock'), with alternating more and less salient tones (Brochard et al., 2003; Potter et al., 2009). In this case, there are no accents in the rhythm to indicate the binary structure, nor any accents that may disconfirm the beat. Accents can be created by varying the temporal structure or acoustic features of a rhythm, and the structure of such accents has been shown to contribute to beat perception (cf. Drake, Penel, & Bigand, 2000, intensity differences; Ellis & Jones, 2009, duration and pitch; Hannon et al., 2004, pitch; Povel & Essens, 1985, temporal structure). In real music, multiple types of accents determine the salience of rhythmic events (Jones & Pfordresher, 1997).

Although it is well established that different types of accents contribute to beat perception, it is unclear whether these different accents contribute to beat perception in differing ways (i.e., are some accents more influential than others, and if so, which?). It is also unknown whether mismatches between the accent structure and the perceived

beat are perceived similarly on the beat versus off the beat (i.e., does an unexpected missing accent on the beat have the same effect on the perception of the beat as an unexpected accent off the beat?). In the current study, we address these issues by examining the contributions to beat perception of two types of accents: *temporal accents* and *intensity accents*. In addition, we explore whether musical expertise affects how sensitive a listener is to the structure of accents in a rhythm.

Temporal accents arise from the structure of the time intervals between events (e.g., note onsets) that make up a rhythm. Rhythmic events are perceived as accented when they are isolated in time, the second of a group of two events, or the first or last of a group of three or more events (Povel & Okkerman, 1981). The relation between the perceived beat and the structure of temporal accents has been described by Povel and Essens (1985) with a complexity score, which is a weighted sum of all beats that do not contain an event and all beats that contain an event but are unaccented. The complexity score is thus a measure of counterevidence against a possible perceived beat and indicates how well a given rhythm fits with the perception of a certain beat. Many studies examining beat perception have used rhythms with temporal accents (hereafter: *temporal rhythms*), designed after the Povel and Essens model, and the relationship between the number of unaccented beats and difficulty in perceiving the beat is well established (Grahn & Brett, 2007; Grube & Griffiths, 2009; Kung et al., 2013; Shmulevich & Povel, 2000).

Contrary to *counterevidence on the beat* (i.e., silences or unaccented events on the beat), *counterevidence off the beat* (i.e., accents occurring between beats) is not taken into account by Povel and Essens (1985). This is in line with the dynamic attending theory (DAT; Large & Jones, 1999), a theory of beat perception that suggests that we are more sensitive to sensory input that coincides with the beat than to input that falls between beats. However, several studies have shown that unexpected intensity accents are more salient off the beat than on the beat (Abecasis, Brochard, Del Río, Dufour, & Ortiz, 2009; Bouwer & Honing, 2015; Geiser et al., 2010, 2009). Accents off the beat may be more salient than on the beat as they disrupt the regularity of the perceived beat. Similar to missing accents on the beat, accents off the beat can be interpreted as counterevidence against a perceived beat. DAT suggests that we are more sensitive to information on the beat than off the beat. However, the salience of intensity accents off the beat raises the question whether counterevidence off the beat may also contribute to beat perception, and whether temporal accents off the beat are as disruptive as intensity accents.

Unlike the relationship between missing temporal accents on the beat and the beat that is perceived, which has been described by the Povel and Essens (1985) model, the relationship between the structure of intensity accents and the beat that is perceived has not been formalized. Despite this lack of formal characterization, many studies have used rhythms with intensity accents (hereafter: *intensity rhythms*) to induce a beat (Bouwer et al., 2014, in press; Chen, Zatorre, & Penhune, 2006; Drake, Penel, et al., 2000; Geiser et al., 2009; Iversen et al., 2009), and models and theories of beat perception stress the importance of intensity accents (Jackendoff & Lerdahl, 2006; Large, 2000; Lerdahl & Jackendoff, 1983a).

In one study, Grahn and Rowe (2009) compared responses to temporal and intensity rhythms directly. They examined beat perception in musicians and non-musicians in response to both types of rhythms using behavioral methods and fMRI. The beat was rated to be more salient in intensity rhythms than in temporal rhythms. However, temporal rhythms elicited more activity than intensity rhythms in the supplementary motor area and the basal ganglia, two brain areas associated with beat perception (Grahn & Brett, 2007; Grahn, 2012; Merchant et al., 2015). Thus, listeners appeared to process temporal and intensity accents differently. In addition, musicians showed greater connectivity between premotor areas and auditory cortex than non-musicians while listening to temporal rhythms that contained a beat, but not while listening to intensity rhythms that contained a beat. Thus, in addition to general processing differences between temporal and intensity rhythms, musical training may selectively increase sensitivity to the beat in temporal, but not intensity rhythms.

Although beat perception develops spontaneously in humans (Merchant et al., 2015), individuals vary widely in their ability to extract a beat from musical rhythm (Grahn & McAuley, 2009; Grahn & Schuit, 2012). Some of this variability may result from musical training, which enhances beat perception abilities (Cameron & Grahn, 2014; Geiser et al., 2010; Vuust et al., 2005). Based on the fMRI findings described above, these musical training enhancements may depend on the type of accents present in the rhythm.

In the current study, we aimed to examine the contributions of different kinds of accents to beat perception in musical experts and musical novices. First, we compared the influence of temporal accents and intensity accents on beat perception. Second, we examined the effects of counterevidence both on the beat and off the beat. Finally, we looked at the influence of musical training. As in previous studies (Grahn & Brett, 2007; Grube & Griffiths, 2009; Kung et al., 2013), we constructed temporal rhythms with varying metrical complexity based on Povel and Essens (1985). However, contrary to previous studies, we not only manipulated how many beats were silent (counterevidence on the beat), but we also varied how many accents occurred off the beat (counterevidence off the beat). In addition, we constructed intensity rhythms that mirrored the temporal rhythms in terms of the number of unaccented beats and the number of accents off the beat.

In Experiment 1, using a web-based setup, we obtained ratings of beat perception difficulty for these rhythms from participants who varied in musical expertise. In Experiment 2, we validated the results from Experiment 1 using a second, more constrained set of rhythms. We expected that an increase in the amount of counterevidence, both on the beat and off the beat, would increase the difficulty of perceiving a beat, both for rhythms with temporal accents and rhythms with intensity accents. Based on Grahn and Rowe (2009), we expected musical training to selectively enhance the sensitivity to the structure of the accents in rhythms with temporal but not intensity accents. Finally, we expected intensity accents to be more salient than temporal accents, and thus to be more perturbing of beat perception than temporal accents when used as counterevidence off the beat.

3.2 Experiment 1

3.2.1 Methods

Participants

The data reported here was retrieved from the online application on February 6, 2015. At that time, a total of 91 people had viewed the start page of the online application for Experiment 1, of whom 78 people had provided consent, 72 had provided their age and years of musical training, 56 had finished reading the instructions, and 54 had listened to the examples. Finally, 48 participants had proceeded to rate one or more rhythms (for more details, see the Procedure section). To improve reliability, 16 participants who rated fewer than 60 rhythms were considered dropouts and were excluded. The dropout rate was thus 33 percent, which is comparable to previous online music cognition experiments (cf. Honing & Ladinig, 2009). The remaining 32 participants were on average 33.3 years old (range 18–66 years, SD = 14.5) and reported on average 11.1 years of musical training (range 0–25 years, SD = 8.3). The study was approved by the Ethics Committee of the Faculty of Humanities of the University of Amsterdam and the Non-Medical Research Ethics Board of the University of Western Ontario.

Stimuli

We generated all possible rhythms of 9 tones and 7 silences aligned to a grid of 16 positions, with the grid positions representing four beats subdivided into four sixteenth tones (see Figure 3.1). By using 16 grid points, which can be divided into groups of two or four, but not into groups of three, we reinforced the perception of a binary metrical structure (Povel & Essens, 1985). We selected a binary metrical structure because the beat is easier to perceive in binary than in ternary meters (Bergeson & Trehub, 2006). Positions 1, 5, 9 and 13 were considered to be on the beat. We assigned accents to events based on Povel and Essens (1985), with isolated events, the second of two consecutive events and the first and last of three or more consecutive events considered accented. Temporal rhythms were subsequently selected based on five constraints. First, only patterns that started with an event were considered. Second, in order to avoid unevenly distributed patterns, we allowed a maximum of five consecutive events and a maximum of three consecutive silences. Third, in order to avoid too much repetition in the rhythms, we only included rhythms in which the four sixteenth notes that made up each of the four beats (notes 1–4, 5–8, 9–12, and 13–16 for the four respective beats) contained a different configuration of events, Thus, rhythms in which multiple beats consisted of the same pattern (for example one eighth note and two sixteenth notes, repeated four times) were not included. Fourth, only patterns with six accented events were used. Finally, as was done previously (Grube & Griffiths, 2009), temporal rhythms with unaccented beats were excluded, allowing silence to be the only type of counterevidence on the beat.

For each rhythm, the number of missing beats and the number of accents off the beat were counted. As the first position always contained an event, the number of missing beats varied between 0 and 3. Although we designed the rhythms to be perceived as four beats subdivided into four sixteenth tones, it is possible to hear a rhythm consisting of 16 grid-points as eight beats subdivided into two eighth tones. We did not want

to exclude this possibility. Therefore, we regarded positions 3, 7, 11 and 15 as ambiguous and did not count evidence in these positions. The number of accents off the beat was thus counted as the number of accents in all even-numbered positions.

Intensity rhythms were constructed to be analogous to the temporal rhythms (see Figure 3.1). Each position on the grid was filled with a tone and intensity accents were introduced on the same positions where temporal accents occurred in the temporal rhythms. Thus, like the temporal rhythms, all intensity rhythms contained six accents. However, unlike the temporal rhythms, in the intensity rhythms a sound occurred on each subdivision of the beat. While the temporal rhythms contained three different event types (accented events, unaccented events and silences), the intensity rhythms only contained two different types (accented and unaccented events). The accented events were always in the same positions for the two types of rhythms, but unaccented events in the intensity rhythms could map onto either unaccented events or silences in the temporal rhythms. Thus, different temporal rhythms could map onto the same intensity rhythm. Therefore, while a total of 670 temporal rhythms adhered to our criteria, only 120 intensity rhythms were possible with the current constraints. Also, within the constraints concerning the total number of accents and events, some combinations of missing beats and accents off the beat were not possible and others were unlikely. For example, when three beats are missing, it is impossible to have nine accents that do not occur off the beat. To be able to test our hypotheses with several different rhythmic patterns per condition, we only included the ten conditions that allowed for six or more different rhythmic patterns (see Table 3.1).

An initial pilot showed that the rhythms were too short for people to make judgments about their metrical complexity. Therefore, for each condition, longer rhythms were constructed by concatenating pairs of different semi-randomly selected rhythms with the same number of missing beats and the same number of accents off the beat into rhythms of 32 grid-points. The randomization was optimized to create as much variety as possible in the rhythms. A final tone was appended to each rhythm to provide metrical closure (Grube & Griffiths, 2009). Figure 3.1 shows an example rhythm for each condition. Sound examples for these rhythms are available as Supplementary Material[1]. During the experiment, participants were specifically asked to detect a beat in the rhythms. Only one of the ten conditions contained strictly metric rhythms (i.e., without any counterevidence). The inclusion of counterevidence may make it hard to hear a beat, especially for musical novices. To prevent them from getting discouraged during the experiment, we did not include an equal number of rhythms from each condition in the experiment, but rather used a larger number of rhythms from the condition without counterevidence than from each condition with counterevidence. Table 3.1 shows the total number of rhythms used for each condition. Figure 3.1, Table 3.1 and the Supplementary Material have the same numbering for the ten conditions used.

---

[1] Supplementary Material for this chapter is available online at http://www.fleurbouwer.nl/publications.

Table 3.1 Characteristics of the rhythms used in Experiment 1. Conditions with fewer than 6 possible rhythms were not included in the experiment. The numbers in the rightmost column correspond to the numbering used in Figure 3.1 and the Supplementary Material.

Missing beats	Accents off the beat	Possible 16 grid-point rhythms		Number of concatenated 32 grid-point rhythms used in Experiment 1		Classification of accents off the beat	No
		Temporal	Intensity	Temporal	Intensity		
0	0	12	6	24	24	Few	1
1	0	36	6	15	15	Few	2
	1	98	18	15	15	Some	3
	2	56	6	15	15	Many	4
2	0	11	3	0	0	Not used	
	1	65	12	15	15	Few	5
	2	143	22	15	15	Some	6
	3	111	16	15	15	Many	7
	4	18	3	0	0	Not used	
3	2	10	2	0	0	Not used	
	3	37	8	12	12	Few	8
	4	53	12	12	12	Some	9
	5	20	6	10	10	Many	10
Total number of rhythms				296			

Temporal rhythms
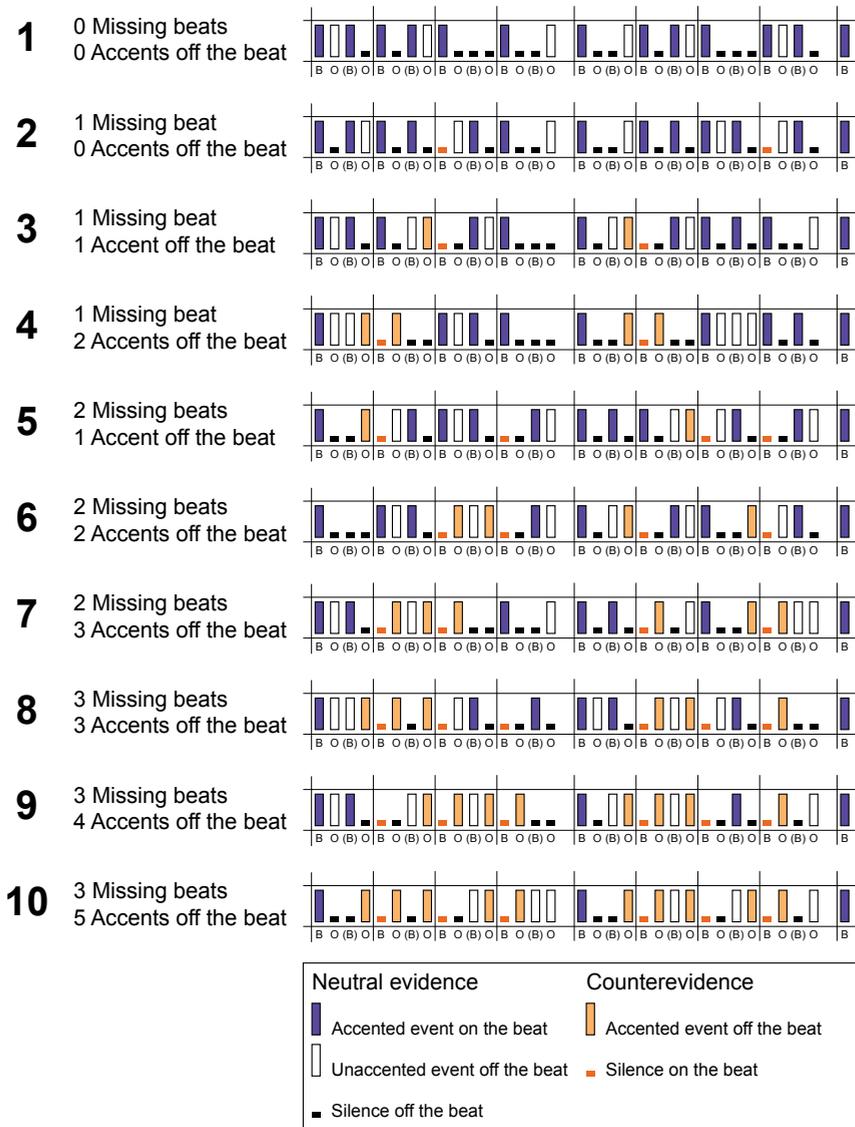
Figure 3.1 Examples of rhythms for each condition. Each rhythm as used in the experiment is constructed from two of the original 16 grid-point rhythms, followed by a final tone, for a total of 33 grid points. The spacing between the two halves of the rhythm and before the final tone is for viewing purposes only. In the concatenation of the rhythms, the isochronicity of the grid-points was preserved.

Intensity rhythms
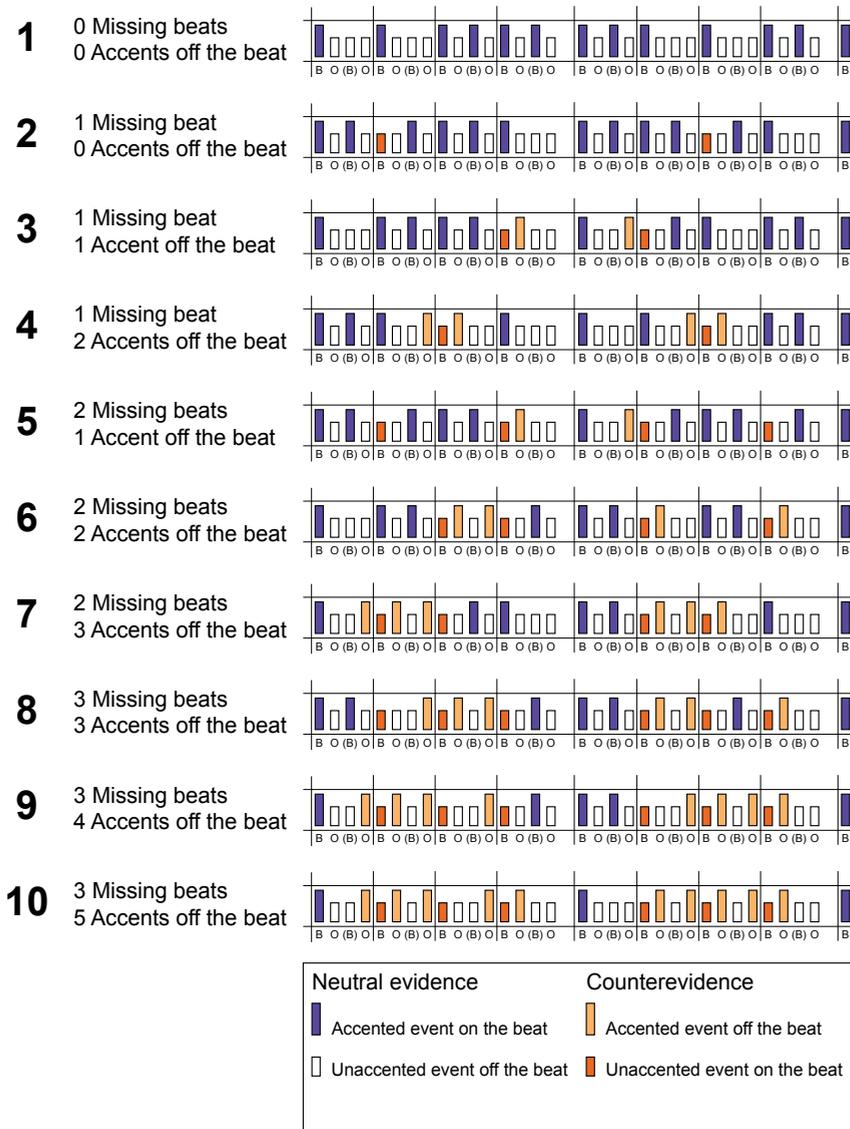
Figure 3.1 (continued) Note that the number of missing beats and number of accents off the beat refer to counterevidence in a rhythm of 16 grid-points. Sound examples for these rhythms are available as Supplementary Material. B = beat; O = off the beat; (B) = ambiguous (off the beat when subdivided into four beats of four sixteenth notes; on the beat when subdivided into eight beats of two eighth notes).

All sounds were woodblock sounds generated in Garageband (Apple Inc.). For the intensity rhythms, the difference between accented and unaccented events was set to 8.5 dB, comparable to the intensity rhythms in Grahn and Rowe (2009). Intensity and temporal rhythms were equated for overall loudness by scaling all sounds in the temporal rhythms to 0.8 dB softer than the accented sounds in the intensity rhythms. The inter-onset interval between grid points was varied to prevent carryover of the perceived beat from one trial to the next. A tempo of 100 beats per minute (inter-beat interval of 600 ms) is the optimal tempo for human adults to perceive a beat at (Drake, Jones, et al., 2000; London, 2012). Assuming a subdivision of the rhythms into beats of four sixteenth tones, this would correspond to an inter-onset interval of 150 ms between grid-points. We used five inter-onset intervals around this rate (140, 145, 150, 155 and 160 ms), corresponding to tempi of 107, 103, 100, 97 and 94 beats per minute.

Procedure

A web-based application to rate auditory stimuli was created using the Google App Engine (Google Inc.). To foster future research, we have released this application as open-source software at https://github.com/dodijk/annotate. For viewing purposes, the application can be accessed online at http://uvamcg.appspot.com. When accessing the website, participants were presented with four obligatory steps before the experiment started. First, they provided informed consent. Second, they provided their age in years and the number of years of formal musical training they had received in their life. Third, they were presented with a written explanation of the experiment. Finally, they were presented with example rhythms. Participants were asked to perform the experiment in a quiet environment and use a computer rather than a mobile device. They received an explanation of the term "beat" and were given the following instructions: "For each rhythm, we ask you to rate on a scale of 1–10 how hard you think it would be to tap along with the beat in that rhythm. Rate each rhythm by clicking on the stars." They were presented with examples of both temporal and intensity rhythms with no missing beats and no accents off the beat (e.g., strictly metric rhythms), which contained the caption "This is an example of a rhythm containing a clear beat, which sounds easy to tap along to. We expect people to give this rhythm 1 star." Examples of temporal and intensity rhythms with three beats missing and several accents off the beat were presented accompanied by the caption "This is an example of a rhythm NOT containing a clear beat, which sounds hard to tap along to. We expect people to give this rhythm 10 stars." Participants could listen to the examples as often as they liked. After listening to the examples, participants could continue with the experiment.

Figure 3.2 shows the interface used for the rating task. Participants were presented once with each rhythm, at a tempo randomly chosen from the five tempi used. After each rating, the application automatically continued with the next rhythm. Once loaded, each rhythm was preceded by 500 ms of silence to allow participants to focus on the start of the trial. After every 30 rhythms (about 5 minutes), a screen appeared indicating a break. Participants could continue the experiment at their own discretion.
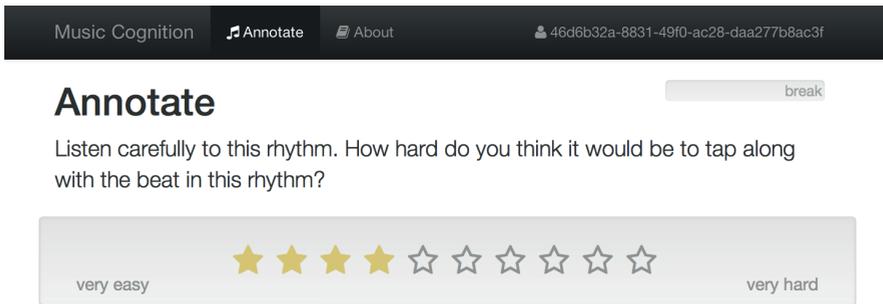
Figure 3.2 Example of the interface used during the online experiment.

Statistical analysis

In total, 5578 ratings were made. After excluding participants who rated less than 60 rhythms, 5297 ratings were included in the analysis. The distribution of the ratings is shown in Figure 3.3A. The distribution is skewed leftwards, indicating a bias for participants to provide low ratings. Such a distribution is often observed for Likert-items (Gardner & Martin, 2007). In general, responses on Likert items can be considered ordinal (Jamieson, 2004), especially when only one item is used (Carifio & Perla, 2008). Both the skewedness of the distribution and the ordinal nature of the responses prohibit the use of parametric statistical tests. Here, we thus used ordinal logistic regression, which provides a normalization of the ordinal data, for the analysis. The normalization of the raw ratings, which ranged from 1 (very easy) to 10 (very hard), is depicted in Figure 3.3B.

The normalized ratings served as the dependent variable in the regression model. Four independent variables were included: missing beats, accents off the beat, type, and musical training. *Missing beats* was defined as the number of beats that were silent (temporal rhythms) or unaccented (intensity rhythms) in each 16 grid-point rhythm (see Figure 3.1). The number of missing beats ranged from 0 to 3. The number of *accents off the beat* ranged from 0 to 5. With the constraints that we put on the rhythms, the absolute number of accents off the beat was strongly dependent on the absolute number of missing beats (for example, 5 accents off the beat could only occur when 3 beats were missing; see Table 3.1). To reduce problems with collinearity between missing beats and accents off the beat, we recoded the number of accents off the beat into three categories: few accents off the beat, some accents off the beat and many accents off the beat (see Table 3.1). *Type* of accents was either temporal or intensity. Finally, the number of years of *musical training* was included in the model as a continuous variable.

Although the categorization of accents off the beat eliminated some of its collinearity with missing beats, the two factors were not completely independent, as all rhythms with no missing beats by definition also had no accents off the beat. To take this into account, while polynomial contrasts were used for the conditions with 1, 2 or 3 beats missing, the condition with 0 missing beats was compared to the other three conditions
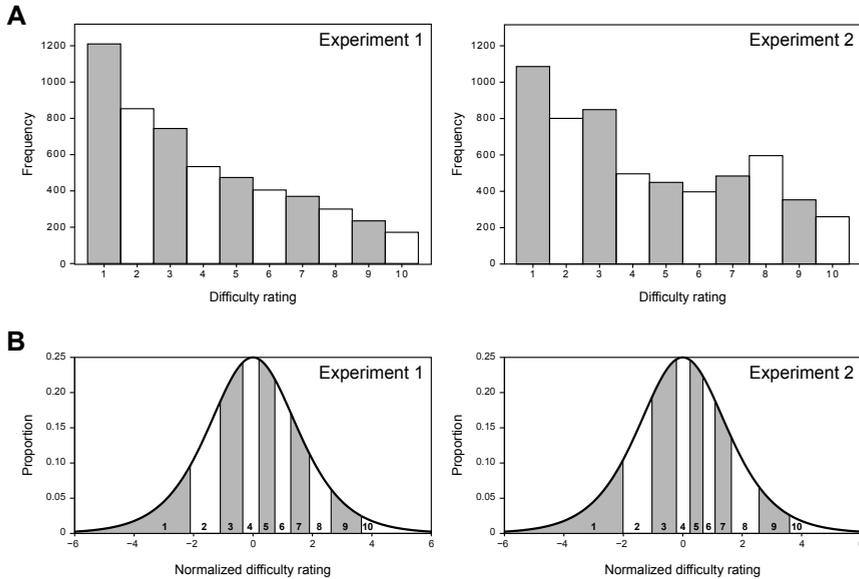
Figure 3.3 Distribution and normalization of ratings. A) Histograms of ratings from Experiment 1 and 2. B) Normalizations obtained with the ordinal regression for Experiment 1 and 2. The area under the curve for each rating corresponds to the proportion of responses for that rating.

using a Helmert contrast. Polynomial contrasts were also used for accents off the beat. To account for between-subject variation, a mixed model was used with a normally distributed random intercept for each participant. All main effects and interactions were then included in the model as fixed effects. The statistical analysis was conducted using R (R Development Core Team, 2008), and the clmm() function of the *ordinal* package (Christensen, 2015) was used.

3.2.2 Results

Figure 3.4 depicts the estimated normalized difficulty ratings for each condition. In the figure, estimates are given separately for participants with less than 2 years of musical training and participants with more than 2 years of musical training. Note that this is just for visualization purposes: In the model, we did not split the participants into two groups, but rather included musical training as a continuous variable. Table 3.2 contains the results of the ordinal regression.

A very small but significant interaction was found between missing beats and type ($X^2_{(3)} = 16.01$, $p = 0.001$, $\eta^2 = 0.003$; see Figure 3.5), showing that participants were more sensitive to missing beats in temporal than intensity rhythms. Planned contrasts showed that the linear association between the number of missing beats and the normalized difficulty was larger for temporal than for intensity rhythms ($z = 2.44$,
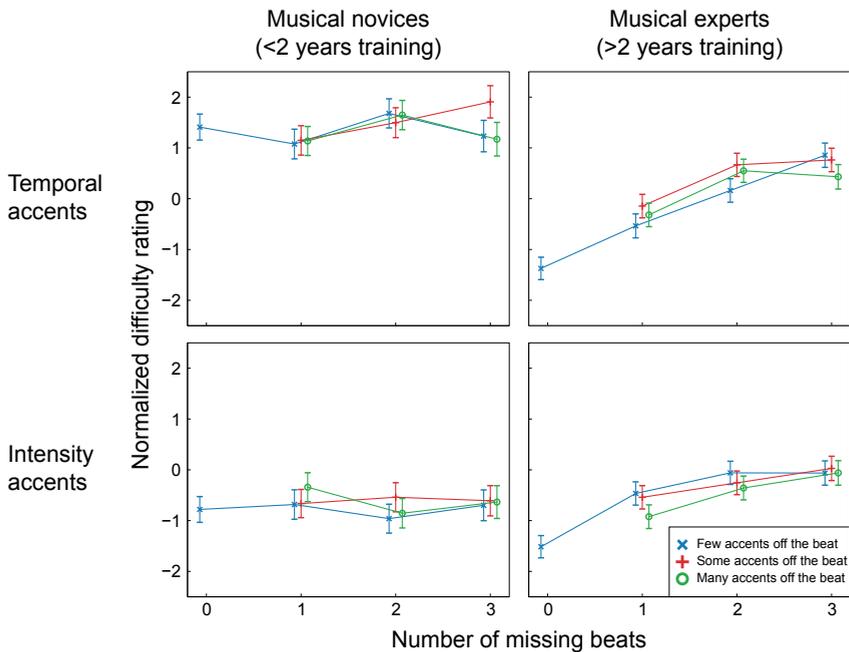
Figure 3.4 Estimated normalized ratings for all conditions in Experiment 1. For visualization purposes, estimates are given for participants with less than 2 years of musical training (musical novices) and participants with more than 2 years of musical training (musical experts). Note that in the model, musical training was included as a continuous variable. Error bars indicate 2 standard errors.

$p = 0.01$, $r = 0.03$). In addition, there was a larger negative quadratic association between number of missing beats and normalized difficulty in temporal than intensity rhythms, showing that for temporal rhythms the increase in difficulty associated with more missing beats showed some curvature, and was larger from 1 to 2 missing beats than from 2 to 3 missing beats ($z = 2.51$, $p = 0.01$, $r = 0.03$). For the Helmert contrast, comparing the difficulty of the rhythms with no beats missing and rhythms with 1 or more beats missing, the interaction with type was not significant. This suggests that the difference between rhythms without counterevidence (e.g., strictly metric rhythms) and rhythms with counterevidence was equally noticeable in temporal and intensity rhythms.

A significant interaction was also found between missing beats and musical training ($X^2_{(3)} = 81.33$, $p < 0.001$, $\eta^2 = 0.015$; see Figure 3.5), showing that musical training increased the sensitivity to missing beats. The linear association between the number of missing beats and the normalized difficulty became larger with more years of musical training ($z = 4.97$, $p < 0.001$, $r = 0.07$). The difference between normalized difficulty for rhythms with and without missing beats (the Helmert contrast) also became larger with more years of musical training ($z = 7.51$, $p < 0.001$, $r = 0.10$).

Table 3.2 Results of the ordinal regression in Experiment 1. The highest order significant effects for each factor are indicated in bold. *Significant at $p < 0.05$; **Significant at $p < 0.01$; ***Significant at $p < 0.001$. LR = Likelihood Ratio. df = degrees of freedom.

Main effects and interactions	LR	df	η^2	p
Missing beats	256.87	3	0.046	<0.001***
Accents off the beat	6.09	2	0.001	0.05*
Type	314.66	1	0.059	<0.001***
Musical training	5.55	1	0.001	0.02*
Missing beats * Accents off the beat	2.50	4	<0.001	0.64
Missing beats * Type	16.01	3	0.003	0.001**
Accents off the beat * Type	4.25	2	<0.001	0.12
Missing beats * Musical training	81.33	3	0.015	<0.001***
Accents off the beat * Musical training	0.01	2	<0.001	0.99
Type * Musical training	57.30	1	0.011	<0.001***
Missing beats * Accents off the beat * Type	8.26	4	0.002	0.08
Missing beats * Accents off the beat * Musical training	0.91	4	<0.001	0.92
Missing beats * Type * Musical training	2.71	3	<0.001	0.44
Accents off the beat * Type * Musical training	−0.22	2	<0.001	1.00
Missing beats * Accents off the beat * Type * Musical training	1.19	4	<0.001	0.88
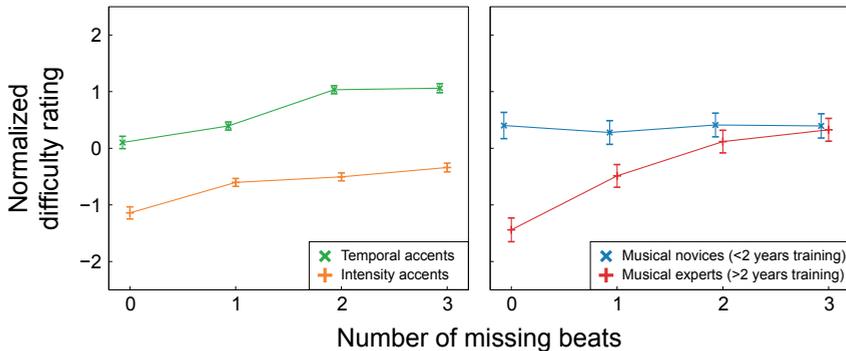
Figure 3.5 Interactions between beats missing and type and between beats missing and musical training in Experiment 1. Note: Error bars indicate 2 standard errors.

A third interaction was found between type and musical training ($X^2_{(1)} = 57.30$, $p < 0.001$, $\eta^2 = 0.011$). Participants with little musical training rated the intensity rhythms as easier than the temporal rhythms. This difference became smaller with more years of musical training ($z = 6.65$, $p < 0.001$, $r = 0.09$). Finally, a main effect was found for accents off the beat ($X^2_{(2)} = 6.09$, $p = 0.05$, $\eta^2 = 0.001$). However, none of the planned contrasts for this factor were significant, and the effect size for the main effect was extremely small.

3.3   Experiment 2

We controlled the rhythms in Experiment 1 for the number of events and accents and allowed a maximum of five consecutive events. However, because of the constraints we used, all temporal rhythms with no beats missing in fact had a maximum of three consecutive events, while in the other conditions, some rhythms could contain four consecutive events (see Figure 3.1). Thus, rhythms in different conditions differed slightly in the distribution of events, creating some rhythms that had a higher local event density. Event density in rhythm has been associated with beat salience and the urge to move to a rhythm (Madison, Gouyon, Ullén, & Hörnström, 2011) and may have thus influenced our ratings. Moreover, all temporal rhythms with no beats missing consisted of five sixteenth notes, two eighth notes, one dotted eighth note and one quarter note, while the distribution of intervals in the other rhythms was more varied. In Experiment 2, we aimed to validate the results from Experiment 1 using the same procedure, while controlling for the possible effects of event density by only including rhythms that had a maximum of three consecutive events. Differences in interval distribution were also controlled, by allowing only rhythms with the same interval distribution that occurred in the strictly metric rhythms (e.g., rhythms without any counterevidence).

Table 3.3 Characteristics of the rhythms used in Experiment 2. Two extra constraints were put on the temporal rhythms, to control for an uneven distribution of intervals and event density. Due to these extra controls, rhythms with one beat missing and rhythms with three beats missing and five accents off the beat were not used in Experiment 2. The numbers in the rightmost column correspond to the numbering used in Figure 3.1 and the Supplementary Material.

Missing beats	Accents off the beat	Possible 16 grid-point rhythms		Number of concatenated 32 grid-point rhythms used in Experiment 2		Classification of accents off the beat	No
		Temporal	Intensity	Temporal	Intensity		
0	0	12	6	28	28	Few	1
1	0	16	6	24	24	Few	2
	1	10	18	24	24	Some	3
	2	0	6	0	0	Many	
2	0	10	3	0	0	Not used	
	1	14	12	18	18	Few	5
	2	14	22	18	18	Some	6
	3	7	16	18	18	Many	7
	4	0	3	0	0	Not used	
3	2	4	2	0	0	Not used	
	3	7	8	10	10	Few	8
	4	9	12	8	8	Some	9
	5	0	6	0	0	Many	
Total number of rhythms					296		

3.3.1 Methods

Participants

We retrieved the data for Experiment 2 from the online application on February 6, 2015. At that time, 217 people had viewed the start page of the online application for Experiment 2, of whom 84 people had proceeded by providing consent and 67 had filled in their age and years of musical training. Among these, 53 people had read the instructions, 51 had listened to the examples, and 48 had rated one or more rhythms in the online application. There were 25 participants who had rated 60 or more rhythms and were thus included in the analysis, a 48-percent dropout rate. The remaining participants were on average 30.8 years old (range 20–69 years, SD = 11.8) and on average had had 7.0 years of musical training (range 0–25 years, SD = 6.4). The study was approved by the Ethics Committee of the Faculty of Humanities of the University of Amsterdam and the Non-Medical Research Ethics Board of the University of Western Ontario.

Stimuli

The stimuli were generated in exactly the same way as for Experiment 1, but with two extra constraints on the temporal rhythms: Only rhythms with no more than three consecutive events and only rhythms consisting of five sixteenth notes, two eighth notes, one dotted eighth note and one quarter note were included. With the extra constraints, some combinations of counterevidence in the temporal rhythms became impossible. The conditions with the combination of many accents off the beat and either one or three beats missing were thus excluded in Experiment 2. Table 3.3 shows the total possible rhythms within the constraints of Experiment 2 and the number of concatenated rhythms randomly constructed to use in the experiment. Note that all rhythms that were used in Experiment 2 could also have occurred in Experiment 1, but not all rhythms that were generated in Experiment 1 were allowed in Experiment 2. From the 296 randomly chosen rhythms in Experiment 2, 57 rhythms also occurred in Experiment 1 (19 percent).

Procedure and statistical analysis

The procedure and statistical analysis were identical to Experiment 1. Figure 3.3A shows the distribution of the data for Experiment 2 and Figure 3.3B shows the normalization obtained with the ordinal regression.

3.3.2 Results

The estimated normalized difficulty ratings for each condition are shown in Figure 3.6 and the results of the ordinal regression can be found in Table 3.4. For viewing purposes, the results are depicted separately for musical novices (<2 years of musical training) and musical experts (>2 years of musical training). In the model, musical training was included as a continuous variable. As in Experiment 1, a small but significant interaction was observed between missing beats and type ($X^2_{(3)} = 10.52$, $p = 0.01$, $\eta^2 = 0.002$; see Figure 3.7), showing that participants were more sensitive to missing beats in temporal than intensity rhythms. The linear association between the number of missing beats and the normalized difficulty rating was larger for temporal than in-

Table 3.4 Results of the ordinal regression in Experiment 2. The highest order significant effects for each factor are indicated in bold. *Significant at $p < 0.05$; **Significant at $p < 0.01$; ***Significant at $p < 0.001$. LR = Likelihood Ratio. df = degrees of freedom.

Main effects and interactions	LR	df	η^2	p
Missing beats	236.75	3	0.039	<0.001***
Accents off the beat	5.58	2	0.011	0.06
Type	64.22	1	0.011	<0.001***
Musical training	0.17	1	<0.001	0.68
Missing beats * Accents off the beat	0.43	2	<0.001	0.81
Missing beats * Type	10.52	3	0.002	0.01*
Accents off the beat * Type	5.46	2	<0.001	0.07
Missing beats * Musical training	55.73	3	0.010	<0.001***
Accents off the beat * Musical training	0.84	2	<0.001	0.66
Type * Musical training	1.99	1	<0.001	0.16
Missing beats * Accents off the beat * Type	0.24	2	<0.001	0.89
Missing beats * Accents off the beat * Musical training	7.10	2	0.001	0.03*
Missing beats * Type * Musical training	3.39	3	<0.001	0.34
Accents off the beat * Type * Musical training	0.73	2	<0.001	0.69
Missing beats * Accents off the beat * Type * Musical training	0.83	2	<0.001	0.66

tensity rhythms ($z = 2.63$, $p = 0.01$, $r = 0.04$). As in Experiment 1, the interaction between the Helmert contrast and type was not significant, showing that participants differentiated between rhythms with no beats missing and rhythms with one or more beats missing equally well in the temporal and intensity rhythms.

A three-way interaction was observed between missing beats, accents off the beat and musical training ($X^2_{(2)} = 7.10$, $p = 0.03$, $\eta^2 = 0.001$). The interaction between missing beats and accents off the beat increased with more years of musical training ($z = 2.12$, $p = 0.03$, $r = 0.03$). Musical novices rated rhythms with some accents off the beat as slightly more difficult than those with few accents off the beat regardless of the number of beats missing. However, musical experts rated rhythms with some accents off the beat as more difficult than those with few accents off the beat only when one beat was missing, but not when two or three beats were missing. With more years of musical training, the interaction between accents off the beat and beats missing became more pronounced and the effect of accents off the beat even reversed in the conditions with three beats missing, with higher difficulty ratings for rhythms with few accents off the beat than for rhythms with some accents off the beat.

Although this three-way interaction was significant, its effect size was very small. Therefore, we also looked at the two-way interaction between missing beats and mu-
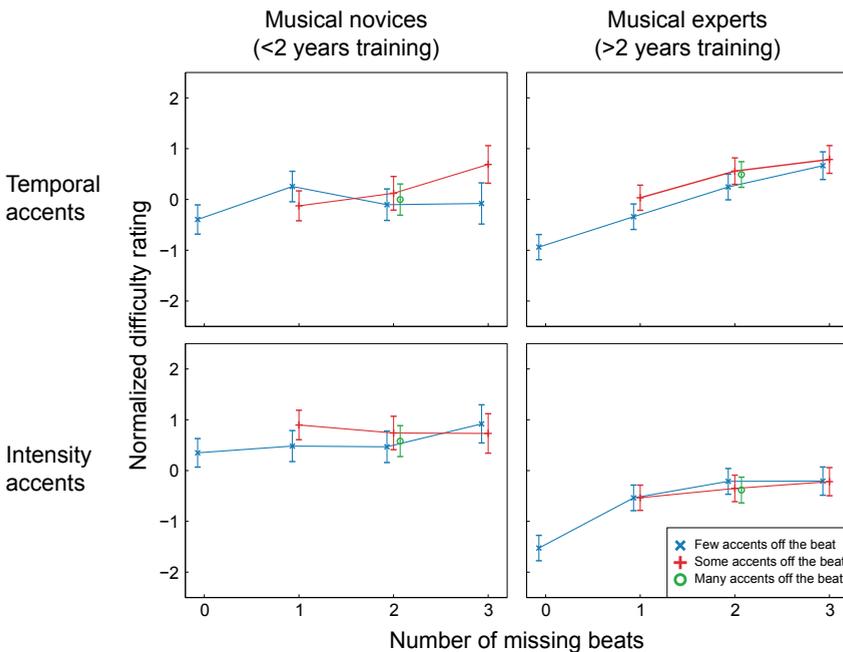
Figure 3.6 Estimated normalized ratings for all conditions in Experiment 2. For visualization purposes, estimates are given for participants with less than 2 years of musical training (musical novices) and participants with more than 2 years of musical training (musical experts). Note that in the model, musical training was included as a continuous variable. Error bars indicate 2 standard errors.
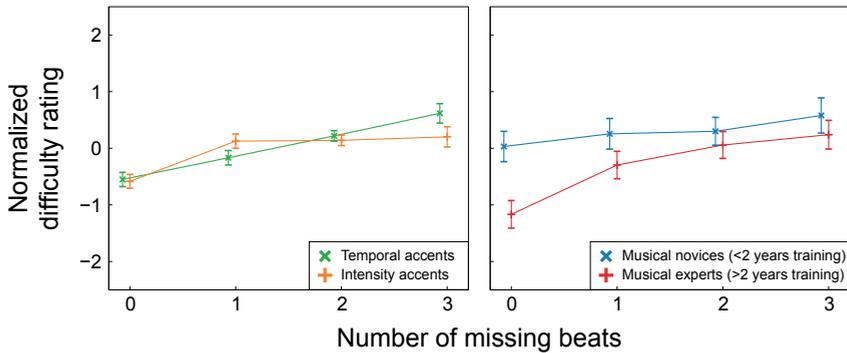
Figure 3.7 Interactions between beats missing and type and between beats missing and musical training in Experiment 2. Note: Error bars indicate 2 standard errors.

sical training ($X^2_{(3)} = 55.73$, $p < 0.001$, $\eta^2 = 0.010$, see Figure 3.7), to compare the results in Experiment 2 to those found in Experiment 1. As in Experiment 1, musical experts differentiated more strongly between rhythms with and without missing beats than musical novices ($z = 6.95$, $p < 0.001$, $r = 0.09$). Contrary to Experiment 1, in Experiment 2 the interaction between type and musical training did not reach significance.

3.4   Discussion

In this study we explored how different types of accents in musical rhythm influence the ease with which listeners with varying musical expertise infer a beat from a rhythm. Both in Experiment 1 and Experiment 2, musical training increased the sensitivity of participants to counterevidence on the beat (e.g., missing beats). For example, musical novices (those with less than two years of training) appeared to be insensitive to the number of beats missing. Contrary to our expectations, this greater sensitivity in musical experts was not selective to temporal rhythms, but also existed for intensity rhythms. Although musical training is not thought to be necessary for beat perception to develop (Bouwer et al., 2014; Merchant et al., 2015), training does seem to affect how a listener processes the structure of accents that indicates where the beat is.

In many previous studies using stimuli designed after Povel and Essens (1985), the effect of musical training on the detection of a beat was not reported (Chapin et al., 2010; Grube & Griffiths, 2009; Povel & Essens, 1985) or only musicians were tested (Kung et al., 2013). Grahn and Brett (2007) did examine the effect of musical training on the detection of a beat in temporal rhythms and did not find significant differences between musicians and non-musicians. However, they used a discrimination task, which implicitly probed beat perception. In a similar study, in which participants rated beat presence, differences were found between musicians and non-musicians (Grahn & Rowe, 2009). That rating task strongly resembled the current task, as it required an explicit rating. Thus, musical novices may be capable of detecting a beat just as well as musical experts but may have less explicit access to the information required to

make a rating of beat presence. In line with this, other work has shown that musical training enhances beat perception only when people attend to rhythm, but not when they ignore it (Bouwer et al., in press). As such, some aspects of beat perception may be more automatic, and independent of musical training, while aspects of beat perception that are related to attention and awareness may be enhanced by training. Future studies could examine potential differences between beat perception and beat awareness in musical novices and experts.

In both Experiment 1 and 2, participants were more sensitive to counterevidence on the beat (missing beats) in temporal than in intensity rhythms. The effect size of this interaction was extremely small, which warrants some caution in interpreting its practical use. Nonetheless, the interaction was highly significant in both experiments, with independent participants, and as such, seems reliable. The greater the number of beats missing in a rhythm, the more difficulty participants reported in finding a beat. This effect was larger for temporal than intensity rhythms when one or more beats were missing (as tested with the polynomial contrast). However, participants differentiated between rhythms with no beats missing and rhythms with one or more beats missing (as tested with the Helmert contrast) equally well for both types of rhythms. Although listeners did differentiate between intensity rhythms that were strictly metric (e.g., did not contain any counterevidence) and intensity rhythms that contained some syncopation (e.g., some counterevidence), they did not differentiate between different degrees of syncopation in the intensity rhythms. This may indicate that the Povel and Essens (1985) model cannot be translated completely to rhythms with intensity accents. As these types of accents are commonly used in real music, studies of beat perception with only temporal rhythms may not provide a full picture of the mechanisms of beat perception in music. Grahn and Rowe (2009) found that the brain networks involved in beat perception differed between intensity rhythms and temporal rhythms, and in the current study responses to the two types of rhythms were qualitatively different. More research is needed to understand how a beat is induced by music, where acoustic information as well as temporal cues are important.

In Experiment 1, musical novices, as expected, rated temporal rhythms as more difficult than intensity rhythms. This effect was generalized over all rhythms and was not modified by the amount of counterevidence. Musicians are more sensitive to the grouping rules that indicate temporal accents than non-musicians (Kung, Tzeng, Hung, & Wu, 2011). Thus, musical novices may have found it more difficult to extract information from the temporal rhythms than musical experts. In addition, musical novices attend more to lower (faster) levels of regularity in a metrical structure than musical experts (Drake, Jones, et al., 2000). In the intensity rhythms, all subdivisions of the beat contained a sound, creating an explicit isochronous pattern at a faster rate than the beat. Musical novices may have focused on this lower level of regularity in judging how easy it was to hear a beat and may have ignored the accents altogether at the hierarchically higher level of the beat, whereas musical experts may have been more attuned to events at all levels of the metrical hierarchy. The interaction between type and musical training, however, was absent in Experiment 2. In the more restricted set of rhythms used in Experiment 2, the variability in the temporal rhythms was less than in Experiment 1, as we controlled for event density and the distribution of the temporal

intervals used. The temporal rhythms in Experiment 2 were therefore more similar to each other than in Experiment 1, and this may have allowed participants to learn to recognize the intervals that were used. This may have made it generally easier for the musical novices to understand the grouping structure of the rhythm and may have therefore eliminated the difference between the two types of rhythms.

The effects of accents off the beat were not consistent over the two experiments, with a main effect in Experiment 1 and an interaction between accents off the beat, missing beats and musical training in Experiment 2. In both experiments, the effect sizes for the influence of accents off the beat were extremely small. This is in line with Dynamic Attending Theory, which predicts more attentional resources on the beat and less detailed processing off the beat (Large & Jones, 1999). However, the weak results for counterevidence off the beat may also have been due to the design of the experiment. The difficulty ratings made by musical experts for temporal rhythms do show a numerical trend in the expected direction, with higher difficulty ratings for rhythms with more counterevidence off the beat. This effect weakens when rhythms become very complex (e.g., when 3 beats are missing). The effects of accents off the beat thus seem to be present only for musical experts, and only for rhythms with little counterevidence on the beat, hence the three-way interaction between accents off the beat, missing beats and musical training in Experiment 2. As the effect of accents off the beat thus is present only in a small subset of the total rhythms, the experiments may have lacked the power to detect the effects of counterevidence off the beat consistently.

The interaction between counterevidence on the beat and off the beat in musical experts can be explained in two ways. First, it is possible that listeners do not differentiate between rhythms once it becomes too difficult to infer a beat. Thus, when three beats are missing, no beat is induced, and any further counterevidence created by accents off the beat cannot reduce beat induction any further. This ceiling effect may also explain the slight curvature in the effect of missing beats. While the difference between no counterevidence at all and some counterevidence is large, once it becomes harder to infer a beat, it does not matter whether more counterevidence is added.

A second explanation for the interaction between counterevidence on the beat and off the beat may be that instead of perceiving a rhythm as more complex, people may shift the phase of the beat when too much counterevidence is present. The perception of a beat unfolds over time (Grahn & Rowe, 2013). In rhythms with a lot of counterevidence (i.e., many silent beats and many accents off the beat), some sections may have contained accents off the beat that were regularly spaced (see Figure 3.1). Locally, a listener could phase-shift the beat to make the rhythm appear less complex, and this may have been easier for musical experts than musical novices. While the effects of accents off the beat were extremely small in our study, the possibility of local phase shifts may be worth considering in stimulus design. If only the number of missing beats is taken into account, beat perception in rhythms that are regarded as very complex (cf. Chapin et al., 2010) may in fact be very easy when accents off the beat allow for phase shifting of the beat.

Two caveats in our stimulus design must be noted. First, the difference between temporal and intensity rhythms in our study can be characterized not only by the nature of the accents, but also by the presence of marked subdivisions in the rhythms. In the intensity rhythms all subdivisions of the beat contained a sound, while in the temporal rhythms some subdivisions were silent. When all subdivisions are marked, which is often the case in real music, people may rely less on accents indicating the beat and instead may infer a duple meter from the isochronous subdivisions themselves (cf. Brochard et al., 2003; Potter et al., 2009). This may explain why the effects of counterevidence in the current study were larger for temporal than intensity rhythms. One way of resolving this issue is by filling all silences in the temporal rhythms with sounds that are softer than the events that indicate the rhythmic pattern. Previously, Kung et al. (2011) used such rhythms, but responses to these have not been compared to responses to temporal rhythms that do not contain all subdivisions. It is not clear whether the extraction of accents from temporal rhythms as proposed by Povel and Essens (1985) and used in the current experiment is the same as when all subdivisions are marked. This issue may be addressed in future research.

Second, we did not equate the different types of accents in terms of salience. However, it has been proposed that the subjective accents perceived in temporal patterns have an imagined magnitude of around 4 dB (Povel & Okkerman, 1981). The accents in the intensity rhythms were much larger (8.5 dB). Nonetheless, participants were more sensitive to the structure of the accents in the temporal rhythms than in the intensity rhythms. Thus, a discrepancy in salience between temporal and intensity accents would have led to an underestimation of this effect and is unlikely to have caused the effect.

3.5 Conclusion

In the current study, we have explored how the structure of different types of accents in rhythm influences the perception of a regular beat. Contrary to our expectations, both musical novices and musical experts were more sensitive to the structure of temporal accents than to the structure of intensity accents. As expected, musical training increased the sensitivity to the accent structure. The large effects of musical training on the perception of the beat may suggest that the use of stimuli with temporal accents in which the complexity is manipulated by varying the number of missing beats, as is often done, may not be meaningful to musical novices. The intensity accents as implemented in the current study did not improve beat perception for musical novices. However, a different combination of accents may be more suited to their beat perception capacities. The use of non-temporal information in beat perception is not well understood and may be important to better understand this ability.

We show here that it is possible to get meaningful data on beat perception by using an online experiment. One could easily use such a set-up to obtain data from a much larger group of people (for example through services like Amazon Turk). Ideally, this could result in a detailed model of how listeners with different backgrounds and experiences deal with different types of accents in rhythm. Our experiment provides a starting point in the search for stimulus material that is more ecologically valid, incorporates more musically relevant features, retains experimental control and has been tested in people varying in musical expertise and cultural background.

Chapter 4

Temporal attending and prediction influence the perception of metrical rhythm: evidence from reaction times and ERPs[*]

The processing of rhythmic events in music is influenced by the induced metrical structure. Two mechanisms underlying this may be *temporal attending* and *temporal prediction*. Temporal fluctuations in attentional resources may influence the processing of rhythmic events by heightening sensitivity at metrically strong positions. Temporal predictions may attenuate responses to events that are highly expected within a metrical structure. In the current study we aimed to disentangle these two mechanisms by examining responses to unexpected sounds, using intensity increments and decrements as deviants. Temporal attending was hypothesized to lead to better detection of deviants in metrically strong (on the beat) than weak (offbeat) positions due to heightened sensitivity on the beat. Temporal prediction was hypothesized to lead to best detection of increments in offbeat positions and decrements on the beat, as they would be most unexpected in these positions. We used a speeded detection task to measure detectability of the deviants under attended conditions (Experiment 1). Under unattended conditions (Experiment 2), we used EEG to measure the mismatch negativity (MMN), an ERP component known to index the detectability of unexpected auditory events. Furthermore, we examined the amplitude of the auditory evoked P1 and N1 responses, which are known to be sensitive to both attention and prediction. We found better detection of small increments in offbeat positions than on the beat, consistent with the influence of temporal prediction (Experiment 1). In addition, we found faster detection of large increments on the beat as opposed to offbeat (Experiment 1), and larger amplitude P1 responses on the beat as compared to offbeat, both in support of temporal attending (Experiment 2). As such, we showed that both temporal attending and temporal prediction shape our processing of metrical rhythm.

4.1   Introduction

In musical rhythm, we often perceive hierarchically organized regular salient moments in time, in the form of a metrical structure. The most salient level of a metrical structure is the beat or pulse. This is the regularity we usually tap and dance to. In addition, we can hear higher-level regularity, termed meter, in the form of alternating strong and weak beats. Metrical saliency often coincides with acoustic saliency in the form of an accent, but the relationship between the acoustic properties of music and the perceived metrical structure is not per se fixed (Honing, 2013; Large, 2008). When presented with an isochronous sequence of identical sounds, people perceive a pattern of alternating strong and weak tones, suggesting they induce a binary metrical structure from a rhythm that does not explicitly contain such a binary structure (Abecasis, Brochard, Granot, & Drake, 2005; Brochard et al., 2003; Potter et al., 2009). This phenomenon, known as subjective rhythmization or subjective accenting, is also a clear example of how a perceived metrical structure can influence the processing of rhythmic events. When listening to a rhythm with identical acoustic events, events in metrically strong positions (on the beat) can be perceived as louder than events in weaker positions (off-beat), even though all events are acoustically identical (Repp, 2010). In addition, a perceived metrical structure causes sound events to be more expected at metrically strong positions than at metrically weak positions (Ladinig et al., 2009). Two possible mechanisms underlying the influence of a perceived metrical structure on the processing of rhythmic events are temporal attending and temporal prediction.

The first mechanism, *temporal attending*[2], is described by the Dynamic Attending Theory (DAT), a prominent theory of the perception of metrical structure. According to DAT, the perception of metrical structure is the result of regular dynamic fluctuations in attentional resources, peaking at metrically strong positions (Jones, 2009; Large & Jones, 1999). Entrainment of neural oscillations to regular rhythmic events has been suggested to underlie these fluctuations in attentional resources (Large, 2008). The availability of more resources at metrically strong positions is thought to cause a general heightened sensitivity for events at those positions. This heightened sensitivity on the beat is supported by studies looking at processing of temporal deviations (Large & Jones, 1999), pitch (Jones et al., 2002), and speech sounds (Quené & Port, 2005). In addition, electrophysiological studies using oddball paradigms have shown larger event-related potentials (ERPs) to unexpected silences or intensity decrements in metrically strong positions than in metrically weak positions (Bouwer et al., 2014; Potter et al., 2009).

---

[2] The use of the term attention in the context of beat perception can lead to confusion (Henry & Herrmann, 2014) as it denotes both the general attentional resources available, usually manipulated by task-relevance of a rhythm and independent from the metrical structure, and the local fluctuations in attentional resources, which, according to DAT, depend on the metrical structure. Here, for the latter we will use the term temporal attending to differentiate it from the general use of the term attention.

Table 4.1 Hypothesized effects of temporal attending and prediction on the detection of intensity increments and decrements in different metrical positions. "+" indicates relatively improved detection; "−" indicates relatively impoverished detection.

Mecha-nism	Hypothesis	Predicted experimental effect			
		Increments		Decrements	
		On the beat	Offbeat	On the beat	Offbeat
Temporal attending	Heightened sensitivity on the beat	+	−	+	−
Temporal prediction	Better processing of events with large prediction error	−	+	+	−

Temporal fluctuations in attentional resources can also explain the occurrence of subjective accents in metrically strong positions. Attention has been proposed to enhance early sensory responses to sound (Lange, 2013). Electrophysiological studies show that auditory evoked potentials (AEPs) are enhanced for events in metrically strong positions as compared to for events in metrically weaker positions (Abecasis et al., 2009; Iversen et al., 2009; Schaefer, Vlek, & Desain, 2010; Tierney & Kraus, 2013). This is in line with more attentional resources being available in metrically strong positions than in metrically weak positions due to temporal attending and may cause events in metrically strong positions to be perceived as subjectively accented.

Recently, Vuust and Witek (2014) have proposed an alternative view on the perception of metrical rhythm, which emphasizes the importance of *temporal prediction*. They suggest that the perception of metrical rhythm can be explained within the framework of predictive coding (Clark, 2013). A metrical structure provides predictions about upcoming events and the degree to which these predictions are met provides a prediction error, which is used to update the perceived metrical structure. Like in DAT, within the framework of predictive coding, the perception of metrical rhythm is thought to be an interplay of top-down, endogenously driven, and bottom-up, exogenously driven processes (Vuust & Witek, 2014). However, the nature of the top-down, endogenous process differs between these two theories, with predictive coding stressing temporal prediction instead of temporal attending, which leads to different hypotheses about the influence of the metrical structure on the processing of rhythmic events (see Table 4.1).

First, DAT predicts better detection of unexpected events in metrically strong than weak positions, due to heightened sensitivity at metrically strong positions. However, loud sounds are more expected in metrically strong positions than in metrically weak positions. As such, the prediction error for an unexpected intensity increment in a metrically weak position is likely bigger than for an unexpected intensity increment in a metrically strong position, which predicts better detection of the former than the latter. Thus, while temporal attending would lead to enhanced processing of any event in a metrically strong position, temporal prediction would lead to enhanced processing of

metrically unpredicted events (cf. Clark, 2013). Indeed, several studies have found better detection of unexpected intensity increments in metrically weak than strong positions (Abecasis et al., 2009; Geiser et al., 2010), in line with temporal prediction but not temporal attending affecting the processing of rhythmic events.

Second, while attention is thought to enhance early responses to auditory events, prediction is thought to attenuate those responses (Lange, 2013; Schafer, Amochaev, & Russell, 1981). In a study comparing the responses to regular and irregular sound sequences, Schwartze et al. (2013) found attenuation of the auditory P1 response to acoustic events in the regular sequences. Similarly, Sanabria and Correa (2013) showed that the auditory N1 response was attenuated for events presented after a predictable time-interval, but not for events presented after an unpredictable time-interval. These studies show that temporal predictability attenuates the response to acoustic events. As such, while temporal attending as proposed in DAT would lead to enhancement of responses to events at metrically strong positions, temporal prediction would lead to attenuation of these responses, as events in metrically strong positions are highly expected (Ladinig et al., 2009).

In the current study, we aimed to examine the influence of temporal attending and temporal prediction to the processing of a metrical rhythm. To be able to disentangle the contributions of temporal attending and temporal prediction, we used an auditory oddball paradigm in which we introduced infrequent unexpected events in the form of both intensity increments and decrements at different metrical positions in an isochronous rhythm (see Figure 4.1). We expected the rhythm to induce a binary metrical structure, with odd positions being metrically strong (on the beat) and even positions metrically weak (offbeat, see Potter et al., 2009). To ensure that people heard the alternating strong and weak tones with the same phase, a click track sound was superposed on the isochronous rhythm every eight tones.

In Experiment 1, we used a speeded detection task in which participants were required to respond to the deviants. As described above, temporal attending is hypothesized to lead to better detection of deviants in metrically strong than weak positions. Temporal prediction is hypothesized to lead to better detection of increments in metrically weak than strong positions and better detection of decrements in metrically strong than weak positions, consistent with larger prediction errors for increments in weak positions and for decrements in strong positions. As such, while temporal attending and temporal prediction are hypothesized to have the same effect on the detection of decrements, the detection of increments differentiates between the presence of these two mechanisms (Table 4.1). It must be noted that temporal attending and temporal prediction may not be independent. Attending in time may lead to strong predictions about the occurrence of an event (Lange, 2013). If both temporal attending and prediction are present, their effects on the detection of increments may cancel each other out. The concurrent presence of both mechanisms would thus lead to large effects of metrical position on the detection of decrements and null or small effects of metrical position of the detection of increments.

In Experiment 2, we examined, using EEG, whether the influence of metrical structure on the processing of rhythmic events persisted with lower general levels of attentional resources directed at the rhythm. Previously, using ecologically valid stimuli in which acoustic saliency and metrical saliency always coincided, we did find differences in processing of unexpected events in metrically strong and weak positions, even when attention was directed elsewhere, showing that the induction of a metrical structure from exogenous cues is possible with lower levels of attentional resources (Bouwer et al., 2014). Contrary to this, Chapin et al. (2010) found that when listening to highly syncopated rhythms, attention was required to recruit the basal ganglia, which has been associated with the perception of metrical structure (Grahn & Brett, 2007; Grahn, 2009b). It is unclear however, whether the lack of basal ganglia activity found by Chapin et al. (2010) when people were not attending to the rhythms was due to the highly syncopated nature of the rhythms or to the lack of acoustic salient accents indicating the metrical structure. We have suggested that the induction of a metrical structure from rhythms without clear acoustic accents may be possible with lower levels of attentional resources, as long as the metrical structure is sufficiently simple (Bouwer et al., 2014). In Experiment 2, we tested this hypothesis by examining the contributions of temporal attending and temporal prediction to the processing of metrical rhythm while attention was directed away from the rhythm.

Whereas in Experiment 1, reaction times and detection rates provided a direct measure of the detectability of the deviant events, in Experiment 2 we used the mismatch negativity (MMN) as an index of deviant detection. MMN is an ERP component that has been shown to occur without attention directed to a sound (Näätänen et al., 2007) and is affected by our predictions in the auditory modality (Winkler, 2007). As such, it is a very useful instrument to examine the perception of metrical structure, especially under conditions when fewer resources are available (Honing et al., 2014). MMN amplitude indexes the magnitude of a regularity violation (Näätänen et al., 2007) and could therefore function as an index of detectability of the deviants in Experiment 2. In general, an effect of metrical structure on the MMN amplitude in response to deviants would indicate that a metrical structure was induced with lower levels of attentional resources directed at the rhythm. The direction of such an effect could serve as additional evidence to differentiate between temporal attending and temporal prediction. In line with the predictions for Experiment 1, temporal attending was hypothesized to lead to larger MMN amplitudes for deviants in metrically strong than weak positions. Temporal prediction was hypothesized to lead to larger MMN amplitudes for increments in metrically weak than strong positions and for decrements in metrically strong than weak positions.

In addition, the use of EEG allowed us to look at the effects of the metrical structure on auditory evoked potentials at different metrical positions, specifically the P1 and N1. These components are generated in the primary and secondary auditory cortices and have been shown to be sensitive to both attention (Picton & Hillyard, 1974; Woldorff et al., 1993) and prediction (Lange, 2009; Schafer et al., 1981). Whereas enhancement of these components on the beat may be indicative of the presence of temporal attending, attenuation would imply the presence of temporal prediction

(Lange, 2013). Thus, in Experiment 2, a possible effect of metrical structure on auditory evoked potentials would provide additional support that a metrical structure was induced with lower levels of attentional resources, with temporal attending leading to enhancement of evoked potentials in response to events in metrically strong positions and temporal prediction leading to attenuation.

Finally, we looked at possible anticipatory effects of temporal attending and prediction, which may be visible before the onset of a stimulus. Indeed, anticipatory processes related to regularity detection have been shown previously using EEG in beta band oscillatory activity (Fujioka et al., 2012). Temporal expectations have also been linked to ERP components, most notably the contingent negative variation (CNV), a negative-going deflection that has been originally associated with the anticipation of a motoric response (Walter, Cooper, Aldridge, McCallum, & Winter, 1964). CNV has also been shown to occur in the absence of an overt response (Mento, 2013) and is sensitive to the temporal interval that is anticipated, peaking at the expected time of an event (Mento, 2013; Praamstra, Kourtis, Kwok, & Oostenveld, 2006). Thus, temporal expectations can be seen in ERPs even before the onset of an event. Therefore, we also looked at possible differences in the ERPs preceding sounds to examine whether we could differentiate between metrical positions on the basis of anticipatory differences.

To summarize, we examined the influence of a perceived metrical structure on the processing of rhythmic events with and without attention directed at the rhythm. We used an isochronous rhythm in which infrequent intensity increments and decrements were introduced to disentangle the contributions of temporal attending and prediction. In the attended condition (Experiment 1), a speeded reaction time task was used to probe the detectability of the deviants. In the unattended condition (Experiment 2), we used the MMN as an index of detectability and additionally looked at the effects of metrical structure on early auditory evoked potentials and anticipatory activity.

4.2 Experiment 1

4.2.1 Methods

Participants
In this experiment we looked at beat perception in an isochronous rhythm. The lack of acoustic cues and the lower attentional resources (cf. Experiment 2) may lead to weaker effects of beat perception (Bouwer et al., 2014). To maximize the chances of inducing a beat under these circumstances we tested only professional musicians. Twenty highly trained musicians (4 males, 16 females) participated in Experiment 1. They were on average 26 years old (range 18–49 years, standard deviation 8 years) and had had an average of 16 years of formal musical training (range 8–23 years, standard deviation 4 years). The instruments they played were clarinet (3), violin (2), viola (1), cello (3), trumpet (1), trombone (2), bassoon (1), flute (1), oboe (1), French horn (2), and piano (1). Two participants were singers. 18 participants were mostly trained and active in classical music, while two participants were trained and active in other genres (pop, world music, jazz). The participants reported an average of 3.3 h of daily practice on their instrument at the time of the experiment (range 1–7 h, standard deviation 1.3 h). All participants provided written informed consent prior to the study.
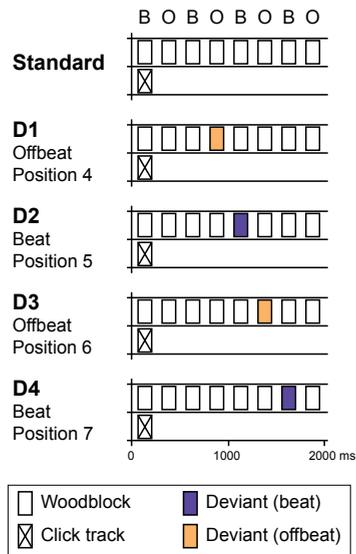
Figure 4.1 Schematic overview of standard and deviant patterns. Standards consisted of eight identical woodblock sounds with an inter-onset interval of 250 ms in which subjects were expected to perceive a binary pattern of alternating beats (B) and offbeats (O). Patterns were presented in a continuous stream. In position 1, a click track sound was superposed on the pattern to ensure phase alignment within the stream of rhythms. Four deviant patterns were used (D1–D4). In two patterns, deviants were introduced in offbeat positions (D1 and D3, positions 4 and 6 respectively). In two patterns, deviants were on the beat (D2 and D4, positions 5 and 7). At each position (D1–D4), two types of deviants were used: intensity increments and intensity decrements. In Experiment 1, deviants of three different magnitudes were used: 4, 6, and 9 dB. In Experiment 2, only 9 dB deviants were used.

The study was approved by the Ethics Committee of the Faculty of Humanities of the University of Amsterdam.

Stimuli

The standard pattern consisted of eight isochronous woodblock sounds with an inter-onset interval of 250 ms (see Figure 4.1). A binary pattern of subjectively accented and unaccented tones at this rate would put the inter-beat interval at 500 ms, close to the preferred tempo for beat perception (Fraisse, 1982; London, 2002). Patterns were presented in a continuous stream. To prevent participants from shifting the phase of the perceived binary pattern, a click track sound was superposed on the pattern in position 1 (see Figure 4.1). The time between two click track sounds was 2000 ms (i.e., every eight events). While this may have induced a regular expectation based on acoustic saliency of the click track sound, it is unlikely that people heard a beat at this very slow rate (London, 2002). The woodblock sound was generated in GarageBand (Apple Inc.). The click track sound was 70 ms long, had a MIDI pitch of 74 (587 Hz) and was generated in Audacity (http://audacity.sourceforge.net/). The peak intensity of the

click track sound was set to 31 dB lower than the peak intensity of the woodblock sound. Figure 4.1 (top) shows a schematic representation of the standard stimulus.

In addition to the standard pattern, we generated patterns containing deviants in four different positions (Figure 4.1, bottom). Two types of deviants were used: intensity increments and intensity decrements. Three different magnitudes of deviants were used: 4, 6, and 9 dB, the smallest being comparable to a subjective accent (Brochard et al., 2003; Povel & Okkerman, 1981). As such, we created a total of 24 different deviant patterns. Deviants were introduced in positions 4, 5, 6, and 7 in the pattern. Previously, using similar stimuli, Bolger et al. (2013) found large effects of metrical expectations in the positions preceding and coinciding with an acoustically salient tone in the first position of an eight-tone pattern. However, as we were specifically not interested in the expectations induced by an exogenous, acoustic cue, we did not use positions 1 and 8, which coincided with and directly preceded the click track sound. In addition, we did not introduce deviants in positions 2 and 3, to avoid confounds due to pattern learning. We have shown that the acoustic context can have a large effect on ERPs in general and MMN in particular, even when difference waves are used (Bouwer et al., 2014; Honing et al., 2014). While difference waves can be used to eliminate the direct effects of acoustic context, the context may have indirect effects on ERPs if a listener has expectations based on the sequential probabilities within a repeating pattern. A deviant in position 2 would have been the only deviant that directly followed the click track sound and as such would have had different sequential properties than the deviants in other positions. While we do not know whether a deviant in position 3 would still be susceptible to this confound, we preferred to err on the side of caution and only introduced deviants in positions 4, 5, 6, and 7 in the pattern.

Procedure
Standard patterns and patterns containing a deviant were presented in a continuous stream (see Supplementary Audio[3]). A deviant could occur in 33% of the patterns. As a deviant was only one out of eight tones in a pattern, of the single tones, 4% was a deviant. Of single tones, 83% were standard woodblock sounds, while 13% were click track sounds. Each of the 24 deviant patterns was presented 25 times. Thus, in total 600 deviant and 1200 standard patterns were presented. The experiment was divided into 12 blocks of 5 min, with each block consisting of 50 deviant and 100 standard patterns. Presentation was pseudo-randomized, with the types and magnitudes of the deviants being completely random while there was always at least one standard pattern between two patterns containing a deviant. Participants were instructed to respond with a button press every time they heard something unexpected in the rhythm. Before the experiment started, they were presented with a practice block of 3 min (60 standard and 30 deviant patterns with the same pseudo-randomization as during the experiment) to get familiarized with the task. If needed, they could repeat this practice block until they felt comfortable doing the task. Stimuli were presented through custom-made speakers that were positioned at an angle of 39° and a distance of 132 cm to both sides

---

[3] Supplementary Audio for this chapter is available online at the publisher's website and at http://www.fleurbouwer.nl/publications.

measured from the back of the chair in which participants were seated. Sound level was set at 60 dB SPL for the standard woodblock sounds, as measured at the back of the chair with a Quest 2800 sound level meter. Presentation® software (Version 14.9, http://www.neurobs.com) was used to present the stimuli.

Analysis

Only responses made between 200 and 1000 ms after presentation of the deviant were included as valid responses. For D1 and D2, this eliminated any responses made after the start of the subsequent pattern. For D3 and D4, this meant responses made after more than 750 and 500 ms respectively were overlapping with the next pattern. For D3, less than 3% of the responses were made after the start of the next pattern. For D4, 29% of responses were made after the start of the next pattern. In the slowest condition at this position (4 dB decrements), 55% of the responses were slower than 500 ms, 85% of the responses were made within 200 ms after the start of the next pattern and 95% were made within 250 ms after the start of the next pattern. As these response times would also have included the motor preparation and response, it is unlikely that they were due to erroneous responses to the next click track sound. Therefore, we did not correct the reaction times beyond the exclusion of reaction times longer than 1000 ms. Average reaction times and miss rates for each condition and each participant were entered into a repeated measures ANOVA with the within subject factors position (D1, D2, D3, and D4), type (increment or decrement) and magnitude (4, 6, or 9 dB difference between the deviant and the standards). We used three orthogonal contrasts to examine possible effects of the position of the deviant. First, to answer our main questions about the contributions of temporal attending and prediction to the processing of metrical rhythm, we compared the responses to deviants on the beat (positions 5 and 7, D2 and D4) with the responses to deviants offbeat (positions 4 and 6, D1 and D3). Second, to examine the possible presence of perceived higher order regularity, we compared the responses to deviants on the third beat (position 5, D2) with the responses to deviants on the fourth beat (position 7, D4). Finally, to check for possible serial position effects, we compared the responses to deviants in the metrically equally weak positions 4 (D1) and 6 (D3). Where applicable, Greenhouse-Geiser corrections were applied to correct for violations of the non-sphericity assumption. The analysis was performed in SPSS Statistics 20.

4.2.2   Results

Figure 4.2 shows the average miss rates for beat and offbeat positions and Figure 4.3 shows the average reaction times. There was a significant interaction between deviant type and metrical position for both miss rates ($F_{(3,57)} = 6.1$, $p = 0.001$, $\eta^2 = 0.24$) and reaction times ($F_{(3,57)} = 10.7$, $p < 0.001$, $\eta^2 = 0.36$). Therefore, we ran additional ANOVAs for increments and decrements separately. For decrements, miss rates were affected by both position ($F_{(3,57)} = 4.9$, $p = 0.004$, $\eta^2 = 0.20$) and magnitude ($F_{(2,38)} = 134.1$, $p < 0.001$, $\eta^2 = 0.88$) of the deviant. Decrements on the beat (D2 and D4) were detected more often than decrements offbeat (D1 and D3; $F_{(1,19)} = 15.4$, $p = 0.001$, $\eta^2 = 0.45$). In addition, decrements on the strong beat in position 5 (D2) were detected more often than decrements on the weaker beat in position 7 (D4; $F_{(1,19)} = 4.6$, $p = 0.045$, $\eta^2 = 0.20$). Reaction times showed a similar pattern of results, with significant effects of position ($F_{(3,57)} = 10.6$, $p < 0.001$, $\eta^2 = 0.36$) and magnitude
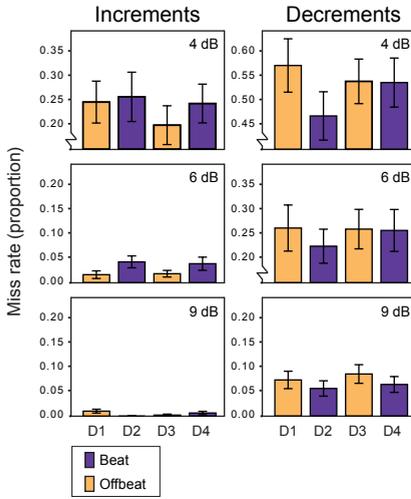
Figure 4.2 Miss rates for all deviants in Experiment 1. Error bars denote one standard error. NB: range of the Y-axis varies between plots for displaying purposes.
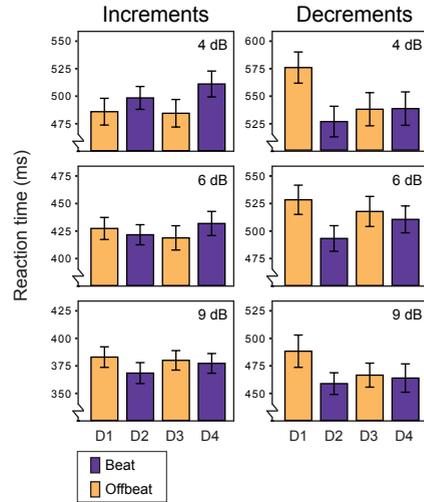
Figure 4.3 Reaction times for all deviants in Experiment 1. Error bars denote one standard error. NB: range of the Y-axis varies between plots for displaying purposes.

($F_{(2,38)} = 57.1$, $p < 0.001$, $\eta^2 = 0.75$). Decrements on the beat were detected faster than decrements offbeat ($F_{(1,19)} = 17.1$, $p = 0.001$, $\eta^2 = 0.47$). Finally, decrements in position 6 (D3) were detected faster than decrements in position 4 (D1; $F_{(1,19)} = 13.7$, $p = 0.002$, $\eta^2 = 0.42$). As this may indicate a serial position effect either hindering detection of D1 or facilitating detection of D3, we performed additional post-hoc contrasts comparing the reaction time for D3 to the reaction times for D2 and D4 separately. While the difference between the reaction times to D2 and D3 was significant ($F_{(1,19)} = 8.4$, $p = 0.009$, $\eta^2 = 0.31$), the comparison between D3 and D4 was not ($F < 0.3$).

For increments, miss rates were also affected by both position ($F_{(3,57)} = 3.6$, $p = 0.020$, $\eta^2 = 0.16$) and magnitude of the deviant ($F_{(2,38)} = 33.2$, $p < 0.001$, $\eta^2 = 0.64$). Contrary to decrements, increments were detected more often offbeat (D1 and D3) than on the beat (D2 and D4; $F_{(1,19)} = 9.9$, $p = 0.005$, $\eta^2 = 0.34$). In addition, increments were detected more often in position 6 (D3) than in position 4 (D1; $F_{(1,19)} = 4.9$, $p = 0.039$, $\eta^2 = 0.21$), which may indicate a similar serial position effect as found for reaction times to decrements. To check whether this may have driven the difference in detection rate between increments on the beat and offbeat, we performed post-hoc tests contrasting the miss rates for D3 with those for D2 and D4. Both comparisons were significant, indicating better detection of increments in position 6 (D3) than positions 5 and 7 (D2; $F_{(1,19)} = 9.5$, $p = 0.006$, $\eta^2 = 0.33$ and D4; $F_{(1,19)} = 6.9$, $p = 0.017$, $\eta^2 = 0.27$).

For reaction times to increments, there was a significant interaction between the position and magnitude of the deviant ($F_{(6,114)} = 2.8$, $p = 0.046$, $\eta^2 = 0.13$). To look at the nature of the interaction effect, we ran ANOVAs for each magnitude separately. The

reaction times for small (4 dB) and large (9 dB) increments were significantly affected by the position of the deviant ($F_{(3,57)} = 3.0$, $p = 0.037$, $\eta^2 = 0.14$ and $F_{(3,57)} = 4.1$, $p = 0.011$, $\eta^2 = 0.18$ respectively). However, metrical position had opposite effects on the detection of small and large increments. Small increments were detected faster offbeat than on the beat ($F_{(1,19)} = 8.4$, $p = 0.009$, $\eta^2 = 0.31$), while large increments were detected faster on the beat than offbeat ($F_{(1,19)} = 13.4$, $p = 0.002$, $\eta^2 = 0.41$). Position did not affect reaction times for 6 dB increments. Finally, 9 dB increments on the strong beat in position 5 (D2) were detected marginally faster than increments on the weaker beat in position 7 (D4; $F_{(1,19)} = 3.9$, $p = 0.062$, $\eta^2 = 0.17$).

4.2.3 Discussion

The results from Experiment 1 suggest that temporal prediction and temporal attending, as well as an interaction between them mediate the effect of metrical position on the perception of rhythmic events. The influence of temporal prediction is apparent from faster and better detection of small increments offbeat than on the beat (see Table 4.1), likely due to the prediction error being larger for increments offbeat than on the beat. The influence of temporal attending is apparent from faster detection of large increments on the beat than offbeat, likely due to heightened sensitivity for events on the beat. The effects of temporal prediction thus seem to be counteracted by temporal attending for large but not small increments. This cannot be explained by assuming additivity of both mechanisms, but instead shows an interaction. Previously, it has been suggested that attention may act to boost the precision of the prediction error (Feldman & Friston, 2010; Kok, Rahnev, Jehee, Lau, & de Lange, 2012). For small increments, the prediction error on the beat was likely very small or even absent, as an increment of this size is comparable in magnitude to a subjective accent (Brochard et al., 2003; Povel & Okkerman, 1981). The weighted prediction error for small increments, taking into account a boost from heightened attentional resources on the beat but not offbeat, was likely still smaller on the beat than offbeat. As the prediction error for large increments would have been substantially bigger, it would have benefitted more from a boost from heightened attentional resources on the beat and this would have outweighed the larger prediction error for increments in offbeat positions. The results for increments as such are consistent not only with the presence of both temporal prediction and temporal attending but also with an interaction between these mechanisms in which attention boosts the precision of predictions. Decrements, as expected, were detected better and faster on the beat than offbeat, which is in line with both temporal prediction and temporal attending.

In addition to differences between the detection of deviants on the beat and offbeat, we also found effects of meter and serial position. Decrements were detected more often and large increments marginally faster on the strong third beat (position 5) than on the weaker fourth beat (position 7), consistent with heightened sensitivity for events in metrically strong positions and thus with temporal attending driving this effect of meter. A serial position effect was apparent from faster detection of decrements and better detection of increments in position 6 than in position 4, while these positions were metrically equally weak. Possibly, the temporal proximity of deviants in position 4 to the click track sound made them harder to detect. When not taking into account position 4, which may have been biased, our post-hoc contrasts show that decrements

on the third beat (position 5) were detected faster than decrements offbeat (position 6) and increments were detected better offbeat (position 6) than on the beat (positions 5 and 7). As such, the observed effects of temporal attending and prediction cannot be explained solely by the presence of a serial position effect.

While the results of Experiment 1 do not allow us to estimate the relative contribution of the two mechanisms involved, we showed that temporal attending, temporal prediction and an interaction between them influence the processing of rhythmic events within a metrical structure. In Experiment 2, using EEG, we examined whether the same mechanisms would be present with lower general levels of attention resources devoted to the rhythm.

4.3   Experiment 2

4.3.1 Methods

Participants
Twenty-four highly trained musicians (8 males, 16 females) participated in Experiment 2, 12 of whom had also participated in Experiment 1. Their average age was 28 years old (range 19–58 years, standard deviation 8 years) and they had received an average of 19 years of formal musical training (range 7–46 years, standard deviation 8 years). The instruments this group of participants played were clarinet (3), violin (5), cello (3), trumpet (1), bassoon (1), flute (2), guitar (2), French horn (3), and piano (3). One participant was a singer. Twenty-two participants were mostly trained and active in classical music, while two participants were trained and active in other genres (pop, world music, jazz). They reported an average of 3.1 h of daily practice on their instrument at the time of the experiment (range 1–5 h, standard deviation 1.1 h).

Stimuli
The stimuli were largely the same as those used in Experiment 1 (see Figure 4.1). However, due to time constraints imposed by the use of EEG we only used deviants of 9 dB, as we expected large deviants to elicit a reliable MMN. Deviants were, similar to Experiment 1, either increments or decrements and were introduced at positions 4, 5, 6, and 7 in the rhythm. In total, we thus used eight deviant patterns. The peak amplitude of the click track sound in Experiment 2 was set to 10 dB lower than the peak intensity of the woodblock sound to ensure participants heard the metrical structure with the same phase alignment under unattended conditions.

Procedure
Increments and decrements were tested in separate sessions using 150 deviants on each of the eight possible positions, resulting in a total of 600 deviant patterns for each type. Deviant patterns represented 33% of the total patterns, with deviant tones making up 4% of total sounds. Thus, a total of 1800 patterns was presented in each session. Patterns were presented in five blocks of 12 min (360 patterns), presented in a continuous stream. As in Experiment 1, patterns were presented in pseudo-randomized order, with at least one standard pattern between two patterns containing a deviant. To minimize possible effects of short-term learning of the rhythmic pattern during the attended behavioral task, those participants that participated in both Experiment 1 and Experiment

2 participated in the EEG task either preceding the behavioral task or on a different day. During the presentation of the rhythms participants watched a self-selected silenced movie with subtitles. They were instructed to concentrate on the movie and to ignore the rhythm. All participants indicated that they could comply with this task. Each condition took around 1 h to complete. Participants could take breaks as needed. The sound equipment was identical to Experiment 1.

EEG Recording

EEG was recorded with a 64-channel Biosemi Active-Two reference free acquisition system (Biosemi, Amsterdam, The Netherlands), using the standard 10/20 configuration and additional electrodes at both mastoids, around the eyes and on the nose. The EEG signal was recorded at 8 kHz.

EEG Analysis

EEG preprocessing was performed in Matlab (Mathworks, Inc.) using the EEGLAB toolbox (Delorme & Makeig, 2004). The statistical analysis was performed in SPSS Statistics 20. For all analyses described below, where applicable Greenhouse-Geiser corrections for non-sphericity were used. For the analysis of ERP responses to both deviants and standards, EEG data was offline re-referenced to linked mastoids and down-sampled to 512 Hz. In eleven participants, one or more bad channels was removed and subsequently interpolated from the surrounding channels. None of these channels is reported here. Independent component analysis was used to remove eyeblinks.

Analysis of ERP Responses to Deviants

For the analysis of the MMN, data were filtered between 0.5 and 20 Hz, using a linear finite impulse response filter and 650 ms epochs were extracted from the continuous data starting 150 ms before the onset of each deviant. Epochs at the same positions were extracted from the standard patterns. Epochs with an amplitude difference of more than 100 microvolts within a 500 ms sliding window were rejected from the analysis, epochs were averaged for each condition separately and baseline corrected using the average activity of the 150 ms pre-stimulus period. Deviant-standard difference waves were calculated by subtracting the ERP obtained in response to the standards from the ERP in response to the deviants aligned in time relative to the start of the pattern. We defined the MMN as the negative peak between 100 and 200 ms after the onset of the deviant. Visual inspection of the group averaged difference waves for the different conditions showed a large difference in morphology between the responses to increments and decrements. To quantify this difference, we performed an analysis of the peak latencies of the MMN at electrode Fz (see Table 4.2). Peak latencies for all participants for all deviants were entered into an ANOVA with factors type (increments and decrements) and position (D1, D2, D3, and D4). The type of deviant significantly affected the peak latency, with later peaks for decrements than increments ($F_{(1,23)} = 11.4$, $p = 0.003$, $\eta^2 = 0.33$). No effects of position on peak latencies was observed, nor an interaction between type and position.

A difference between the responses to increments and decrements has previously been observed (Rinne et al., 2006) and may be due to overlap with other ERP components

that are affected by the intensity of the deviants. As the responses to the different deviant types were qualitatively different, we performed the statistical analysis separately for increments and decrements. We calculated the average difference waves for increments and decrements collapsed over the four metrical positions. These difference waves are shown in Figure 4.4 (top). The MMN for increments peaked at a latency of 140 ms, while the MMN for decrements peaked at 169 ms. At the peak latency, the MMN for increments showed a right-frontal scalp distribution, while the MMN for decrements was slightly more centrally located. For both types, we defined a region of interest for the analysis of the MMN encompassing the 6 electrodes with highest amplitudes at the peak latency. These regions of interest are indicated in Figure 4.4 (top).

For the analysis, the MMN amplitude was defined as the average amplitude in a 60 ms window around the peak of the MMN for each type collapsed over positions. As such, we defined the window for analysis independent from the metrical positions, while acknowledging the differences due to the different types of deviants. MMN amplitudes were entered into a repeated measures ANOVA with the within subject factor position (D1, D2, D3, and D4). The same contrasts as in Experiment 1 were used to explore the effect of the position of the deviant on the MMN amplitude. To examine the effect of metrical structure, the responses to deviants in offbeat positions (D1 and D3) were compared to the responses to deviants on the beat (D2 and D4). To examine the possible presence of higher order regularity in the form of meter, we compared the response to deviants on the third beat (D2) to the response to deviants on the, theoretically less salient, fourth beat (D4). Finally, to look at possible serial position effects, we compared the responses to deviants in positions 4 and 6 (D1 and D3), which were both metrically weak.

Analysis of ERP Responses to Standards
Regarding the analysis of AEPs in response to the standards, we were mainly interested in the P1 and N1 components. To optimize the analysis of the standards to these shorter latency components, we filtered the data using linear finite impulse response filtering between 5 and 75 Hz (see Schwartze et al., 2013, for a discussion of these filter settings). Epochs starting at 50 ms before the onset of each sound in the standard patterns and ending at 250 ms after the onset of each sound were extracted from the continuous data. Epochs with an amplitude difference larger than 150 microvolts were rejected and epochs were averaged for each position separately to obtain ERPs. ERPs were averaged over blocks of deviant types, as the standards were exactly the same in both conditions. No baseline correction was applied. With a stricter high-pass filter, the effects of slow amplitude changes are much less pronounced, making baseline correction unnecessary. Also, while for the MMN analysis we were interested in the reaction to the deviants, which starts the moment the deviant sound is heard, for the analysis of the standards, we were also interested in possible differences in anticipatory activity. If these effects would indeed be present, a baseline correction would falsely eliminate any differences between conditions, while possibly falsely creating differences between conditions in the P1 or N1 responses due to differences in the baseline.

The amplitude of the P1 and N1 was defined as the average amplitude in a 40 ms window around the average latency of the peaks of these components for all four positions. The peak latency of the P1 response was 63 ms and the peak latency of the N1
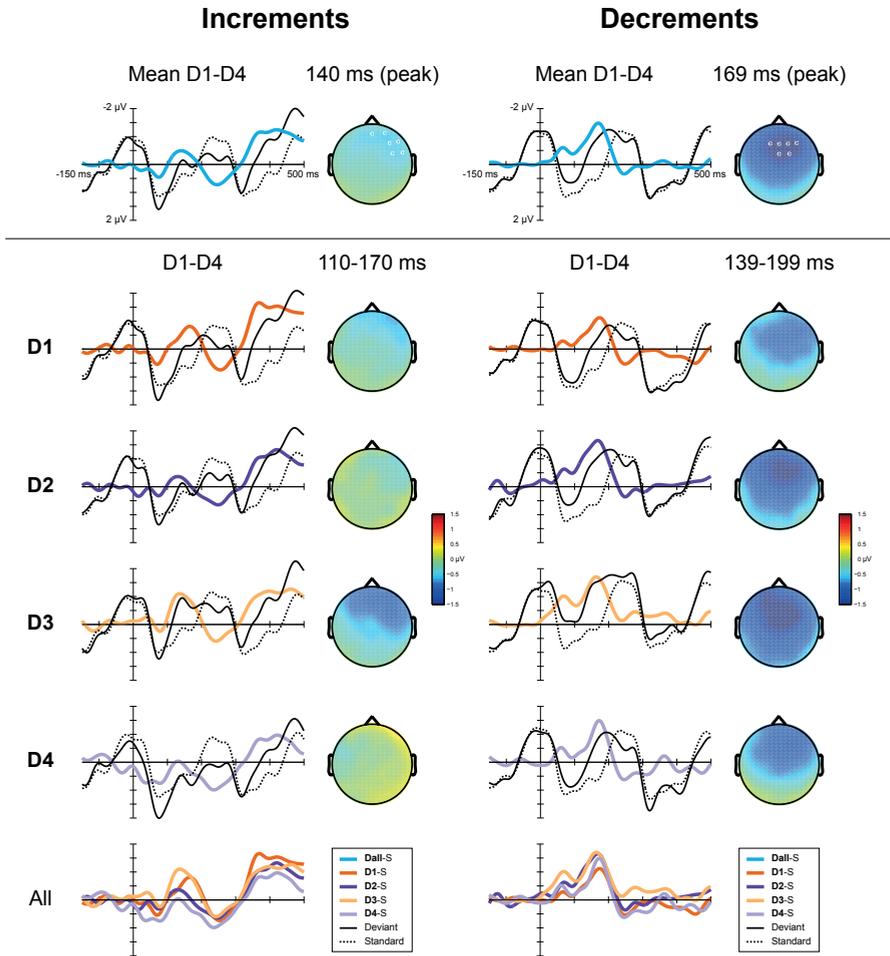
Figure 4.4 ERP responses to the deviants in Experiment 2 for increments (left) and decrements (right). Top panels show the difference waves for both types collapsed over positions, the scalp distributions at the peak latency of the MMN and the regions of interest used for the analysis. Middle panels show, for each position separately, group averaged ERPs elicited by the deviants, the standards (S), the derived difference waves and the scalp distribution of the MMN averaged over the analysis window. The bottom panel shows all difference waves combined.

Table 4.2 Mean average peak latencies and average amplitudes of the MMN to deviants. Peak latencies are the negative peak between 100 and 200 ms on Fz. Amplitudes are as used for the analysis, measured on ROIs as specified in Figure 4.4 from a 60 ms window around the peak for the averaged increments and decrements separately. Standard deviations in brackets.

Deviant	Average Peak Latency (ms)		Average Amplitude (µV)	
	Increments	Decrements	Increments	Decrements
D1	157 (32)	165 (27)	−0.52 (1.47)	−0.93 (1.41)
D2	140 (27)	165 (23)	−0.26 (1.19)	−1.44 (1.10)
D3	148 (30)	161 (27)	−0.97 (1.50)	−1.48 (1.13)
D4	149 (31)	164 (30)	0.11 (1.70)	−1.14 (1.18)

response was 133 ms. For anticipatory activity, the 40 ms window was centered around 0, where anticipatory activity was expected to be maximal. Statistical analysis was thus conducted for three time windows: 43–83 ms for P1, 113–153 ms for N1 and –20–20 ms to look at differences in anticipatory activity. For the analysis of the standards, we used a region of interest containing fronto-central midline electrodes (Cz, FCz, and Fz). Like for the deviants, we only included the ERPs in response to sounds in positions 4–7 in the analysis, to avoid confounds due to the click track sound. We tested the same orthogonal contrasts as described for the analysis of the MMN.

4.3.2 Results

ERP Responses to Deviants

Figure 4.4 (bottom) shows the difference waves for all deviants. Table 4.2 shows the average amplitudes and peak latencies for all conditions. For increments, we found a marginal effect of metrical position, with a larger amplitude MMN offbeat than on the beat ($F_{(1,23)} = 3.0$, $p = 0.097$, $\eta^2 = 0.12$). In addition, the MMN to increments on the strong third beat (D2) was marginally larger than the MMN to increments on the weaker fourth beat (D4; $F_{(1,23)} = 2.9$, $p = 0.10$, $\eta^2 = 0.11$). For decrements, the MMN to deviants on position 4 (D1) was smaller than the MMN to deviants on position 6 (D3; $F_{(1,23)} = 5.6$, $p = 0.026$, $\eta^2 = 0.20$), possibly indicating a serial position effect.
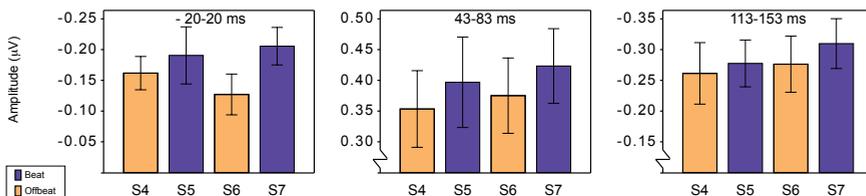
Figure 4.5 Average magnitudes of ERP components in response to standards on the beat and offbeat in Experiment 2. Anticipatory negativity (left), P1 (middle), and N1 (right). Responses are shown for positions 4–7 in the standards, corresponding to the positions in which deviants D1–D4 could occur.

Table 4.3 Mean average peak latencies and average amplitudes of the ERP responses to standards. For the anticipatory negativity, amplitudes are as measured from a 40 ms window around the onset of the sound. We do not report peak latencies for this component as we cannot estimate the peak from our data. Peak latencies for P1 are defined as the positive peak between 40 and 100 ms on midline electrodes. Peak latencies for N1 are defined as the negative peak between 100 and 180 ms on midline electrodes. Amplitudes are as used for the analysis, measured on midline electrodes as specified in Figure 4.6 from a 40 ms window around the peak for each component averaged over conditions. Standard deviations in brackets.

Standard	Anticipatory	P1		N1	
	Average Amplitude (μV)	Average Peak Latency (ms)	Average Amplitude (μV)	Average Peak Latency (ms)	Average Amplitude (μV)
S4	−0.16 (0.13)	72 (16)	0.35 (0.31)	129 (18)	−0.26 (0.25)
S5	−0.19 (0.23)	71 (14)	0.40 (0.36)	132 (15)	−0.28 (0.19)
S6	−0.13 (0.16)	65 (11)	0.38 (0.30)	127 (16)	−0.28 (0.22)
S7	−0.21 (0.15)	66 (12)	0.42 (0.30)	130 (13)	−0.31 (0.20)

ERP Responses to Standards

Figure 4.5 shows the average amplitudes for all positions in the standard pattern of all time windows of interest. Table 4.3 lists the average amplitudes and peak latencies. ERPs for positions 4–7, collapsed over metrical levels, are shown in Figure 4.6. Around the baseline, the anticipatory activity was more negative for sounds on the beat than offbeat ($F_{(1,23)} = 5.2$, $p = 0.033$, $\eta^2 = 0.18$). The P1 amplitude was larger on the beat than offbeat ($F_{(1,23)} = 4.30$, $p = 0.049$, $\eta^2 = 0.16$). None of the other contrasts was significant.
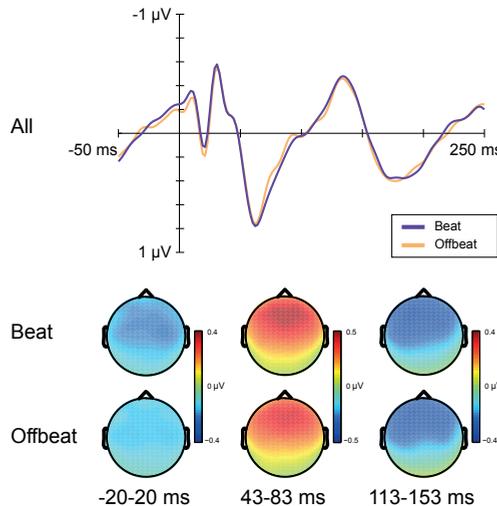
Figure 4.6 ERP responses elicited by standards on the beat and offbeat. Top panel shows ERPs collapsed over metrical position (on the beat: positions 5 and 7; offbeat: positions 4 and 6). Bottom panel shows scalp distributions for analysis windows.

4.3.3   Discussion

The results of Experiment 2 regarding the responses to deviants suggest that even with lower levels of attentional resources available for the perception of a rhythm, temporal prediction and temporal attending affect processing of regular rhythmic events. The MMN amplitude for intensity increments was marginally larger offbeat than on the beat. This is in line with a larger prediction error for increments offbeat than on the beat and thus suggests the presence of temporal prediction. In addition, the MMN amplitude for increments on the strong third beat was marginally larger than for increments on the weaker fourth beat. This is in line with heightened sensitivity for events in metrically salient positions and thus suggests the presence of temporal attending. However, the results for the deviants are tentative at best, with no effect of metrical position on the MMN responses to intensity decrements, and only marginally significant effects of metrical position on the MMN responses to increments. The latter may be due to the effects of temporal attending and temporal prediction canceling each other out. However, for decrements, the simultaneous presence of both mechanisms should have strengthened the results. Also, like in Experiment 1, serial position effects could be observed for decrements, with smaller responses to decrements in position 4 than in position 6. As such, we have to be cautious in interpreting the findings regarding the influence of metrical position on MMN amplitude.

The results of Experiment 2 regarding the responses to standards provide additional support for the presence of temporal attending. The P1 response was larger for events on the beat than offbeat, consistent with the results of Tierney and Kraus (2013). This enhancement of the response to sounds on the beat may be due to attention peaking at metrically strong moments in time and leading to enhancement of early sensory processing (Lange, 2013). We did not find any effect of metrical position on the amplitude of the N1. A similar enhancement due to attention of the P1 but not the N1 has been reported previously (Karns & Knight, 2009; Tierney & Kraus, 2013). However, the opposite effects, attenuation of the P1 and enhancement of the N1, have also been shown simultaneously in a study manipulating the temporal predictability of auditory events (Rimmele, Jolsvai, & Sussman, 2011). These different results are likely due to differences in stimuli and tasks that influenced the relative contributions of temporal attending and prediction.

Finally, in anticipation of standard events on the beat, ERPs were more negative than in anticipation of standard events offbeat. The fact that this difference was present at the onset of the events and that the activity for more expected events (on the beat) was negative relative to the activity for less expected events (offbeat) makes it reminiscent of the contingent negative variation (CNV; Walter et al., 1964), a negative-going ERP component peaking at the expected time of an event. Whether the processes underlying the CNV are relevant to the perception of a metrical structure is unclear, but our results show that it may be fruitful to acknowledge possible differences in brain activity preceding the onset of events when examining the perception of metrical rhythm using ERPs.

One final remark must be made about the ERP results. While all participants reported being able to focus on the movie during the experiment, we cannot completely rule out

that the results we found are due to lapses in attention. We feel confident that participants were listening to the rhythms with lower levels of attentional resources while watching the movie than while performing a task on the rhythm itself. However, to draw stronger conclusions about the influence of attentional resources on the perception of metrical structure, results with and without attention directed at the rhythm should be acquired using the same method. Furthermore, to be able to prevent and control for attentional lapses a continuous task should be used to direct attention away from the rhythm. Within the context of EEG research, this provides practical challenges that future experiments will have to tackle.

4.4   General Discussion

We have shown that the induced metrical structure influences the processing of rhythmic events through the influence of both temporal attending and temporal prediction. Moreover, our data suggest that both temporal attending and prediction are involved in processing of metrical rhythm when attention is directed away from the rhythm. Temporal attending was apparent from heightened sensitivity for events in strong metrical positions. Unexpected intensity decrements and large increments were detected better and faster on the beat than offbeat and decrements were detected better on the strong third beat than on the weaker fourth beat (Experiment 1). In addition, the auditory P1 for standard events on the beat was enhanced and the MMN amplitude for increments was marginally larger on the strong third beat than on the weaker fourth beat (Experiment 2). Temporal prediction was apparent from better detection of events that elicited a large prediction error. Small increments were detected faster and better offbeat than on the beat (Experiment 1) and the MMN amplitude for increments on the beat was marginally larger than for increments offbeat (Experiment 2). Finally, an interaction between temporal attending and prediction was evident from the interaction between the magnitude of the deviant and the effect of metrical position in Experiment 1. This interaction is in line with temporal attention boosting the precision and the weighting of the prediction error (Kok et al., 2012).

The complex interplay of temporal attending and temporal prediction may explain previous conflicting findings regarding the processing of metrical rhythm. While some studies found enhancement of early sensory processing in metrically strong positions (Tierney & Kraus, 2013), others found attenuation (Schwartze et al., 2013). Interestingly, while the former study used real music, and as such had stimuli with presumably multiple levels of regularity present, the latter used isochronous sequences. Arguably, while this tests regularity detection, it is not necessarily examining metrical structure, which by nature has a hierarchical component (Fitch, 2013; Vuust & Witek, 2014). In the current study, consistent with temporal attending, we found enhancement of the auditory P1 in metrically strong positions. We compared responses on the beat with responses offbeat, which constitute different levels in a metrical hierarchy. At a higher level, the differences in responses to deviants on the strong third and weak fourth beat were also consistent with heightened sensitivity for events in metrically strong positions and thus with temporal attending. Possibly, temporal attending plays a relatively larger role than temporal prediction in shaping our perception when different hierarchical levels are used. This would fit nicely with a neural resonance account of metrical

perception, which presumes that multiple emergent oscillators cause dynamic fluctuations in attentional resources and the perception of regularity at multiple hierarchical levels (Large, 2008).

Several other factors may influence the relative contributions of temporal attending and prediction on the processing of metrical rhythm. First, it has been suggested that temporal attending is an endogenously driven process, while temporal prediction is driven by bottom-up cues (Sanabria & Correa, 2013). While we found evidence of both processes using stimuli that required mainly endogenous generation of the metrical structure, it is possible that the relative contribution of temporal prediction would be bigger when using stimuli with more exogenous cues indicating the metrical structure. Second, the balance between temporal attending and prediction may be affected by the amount of resources available for processing a rhythm. With the current design, we cannot compare the results of the attended behavioral experiment and the unattended EEG experiment directly. Third, different ERP components may be affected differently by temporal attending and prediction. MMN has been specifically linked to predictive coding (Winkler & Czigler, 2012), and may therefore be more sensitive to the effects of temporal prediction than temporal attending. Also, in the current study, the effect of metrical structure on the amplitude of the auditory P1 but not the N1 may indicate a difference in the sensitivity of these components to temporal attending and prediction. This would also explain the inconsistent findings for these components in previous studies (Rimmele et al., 2011; Schwartze et al., 2013; Tierney & Kraus, 2013).

The effects of temporal attending and prediction we found in Experiment 2, with lower levels of attentional resources directed at the rhythm, were very small, despite the high level of musical expertise of our participants. Previously, we have shown that musically untrained individuals can induce a metrical structure from a rhythm with clear acoustic accents even with lower levels of attentional resources (Bouwer et al., 2014). Whether musical training is necessary to induce a metrical structure from stimuli without acoustic accents under these circumstances remains to be tested. However, as we have shown here that multiple processes contribute to the processing of metrical rhythm, it may be fruitful to look at the influence of musical training on temporal attending and prediction separately. Possibly, temporal attending is a process arising from the properties of the brain itself (Large, 2008) and as such independent of musical training, while temporal prediction relies more on long term learning of musical structure (Vuust & Witek, 2014) and thus may be more susceptible to musical training. As such, temporal predictions may in fact be derived from the perceptual effects of temporal attending. The relationship between temporal attending and prediction and whether musical training, attentional resources and the presence of hierarchy and exogenous cues in a rhythm indeed affect their relative contributions to the processing of metrical rhythms is an interesting topic for future studies.

4.5   Conclusion

We provided evidence in support of concurrent effects of both temporal attending and temporal prediction on the processing of metrical rhythm. This was shown both in an

attended behavioral task and in an EEG experiment with attention directed away from the rhythm. These mechanisms can provide useful notions in decomposing the top-down influence of a metrical structure on the processing of rhythm. This opens up interesting possibilities for future work, which should take into account that the perception of metrical rhythm is not simply one process. In addition, the relationship between these processes may inform us about mechanisms underlying the human ability to perceive a metrical structure in musical rhythm, which while being a fundamental aspect of music cognition (Honing, ten Cate, Peretz, & Trehub, 2015), is still ill understood.

Chapter 5

Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study[*]

The perception of a regular beat is fundamental to music processing. Here we examine whether the detection of a regular beat is pre-attentive for metrically simple, acoustically varying stimuli using the mismatch negativity (MMN), an ERP response elicited by violations of acoustic regularity irrespective of whether subjects are attending to the stimuli. Both musicians and non-musicians were presented with a varying rhythm with a clear accent structure in which occasionally a sound was omitted. We compared the MMN response to the omission of identical sounds in different metrical positions. Most importantly, we found that omissions in strong metrical positions, on the beat, elicited higher amplitude MMN responses than omissions in weak metrical positions, not on the beat. This suggests that the detection of a beat is pre-attentive when highly beat inducing stimuli are used. No effects of musical expertise were found. Our results suggest that for metrically simple rhythms with clear accents beat processing does not require attention or musical expertise. In addition, we discuss how the use of acoustically varying stimuli may influence ERP results when studying beat processing.

5.1   Introduction

In music, people often perceive regularly recurring salient events in time, known as the beat (Cooper & Meyer, 1960; Honing, 2013). Beat perception has been suggested to be a fundamental and innate human ability (Honing, 2012) and has been explained as neural resonance at the frequency of the beat (Fujioka et al., 2012; Large, 2008; Nozaradan et al., 2011, 2012) caused by regular fluctuations in attentional energy (Large & Jones, 1999). While the ease with which humans can pick up a beat is remarkable, it remains an open question how much attentional resources are needed to detect a beat. Some suggested that focused attention is necessary both for beat perception (Chapin et al., 2010; Geiser et al., 2009) and regularity detection in general (Schwartze et al., 2011). Others argued that beat processing and possibly even the processing of meter – alternating stronger and weaker beats – are in fact pre-attentive (Bolger et al., 2013; Ladinig et al., 2009, 2011) and that beat processing might even be functional in (sleeping) newborns (Winkler et al., 2009).

In the former studies, in which no evidence of beat processing without attention was found, only the temporal structure of the rhythm was varied to indicate the metrical structure (Geiser et al., 2009) and highly syncopated rhythms were used (Chapin et al., 2010). Conversely, the latter studies (Ladinig et al., 2009; Winkler et al., 2009) used strictly metrical stimuli with not only variation in the temporal structure of the rhythm, but also variation in the timbre and intensity of tones to convey the metrical structure. The use of such acoustically rich, ecologically valid stimuli could be essential to allow the listener to induce a beat pre-attentively (Bolger et al., 2013), arguably because multiple features in the stimuli carry information about the metrical structure. However, in these studies a beat was induced by using different sounds for metrically strong and metrically weak positions. While these different sounds may have aided in inducing a beat, this leaves open the possibility that different responses to tones in different metrical positions are due to acoustic differences rather than beat processing (Honing et al., 2014). To rule out this explanation, in the current study, we test whether beat processing is pre-attentive using stimuli that resemble real music whilst probing positions varying in metrical salience but with identical acoustic properties.

We examine beat processing with a mismatch negativity (MMN) paradigm. The MMN is an auditory ERP component that is elicited when acoustic expectations are violated (Bendixen, Schröger, & Winkler, 2009; Winkler, 2007). The MMN is known to be independent of attention and the amplitude of the MMN response indexes the magnitude of the expectancy violation (Näätänen et al., 2007). Also, the MMN response has been shown to correlate with behavioral and perceptual measures of deviance detection (Jaramillo, Paavilainen, & Näätänen, 2000; Näätänen et al., 2007; Novitski et al., 2004; Tiitinen, May, Reinikainen, & Näätänen, 1994). We compare the pre-attentive MMN response to unexpected omissions of sounds in different metrical positions in a music-like rhythm. As the omission of a sound in a metrically strong position is a bigger violation of the metrical expectations than the omission of a sound in a metrically weak position, we expect the MMN response to depend on the metrical position of the omissions, with larger responses for omissions in metrically stronger positions.

Finally, we compare the responses of musicians and non-musicians. Earlier, it has been shown that musical training affects beat processing (Chen, Penhune, & Zatorre, 2008b) and can enhance several aspects of pre-attentive auditory processing, including melodic encoding (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004), detection of numerical regularity (Van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005) and sequence grouping (Van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004). Here we assess whether musical training can also affect the pre-attentive processing of temporal regularity. If beat processing is indeed a fundamental human ability, we expect to find no difference between musicians and non-musicians. However, if beat processing is learned behavior, we expect this ability to be influenced by musical expertise and thus we expect a bigger effect of metrical position on the MMN responses in musicians than in non-musicians.

5.2 Materials and Methods

5.2.1 Ethics Statement

All participants gave written informed consent before the study. The experiment was approved by the Ethics Committee of the Faculty of Social and Behavioral Sciences of the University of Amsterdam.

5.2.2 Participants

Twenty-nine healthy adults participated in the experiment. Fourteen were professional musicians, or students enrolled in a music college (mean age, 29 years; age range, 22–57 years; 8 females). On average, they had received 18.5 years of musical training (range 9–36 years) and they reported playing their instrument at the time of the experiment on average 3.4 hours per day (range 1–5 hours). This group was considered musicians. Fifteen participants (mean age, 31 years; age range, 22–55 years; 9 females) did not play an instrument at the time of the experiment and had received on average 1.2 years of musical training (range 0–2 years), ending at least 10 years prior to the experiment. These participants were considered non-musicians. All participants had received college education or higher and none reported a history of neurological or hearing problems.

5.2.3 Stimuli

We presented participants with a continuous stream of varying rhythm designed to induce a regular beat in a music-like way (for studies using a similar paradigm, see Honing et al., 2012; Ladinig et al., 2009; Winkler et al., 2009). We used a rhythmic sequence composed of seven different patterns. Of these patterns, four were used as standard patterns (S1–S4) and three were used as deviant patterns (D1–D3). Figure 5.1 shows an overview of all patterns. The base pattern (S1) consisted of eight consecutive sounds, with an inter-onset interval of 150 ms and a total length of 1200 ms. Hi-hat, snare drum and bass drum sounds were organised in a standard rock music configuration. We created sounds using QuickTime's drum timbres (Apple Inc.). The bass drum and snare drum sounds always occurred together with a simultaneous hi-hat sound. For the remainder of this paper, we will refer to these combined sounds as bass drum sound
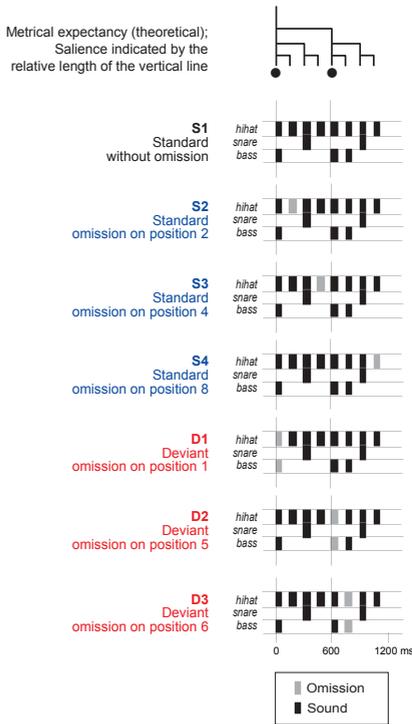
Figure 5.1 Schematic illustration of the rhythmic patterns used in the experiment. The pattern consisted of eight sounds and was designed to induce a rhythm with a hierarchical metrical structure (see tree-structure at the top; beats are marked with dots). The omissions occurred in positions varying in metrical salience, with the omissions in D1 on the first beat, the omissions in D2 on the second beat and the other omissions in equally weak metrical positions.
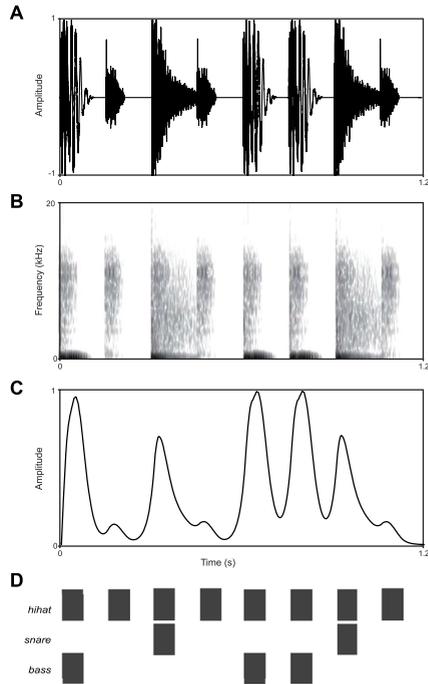
Figure 5.2 Acoustic analyses of stimulus S1. A) Waveform, B) spectrogram, C) amplitude envelope, and D) diagram of stimulus S1 (cf. Figure 5.1). The spectrogram was calculated with a Short Time Fourier Transform, Gaussian window, window size 2 ms, time resolution 5 ms, frequency resolution 20 Hz, and 50 dB dynamic range. The amplitude envelope was calculated using a loudness model as described in (Moore, Glasberg, & Baer, 1997).

(positions one, five and six, see Figure 5.1) and snare drum sound (positions three and seven, see Figure 5.1). Sound durations were 50, 100 and 150 ms for hi-hat, bass drum and snare drum respectively.

Figure 5.2 depicts the acoustic properties of the base pattern (S1). The intensity of the bass drum sound was largest, followed by the intensity of the snare drum sound. The hi-hat sound had the lowest intensity. Therefore, the latter, the shortest and softest sound, would likely be interpreted as metrically weakest, while the bass drum sound would likely be interpreted as metrically strongest. This is in line with the way this pattern is often used in Western music, in which the bass drum indicates the downbeat,

the snare drum indicates the offbeat and the hi-hat is used for subdivisions at the weakest metrical level. We expected the bass drum sounds at positions one and five to be interpreted as beats as they occurred with a regular inter-onset interval of 600 ms. As such, the pattern was expected to induce a beat at 100 beats per minute, a tempo close to the preferred rate for beat perception (London, 2012). At this rate, each pattern encompassed two beats. The first and fifth position of the pattern coincided with respectively the first and second beat, while the second, fourth, sixth and eighth position were metrically weak positions (Figure 5.1).

The base pattern (S1) was varied to create three additional standard patterns (S2–S4). In these patterns a hi-hat sound was omitted in positions two (S2), four (S3) and eight (S4). As such, the omissions in the standard patterns were all in metrically weak positions, that is, not on the beat. Together, the four standard patterns created a rhythm in which the surface structure varied, as is the case in natural music, but in which the metrical structure was left intact, to be maximally beat inducing. The standard patterns accounted for 90% of the total patterns.

The standard patterns were interspersed with three infrequent deviant patterns, accounting for the remaining 10% of the total patterns. In the deviant patterns (D1–D3) a bass drum sound was omitted. In deviant pattern D1 the sound on the first beat (position one), the most salient position in the pattern, was omitted. In deviant pattern D2 the sound on the second beat (position five) was omitted. Both in pattern D1 and in pattern D2 the omission of a sound on the beat violated the metrical structure and created a syncopation. In the third deviant pattern (D3), the same sound was omitted as in patterns D1 and D2, but in a metrically weak position (position six), leaving the metrical structure of the pattern intact.

We examined the presence of pre-attentive beat and meter processing by comparing the MMN responses to the omissions in the deviant patterns. We expected the magnitude of the MMN response to be affected by the metrical position of the omissions in two ways. First, we expected the amplitude of the MMN to omissions in D1 and D2, which were on the beat and thus violated the metrical expectations, to be larger than the amplitude of the MMN to omissions in D3, which was not on the beat and thus left the metrical structure intact. Such a difference would indicate that a beat was detected by the auditory system. Second, we expected to find a larger MMN response to omissions in D1 (on the first beat) than to omissions in D2 (on the second beat) as the former are bigger violations of the metrical expectations than the latter. Such a difference would suggest that a hierarchy between consecutive beats was detected, hence would be evidence for meter processing.

Importantly, the omissions in patterns D1, D2 and D3 could not be distinguished from each other based on the acoustic properties of the sound that was omitted (a bass drum sound) or their probability of occurrence (0.033 for each deviant pattern). Thus, we probed three metrically different positions with exactly the same procedure. Post hoc, we also assessed the effects of the acoustic variation in the stimuli by comparing the

MMN responses to omissions of acoustically different sounds that were all in metrically equally weak positions, that is, the omissions in patterns D3 (a bass drum sound), S2, S3 and S4 (hi-hat sounds).

The patterns were delivered as a randomized continuous stream, without any gaps between consecutive patterns (see Sound S1[4] for a short example of the stimuli in a continuous stream). There were two constraints to the randomization. First, a deviant pattern was always preceded by at least three standard patterns. Second, no deviant pattern could be preceded by standard pattern S4, because this could potentially create two consecutive gaps. In the EEG experiment the stimuli were presented in 20 blocks of 300 patterns. Of these, 10% were deviant patterns, making the total number of trials for each of the three positions 200. Six additional standard patterns were added to the beginning (5) and end (1) of each block. Thus, each block lasted just over 6 minutes and the total number of standard patterns in the whole experiment was 5520, or 1380 trials for each of the four standard patterns. Stimuli were presented through two custom made speakers at 60 dB SPL using Presentation software (Version 14.9, www.neurobs.com).

5.2.4 Procedure

Participants were tested individually in a soundproof, electrically shielded room at the University of Amsterdam. During presentation of the sounds, they watched a self-selected, muted, subtitled movie on a laptop screen. Every block of stimuli was followed by a break of 30 seconds. Longer breaks were inserted at the participants' need. Participants were instructed to ignore the sounds and focus on the movie. In a questionnaire administered after the experiment all of the participants reported being able to adhere to these instructions. This questionnaire was also used to obtain information about their musical experience. Including breaks, the entire experiment took around 2,5 hours to complete.

5.2.5 EEG recording

The EEG was recorded with a 64 channel Biosemi Active-Two reference-free EEG system (Biosemi, Amsterdam, The Netherlands). The electrodes were mounted on an elastic head cap and positioned according to the 10/20 system. Additional electrodes were placed at the left and right mastoids, on the tip of the nose and around the eyes to monitor eye movements. The signals were recorded at a sampling rate of 8 kHz.

5.2.6 EEG analysis

EEG pre-processing was performed using Matlab (Mathworks, Inc.) and EEGLAB (Delorme & Makeig, 2004). The EEG data was offline re-referenced to linked mastoids, down-sampled to 256 Hz and filtered using 0.5 Hz high-pass and 20 Hz low-

---

[4] In example Sound S1, each deviant appears once and in total 30 patterns have been concatenated. The order of appearance of the stimuli in this example is: S1–S4–S3–S1–S2–S1–S2–D2–S4–S2–S3–S2–S3–S3–S4–S1–S3–D3–S1–S4–S1–S2–S1–D1–S2–S4–S3–S4–S2–S4. This sound example is available online at the publisher's website and at http://www.fleurbouwer.nl/publications.

pass FIR filters. For seven participants, one bad channel was removed and replaced by values interpolated from the surrounding channels. None of these channels is included in the statistical analysis reported here. Independent component analysis as implemented in EEGLAB was conducted to remove eye blinks. For the deviant patterns (D1–D3) and the three standard patterns containing omissions (S2–S4), epochs of 800 ms were extracted from the continuous data starting 200 ms before the onset of the omission. Epochs with an amplitude change of more than 75 µV in a 500 ms window on any channel were rejected. Finally, epochs were baseline corrected by the average voltage of the 200 ms prior to the onset of the omission and averaged to obtain ERPs for omissions in each position for each participant.

The omissions in the various patterns could be preceded by a bass drum sound (D3 and S2), a snare drum sound (S3 and S4) or a hi-hat sound (D1 and D2). To control for the possible effects of this contextual difference we calculated difference waves. For all patterns containing omissions, from the ERP obtained in response to the omissions we subtracted the temporally aligned ERP obtained from base pattern S1. This procedure yielded difference waves for each participant that were thought to reflect only the additional activity elicited by the omission in that particular position.

Visual inspection of the group averaged difference waves showed negative deflections peaking between 100 and 200 ms after the onset of each omission with a frontocentral maximum. This is consistent with the latency and scalp distribution of the MMN (Näätänen et al., 2007). Hence, MMN latencies were subsequently defined as the negative peak on electrode FCz between 100 and 200 ms. Single subject amplitudes were defined for each condition as the average amplitude in a 60 ms window around the condition specific peaks obtained from the group averaged difference waves.

The group averaged difference waves also showed positive deflections consistent in latency and scalp distribution with a P3a (Polich, 2007). However, in the latency range of the P3a the ERPs could possibly contain contributions from activity related to the tone following the omission, which occurred 150 ms after the omission. While the use of difference waves might eliminate some of this activity, the tones following an omission could possibly elicit an enhanced N1 response due to fresh afferent neuronal activity. This additional activity may be absent in the ERPs for S1, which we used to obtain the difference waves and thus would not be eliminated by the subtraction procedure. Due to the different sounds following the omissions in the deviants (Figure 5.1), such an effect would be different for each deviant. Differences between the ERPs in the latency range of the P3a are thus hard to interpret. Therefore, here we will only consider the MMN results.

5.2.7 Statistical analysis

To confirm that the MMN peaks were significantly different from zero, we performed T-tests on the MMN amplitudes for each condition separately on electrode FCz. Our primary interest concerned the difference in response to omissions in the deviant patterns, to evaluate the effects of metrical position and musical expertise. Thus, first we compared the amplitude and latency of the MMN response to the omissions in the deviant patterns in a repeated measures ANOVAs, with position (D1, D2, D3) as a

within subject factor and musical expertise (musician, non-musician) as a between subject factor. In addition, to examine the effects of using acoustically varying stimuli we compared the MMN responses to omissions in D3, S2, S3 and S4 in ANOVAs with the same structure. Greenhouse-Geisser corrections were used when the assumption of sphericity was violated. For significant main effects, Bonferroni-corrected post hoc pairwise comparisons were performed. The statistical analysis was conducted in SPSS (Version 20.0). We report all effects that are significant at $p < 0.05$.

5.3   Results

Table 5.1 shows the average mean amplitudes and peak latencies of the MMN for omissions in all patterns. T-tests confirmed that the amplitudes of the negative peaks in the difference waves between 100 and 200 ms from the onset of the omissions were significantly different from zero for both musicians and non-musicians and for omissions in all positions (all p values < 0.001), showing that an MMN was elicited by all omissions.

5.3.1   Response to omissions in deviant patterns

Figure 5.3 shows the group averaged ERPs and difference waves for omissions in the three deviant patterns (D1, D2 and D3) for electrode FCz for both musicians and non-musicians. The position of the omissions in the deviant patterns had a significant effect on both the amplitude ($F_{(2,54)} = 19.4$, $p < 0.001$, $\eta^2 = 0.42$) and the latency ($F_{(2,54)} = 24.0$, $p < 0.001$, $\eta^2 = 0.47$) of the MMN. Post hoc pairwise comparisons revealed that this was due to the MMN to the omissions in D3 being smaller in amplitude and earlier in latency than the MMN to the omissions in both D1 and D2 (all p values < 0.001). The amplitudes of the responses to omissions in D1 and D2 did not differ from each other (amplitude, $p = 0.191$; latency, $p = 1.000$). Neither the effect of musical expertise (amplitude, $F_{(1,27)} = 0.21$, $p = 0.647$, $\eta^2 = 0.008$; latency, $F_{(1,27)} = 0.42$, $p = 0.521$, $\eta^2 = 0.015$) nor the interaction between musical expertise and position (amplitude, $F_{(2,54)} = 0.09$, $p = 0.911$, $\eta^2 = 0.003$; latency, $F_{(2,54)} = 2.37$, $p = 0.103$, $\eta^2 = 0.081$) was significant.

Table 5.1 Mean average amplitudes and average peak latencies of the MMN to omissions in all conditions. Note: Standard deviations in brackets.

Omission	Average Amplitude (µV)		Average Peak Latency (ms)	
	Musicians (N = 14)	Non-musicians (N = 15)	Musicians (N = 14)	Non-musicians (N = 15)
D1	−3.49 (1.43)	−3.70 (1.96)	146 (22)	142 (19)
D2	−3.12 (1.18)	−3.26 (1.73)	144 (16)	148 (16)
D3	−2.05 (1.26)	−2.38 (1.14)	129 (21)	117 (17)
S2	−1.55 (0.64)	−1.64 (0.86)	136 (17)	135 (19)
S3	−1.09 (0.69)	−0.97 (0.79)	151 (33)	157 (37)
S4	−1.15 (0.75)	−1.03 (0.76)	136 (28)	157 (31)

Musicians **Non-musicians**
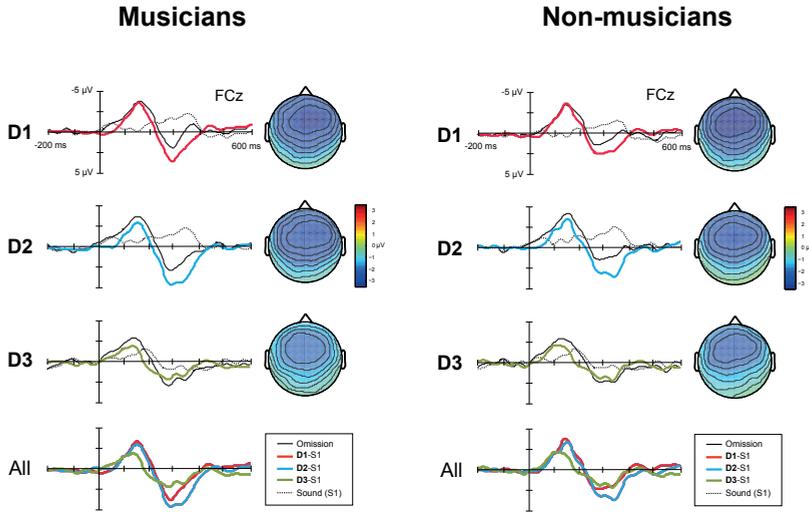
Figure 5.3 ERP responses for D1, D2 and D3 for musicians (N = 14, left) and non-musicians (N = 15, right). The panels labeled D1, D2 and D3 show the group averaged ERPs for electrode FCz elicited by omissions, the corresponding position in S1, the derived difference waves and the scalp distributions of the difference waves. The panel labeled *All* shows all difference waves combined. Time 0 is the onset of the omission, or, in the case of S1, the onset of the corresponding sound. The omissions in D1, D2 and D3 were equally rare in occurrence (0.033) and in all cases, a bass drum sound was omitted.

5.3.2   Response to omissions in metrically weak positions

Figure 5.4 shows the ERPs elicited by all omissions in metrically weak positions (in patterns D3, S2, S3 and S4). The amplitude and latency of the MMN were significantly affected by the position of the omissions (amplitude, $F_{(3,81)} = 25.4$, $p < 0.001$, $\eta^2 = 0.48$; latency, $F_{(3,81)} = 9.99$, $p < 0.001$, $\eta^2 = 0.27$) but not by the factor musical expertise (amplitude, $F_{(1,27)} = 0.03$, $p = 0.864$, $\eta^2 = 0.001$; latency, $F_{(1,27)} = 0.31$, $p = 0.580$, $\eta^2 = 0.012$) or an interaction between musical expertise and position (amplitude, $F_{(3,81)} = 0.96$, $p = 0.415$, $\eta^2 = 0.034$; latency, $F_{(3,81)} = 2.37$, $p = 0.077$, $\eta^2 = 0.081$).

Post hoc pairwise comparisons revealed that the significant effect of position on MMN amplitude was due to the MMN to omissions in D3 being larger in amplitude than the MMN to omissions in S2 ($p = 0.002$), S3 ($p < 0.001$) and S4 ($p < 0.001$). Interestingly, the amplitude of the MMN to the omissions in standard S2 was significantly larger than the amplitude of the MMN to the omissions in standards S3 ($p = 0.005$) and S4 ($p = 0.011$). Finally, the MMN to omissions in D3 was earlier in latency than the MMN to omissions in S2 ($p = 0.040$), S3 ($p = 0.001$) and S4 ($p = 0.001$).
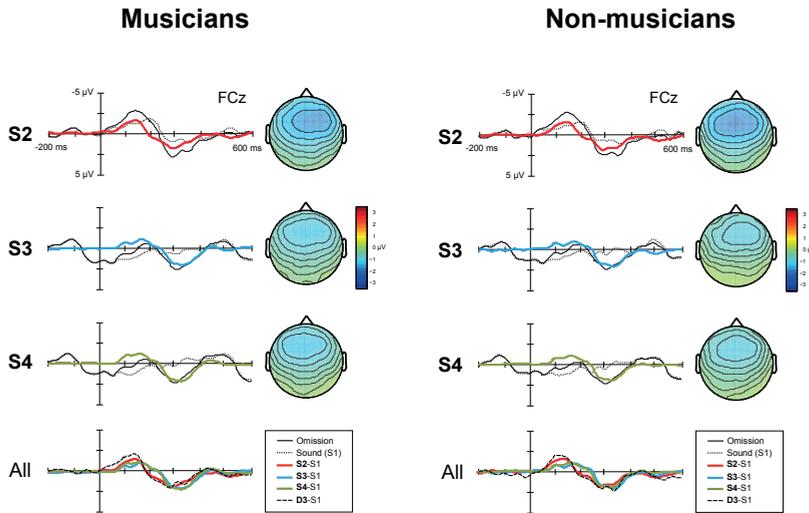
Figure 5.4 ERP responses for S2, S3 and S4 for musicians (N = 14, left) and non-musicians (N = 15, right). The panels labeled S2, S3 and S4 show the group averaged ERPs for electrode FCz elicited by omissions in the standards, the corresponding position in S1, the derived difference waves and the scalp distributions of the difference waves. The panel labeled *All* shows all difference waves combined. Time 0 is the onset of the omission, or, in the case of S1, the onset of the corresponding sound. The omissions in S2, S3 and S4 were equally rare in occurrence (0.225) and in all cases, a hi-hat sound was omitted. For clarity, here we add the difference wave for D3 (see Figure 5.3 for the separate ERPs) to make a comparison with the difference waves derived for the standards possible. The omissions in D3 were in equally weak metrical positions as in S2, S3 and S4.

5.4 Discussion

The data show that the MMN responses to omissions on the beat (D1, D2) were larger in amplitude than the MMN response to omissions in a metrically weak position (D3), indicating that the former, which violated the metrical structure, were processed as more salient than the latter, which left the metrical structure intact (Figure 5.3). The omissions could not be differentiated from each other based on their acoustic characteristics, suggesting that auditory system of the participants detected the beat pre-attentively.

Each pattern encompassed two beats. To examine whether participants detected a hierarchy between the two beats, we compared the MMN responses to omissions on the first (D1) and second (D2) beat (Figure 5.3). We found no differences in amplitude or latency, suggesting that processing of meter – higher order regularity in the form of alternating stronger and weaker beats – is not pre-attentive. However, while the lack of an effect of the position of the beat may be indicative of a true absence of meter perception, two caveats must be noted. First, the MMN amplitude for omissions in both D1 and D2 was very large (< -3 µV) and maybe near ceiling, as it might contain

the additive effects of multiple regularity violations, not only violations of the metrical structure, but also violations of the acoustic regularity (see below). This may have caused the tendency towards larger amplitude responses to D1 than D2, present in both musicians and non-musicians, not to reach significance. Second, while we assumed that the pattern was perceived as two consecutive beats, with D1 containing an omission on the first beat and D2 containing an omission on the second beat, the patterns in fact did not contain any accents indicating a hierarchy between a first and second beat. Therefore, it is possible that some participants processed the fifth position in the pattern as the first beat and the first position as the second beat. To address these issues and to examine meter processing, a paradigm more specifically tuned to inducing and measuring a hierarchy between beats is needed.

The MMN responses of musicians and non-musicians did not differ (Figure 5.3; Table 5.1). Thus, not only may beat processing not require attention, but also it may be independent of musical expertise. Our findings are in contrast with earlier studies proposing a role for both attention (Chapin et al., 2010; Geiser et al., 2009) and expertise (Geiser et al., 2010) in beat processing. These conclusions were based on experiments in which the beat was marked only by temporal variation in the surface structure of the rhythm. In the current study, acoustically more varied stimuli were used, in which the beat was marked by both the surface structure of the rhythm and timbre and intensity differences. Arguably, the additional information contained in the acoustic properties of the sounds may make it easier to induce a beat, as accents are simply indicated by intensity differences and do not have to be deduced from the temporal organization of the rhythm. Therefore, we propose that conflicting findings regarding the role of attention and musical expertise in beat processing may be explained by looking at the temporal and acoustic complexity of the musical stimuli.

This view is further supported by studies suggesting that the use of real music leads to bigger effects of beat processing than the use of more abstract sequences of tones (Bolger et al., 2013; Tierney & Kraus, 2013), which may also be attributable to the real music containing multiple clues for the metrical structure. Finally, in a study directly comparing beat processing with only temporal accents and beat processing with only intensity accents it was suggested that the latter required less internal effort than the former (Grahn & Rowe, 2009). Together with our results, these findings stress the importance of using more acoustically varied stimuli when testing beat processing. The use of highly abstract sequences of tones, with only variation in the temporal organization of the rhythm, may result in an underestimation of the beat processing abilities of untrained individuals.

While attention and expertise did not seem to affect beat processing with the current, highly beat inducing stimuli, we cannot rule out that beat processing, especially when more complex stimuli are used, is mediated to some extent by attention and expertise. However, our results support the view that for metrically simple, acoustically varied music-like rhythms, beat processing is possible without attention or expertise and may indeed be considered a very fundamental human ability (Honing, 2012).

To examine, exploratory, possible effects of acoustically rich stimuli on ERPs we compared the responses to omissions that varied acoustically but were all in metrically equally weak positions. As in each pattern only one out of eight tones was omitted, all these omissions could be considered rare events within a pattern, and as such, elicited an MMN (Figure 5.4). The comparison between these MMN responses yielded two interesting effects. First, the MMN to omissions in pattern D3 was larger in amplitude than the MMN to omissions in the standard patterns (S2, S3 and S4). As it is known that low probability events cause higher amplitude MMN responses (Sabri & Campbell, 2001), this was presumably due to the omission of a bass drum sound, as in D3, being more rare than the omission of a hi-hat sound, as in S2, S3 and S4. Interestingly, to detect this probability difference, not only acoustic information but also information about the sequential order of the sounds is required. Thus, the auditory system formed a representation at the level of the complete pattern. This is consistent with the view that patterns as long as 4 seconds can be represented as a whole by the MMN system, whilst this system can operate at multiple hierarchical levels, representing both patterns and sounds within patterns simultaneously (Herholz, Lappe, & Pantev, 2009).

Second, unexpectedly, the amplitude of the MMN to omissions in S2 was larger than the amplitude of the MMN to omissions in S3 and S4 (Figure 5.4). These omissions were all in metrically weak positions and in all cases a hi-hat sound was omitted. However, in S2, the omissions followed a bass drum sound, while in S3 and S4 the omissions followed a snare drum sound (Figure 5.1). While we used difference waves to eliminate any direct effects of the acoustic context on the waveforms, the sounds preceding the omissions may have affected the MMN response indirectly by affecting the regularity representation (Sussman, 2007) through forward masking (Carlyon, 1988). Forward masking decreases with an increasing interval between the masking sound and the masked sound, the masker-signal delay (Zwicker, 1984). Thus, the hi-hat sounds in positions four and eight, which immediately followed the snare drum sound with a delay of 0 ms, may have been perceptually less loud than the hi-hat sound in position two, which followed the bass drum sound with a delay of 50 ms. The omission of the former, in S3 and S4, may therefore have been perceived as acoustically less salient than the omission of the latter, in S2, explaining the difference in MMN amplitude.

The presence of this effect could potentially weaken our conclusions regarding pre-attentive beat processing, as the acoustic context of the omissions in D1 and D2, following a hi-hat sound with a delay of 100 ms, differed from the acoustic context of the omissions in D3, following a bass drum sound with a delay of 50 ms. However, it has been shown that increases in masker-signal delay affect the magnitude of masking nonlinearly, with more rapid decreases in masking at smaller masker-signal delays than at larger masker-signal delays (Dau, Püschel, & Kohlrausch, 1996; Zwicker, 1984). Therefore, any effect of masking on the MMN responses to omissions in D1, D2 and D3, with delays from 50 to 100 ms, should be the same or smaller than the effect of masking on the MMN responses to omissions in S2, S3 and S4, with delays from 0 to 50 ms. Yet the difference between the MMN responses to omissions in D3 and in D1 and D2 was much larger than the difference between the MMN responses to omissions in S2 and in S3 and S4. Consequently, contextual differences alone are

unlikely to account for the difference between the response to omissions on the beat (D1 and D2) and omissions in metrically weak positions (D3).

To summarize, the differences in the responses to acoustically varying omissions in metrically weak positions show how the same sound differences that allow people to perceive a beat can cause difficulty in the interpretation of ERP results. Here, we controlled for these acoustic differences and show that adults differentiate pre-attentively between omissions in different metrical positions, based solely on their position. However, our results suggest that some caution has to be taken in interpreting earlier results in newborns (Winkler et al., 2009). It is unclear whether newborns, like adults in the current study, detected the beat solely based on its position in the rhythm. While not in conflict with these previous findings (Winkler et al., 2009), our results do suggest the need for additional testing to fully confirm their conclusions.

The use of acoustically rich stimuli can be advantageous when testing beat processing (Bolger et al., 2013; Tierney & Kraus, 2013). One way of addressing the possible pitfalls associated with such stimuli is by improving stimulus design, as in the current study. Alternatively, beat processing can be probed with alternative methods, which perhaps are less sensitive to acoustic factors than ERPs. Promising results have been obtained by looking at neural dynamics (Fujioka et al., 2012; Snyder & Large, 2005) and steady-state potentials (Nozaradan et al., 2011, 2012), but so far only using simple isochronous or highly repetitive sequences. Combining these methods with acoustically rich and temporally varied stimuli may provide valuable information about beat processing and warrants further research.

5.5   Conclusions

We have provided evidence suggesting that beat processing with metrically simple and acoustically varied stimuli does not require attention or musical expertise. Furthermore, we have shown that the MMN response to omissions in a rhythm is indeed sensitive to metrical position and as such can be a useful tool in probing beat processing, even if acoustically varied stimuli are used. Our conclusions are in line with previous findings in adults (Ladinig et al., 2009, 2011) and newborns (Winkler et al., 2009). However, we also showed that the ability of the listener to recognize longer patterns and the acoustic context of an omission can influence the ERP response to sound omissions in a rhythm. While the present results are not in conflict with previous findings, controls for these issues were lacking in earlier experiments (Honing et al., 2012; Ladinig et al., 2009, 2011; Winkler et al., 2009). To be certain that any effects observed are due to metrical position and not pattern matching or acoustic variability, future experiments will have to take these factors into account. At the same time, if sufficiently controlled, the use of stimuli with acoustic variability may be a big advantage when testing beat processing.

The current study thus not only contributes to the growing knowledge on the functioning of beat processing, it also nuances findings that were novel and exciting, but that are in need of additional testing to be fully confirmed. As such, the current study fits in a general trend that stresses the importance of replication in psychological research (Carpenter, 2012; Pashler & Wagenmakers, 2012).

Chapter 6

Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm[*]

Beat perception is the ability to perceive temporal regularity in musical rhythm. When a beat is perceived, predictions about upcoming events can be generated. These predictions can influence processing of subsequent rhythmic events. However, statistical learning of the order of sounds in a sequence can also affect processing of rhythmic events and must be differentiated from beat perception. In the current study, using EEG, we examined the effects of attention and musical abilities on beat perception. To ensure we measured beat perception and not absolute perception of temporal intervals, we used alternating loud and soft tones to create a rhythm with two hierarchical metrical levels. To control for sequential learning of the order of the different sounds, we used temporally regular (isochronous) and jittered rhythmic sequences. The order of sounds was identical in both conditions, but only the regular condition allowed for the perception of a beat. Unexpected intensity decrements were introduced on the beat and offbeat. In the regular condition, both beat perception and sequential learning were expected to enhance detection of these deviants on the beat. In the jittered condition, only sequential learning was expected to affect processing of the deviants. ERP responses to deviants were larger on the beat than offbeat in both conditions. Importantly, this difference was larger in the regular condition than in the jittered condition, suggesting that beat perception influenced responses to rhythmic events in addition to sequential learning. The influence of beat perception was present both with and without attention directed at the rhythm. Moreover, beat perception as measured with ERPs correlated with musical abilities, but only when attention was directed at the stimuli. Our study shows that beat perception is possible when attention is not directed at a rhythm. In addition, our results suggest that attention may mediate the influence of musical abilities on beat perception.

6.1   Introduction

The perception of a regular beat in music allows us to predict the timing of musical events and thus to synchronize and dance to music together, activities that may be crucial in understanding the origins of musicality (Honing et al., 2015). A musical beat can be defined as a regularly recurring salient moment in time (Cooper & Meyer, 1960) and is the regularity in music that we clap and dance to. The hierarchical structure of more and less salient moments in time is referred to as the metrical structure. Often, metrical salience in the form of a beat coincides with musical salience in the form of an accented event (Honing et al., 2014). However, once a beat is perceived, its perception can remain stable even if accents locally do not conform to the metrical structure. Thus, a perceived beat is a psychological construct and not necessarily physically present in a stimulus (Merchant et al., 2015).

Beat perception has been explained by Dynamic Attending Theory (DAT) as regular fluctuations in attentional resources, peaking at metrically salient positions (Large & Jones, 1999). Computationally and at a neural level, DAT has been linked to oscillator models (Henry & Herrmann, 2014; Large, 2008), with multiple oscillators present for multiple levels of regularity in a metrical hierarchy. When listening to music, internal oscillators entrain to the external regularity in a rhythm (Drake, Jones, et al., 2000), and this allows a listener to generate precise temporal predictions about the occurrence of rhythmic events (Large, 2000; Phillips-Silver, Aktipis, & Bryant, 2011). Beat perception has been shown to be mediated by motor networks in the brain, and specifically the basal ganglia (Grahn & Brett, 2007). These motor areas are active during beat perception even when no movement is involved (Merchant et al., 2015). This suggests that the mere perception of a beat relies on interactions between auditory and motor areas in the brain (Zatorre, Chen, & Penhune, 2007). One of the hypothesized roles of the motor areas in beat perception is the generation of temporal predictions (Grahn & Rowe, 2013; Merchant et al., 2015).

The predictions generated by a perceived beat not only allow for synchronization of movement to a beat, but can also affect processing of rhythmic events within a metrical structure. When predictions are generated about upcoming events, processing of auditory events that violate these predictions is enhanced, as is evidenced by three ERP components that have been specifically linked to processing of unexpected auditory events: the mismatch negativity (MMN), the N2b and the P3a. The larger the violation of expectations, the larger is the amplitude of these components (Näätänen et al., 2007; Polich, 2007). As such, these components provide a very useful way to examine beat perception. The perception of a beat leads to the prediction of events on the beat, while no events or softer events are predicted offbeat (Bouwer & Honing, 2015; Large, 2000). A perceived metrical structure can be probed by violating these predictions and measuring the ERP responses to prediction violations (Honing et al., 2014).

Earlier, using the strategy described above, we examined beat perception by comparing the ERP responses to silences on the beat, where they are unexpected, and offbeat, where they are more expected, and we showed that beat perception is independent of attention or explicit musical training (Bouwer et al., 2014). However, in studies using

a similar approach it has been argued that attention is necessary to perceive temporal regularity in an auditory sequence (Geiser et al., 2009; Schwartze et al., 2011) and that musical training enhances the perception of a beat (Geiser et al., 2010). These conflicting findings may be due to the differences in materials used in these studies, ranging from stimuli resembling real music (Bouwer et al., 2014), to rhythms with a varying temporal pattern but with identical sounds (Geiser et al., 2010, 2009), to monotonous isochronous sequences (Schwartze et al., 2011). Many tasks aimed at measuring beat perception can in fact be accomplished by recruiting mechanisms that are not related to beat perception per se (Tranchant & Vuvan, 2015). In natural music, there is an abundance of cues indicating the metrical structure. This may additionally lead to recruitment of mechanisms related to beat perception that are not used when listening to an isochronous sequence. To understand how attention and musical training influence the perception of a beat, disentangling beat perception from other mechanisms (i.e., those that may contribute to or interact with beat perception) may be crucial.

First, it is important to note that beat perception relies on the perception of the relative proportions of the time intervals that make up a rhythm (Honing, 2013; Leow & Grahn, 2014). Relative or beat-based perception of rhythm is considered distinct from the perception of absolute time intervals in rhythm (Merchant & Honing, 2014; Teki et al., 2011). To separate beat-based perception from absolute interval perception, several studies have compared the responses to temporally regular, isochronous sequences with the responses to temporally irregular, jittered sequences (Fujioka et al., 2012; Schwartze et al., 2011; Teki et al., 2011). The prediction of events in jittered sequences has been suggested to rely on absolute interval perception, while the prediction of events in isochronous sequences has been suggested to recruit beat-based perception (Fujioka et al., 2012; Schwartze et al., 2011). However, humans can predict a sequence of temporal intervals relying solely on absolute interval perception, as is apparent from the possibility for humans to reproduce rhythms that do not contain a beat at all (Cameron & Grahn, 2014). A similar phenomenon is observed in nonhuman primates. While macaques have little or no ability to perceive a beat (Honing et al., 2012; Merchant & Honing, 2014), they respond more accurately to temporally regular than jittered sequences, suggesting a capacity for making temporal predictions (Zarco et al., 2009), which most likely depends on absolute interval perception (Merchant & Honing, 2014). Thus, it cannot be ruled out that humans, like macaques, can predict temporal intervals in an isochronous sequence based on absolute interval perception. Differences between responses to regular and jittered sequences (as reported by Fujioka et al., 2012; Schwartze et al., 2011) may be caused by enhanced predictions generated through absolute interval perception when temporal variability of a sequence is low. Therefore, the use of isochronous sequences may not be optimal for examining beat perception, as it is unclear whether the prediction of events in an isochronous sequence depends on beat-based perception, absolute interval perception, or both. To ensure that beat perception is measured, and not absolute interval perception, it is necessary to introduce some level of hierarchy in a rhythm to create a metrical structure. The perceived metrical structure can then be probed by comparing responses to events in different metrical positions, which differ in metrical salience, but have the same temporal properties.

One often-used way of introducing metrical hierarchy in a rhythm is by varying the temporal structure of the rhythm, while keeping all sounds identical. The temporal grouping of events in a rhythm can induce perceptual accents, which, if regularly spaced in time, can induce a beat (Povel & Essens, 1985). In two studies using such a non-isochronous rhythm with temporal accents, Geiser et al. (2009, 2010) found that ERP responses to unexpected intensity increases were larger offbeat than on the beat. Interestingly, in one of the studies (2009), this effect was only present when attention was directed towards the stimuli, while in the other (2010), the effect was also present when attention was directed away. Moreover, in the first study (2009), no effect of musical training was found, while in the second study (2010), musical training enhanced the difference between responses to events on the beat and offbeat. Thus, it is unclear how attention and musical abilities affect responses to non-isochronous rhythms with temporal accents. In an fMRI study using both rhythms with temporal accents and rhythms with acoustic cues indicating the metrical structure, Grahn and Rowe (2009) found that musicians showed more connectivity between premotor areas and auditory cortex than non-musicians, but only for the rhythms with temporal accents. This suggests that musical training may enhance the perception of a beat in rhythms when information about the metrical structure is only present in the temporal grouping of events. Acoustic cues to the beat as in real music may help especially musical novices to extract a beat and may thus be important to use when testing beat perception in musical novices.

In studying beat perception with more natural stimuli, such acoustic cues can be used to indicate the salience of events and thus to induce a hierarchical metrical structure (Ellis & Jones, 2009; Honing et al., 2014), ensuring that predictions cannot be solely made by relying on absolute interval perception. However, apart from being regularly spaced in time, metrical accents may also exhibit statistical regularity in the order of different events, which can influence the expectations of auditory events. To ensure that beat perception is measured when examining responses to rhythm, beat perception should thus be differentiated from statistical learning of the order of events in a rhythmic sequence (hereafter: *sequential learning*). For example, in the highly beat inducing sequences used by Bouwer et al. (2014), a comparison was made between ERP responses to unexpected omissions of events on the beat and offbeat. Beat perception was hypothesized to lead to strong expectations for the occurrence of events on the beat, making omissions on the beat less expected than omissions offbeat. In line with this, larger responses to omissions on the beat than offbeat were found. However, the patterns of bass drum, hi-hat and snare drum sounds that were used to induce a beat exhibited statistical regularity in the order and the transitional probabilities of the different sounds. While the probability of an omission in general was relatively small, the probability of a hi-hat sound being followed by an omission was smaller (0.029) than the probability of a bass drum sound being followed by an omission (0.089). As an omission on the beat always followed a hi-hat sound and an omission offbeat always followed a bass drum sound, it could be that the omissions on the beat were less expected than the omissions offbeat not only because of metrical expectations, but also because of differences in transitional probabilities. Humans possess the ability to learn such transitional probabilities in both linguistic (Saffran, Aslin, & Newport, 1996) and non-linguistic sequences (Saffran, Johnson, Aslin, & Newport, 1999; Tillmann &

McAdams, 2004). In addition, learning of the statistical properties of sequences is possible, in principle, without attention (Schröger, Bendixen, Trujillo-Barreto, & Roeber, 2007; Van Zuijen, Simoens, Paavilainen, Näätänen, & Tervaniemi, 2006). Thus, one can argue that sequential learning rather than beat perception may have influenced responses to rhythms in previous studies (e.g., Bouwer et al., 2014; Ladinig et al., 2009; Vuust et al., 2005, 2009; Winkler et al., 2009).

In the current study we aimed to confirm previous findings showing that beat perception is independent of attention and musical training. We used rhythms with multiple acoustic cues indicating the metrical structure to facilitate beat perception for musical novices. We explicitly sought to disentangle beat perception from sequential learning, which may have biased results in previous studies (Bouwer et al., 2014; Winkler et al., 2009). Moreover, we used stimuli with a hierarchical structure to ensure that we measured beat perception and not absolute interval perception. We used a binary rhythmic pattern with alternating loud bass drum and softer hi-hat sounds indicating accented beats and unaccented offbeats. The bass drum and hi-hat sounds differed not only in intensity, but also in length and timbre, providing many cues for the listener to differentiate accented beats from unaccented offbeats. The alternating accented and unaccented sounds created a pattern with two metrical levels, the beat and subdivisions of the beat. We measured ERP responses to unexpected deviant tones in the form of intensity decrements on the beat and offbeat, both while participants were actively attending to the rhythm and while they directed their attention to a silent movie. Specifically, we were interested in the N2b response, which is recorded when people attend to a stimulus, and the MMN and P3a responses, which are recorded both under attended and unattended conditions. Intensity decrements are less expected on the beat than offbeat. Thus, when a beat is perceived, these ERP components, that are known to index the magnitude of a regularity violation (Näätänen et al., 2007; Polich, 2007) are expected to be larger in response to intensity decrements on the beat than offbeat (Bouwer & Honing, 2015; Potter et al., 2009).

ERPs are highly sensitive to the preceding acoustic context (Bouwer et al., 2014; Honing et al., 2014; Woldorff & Hillyard, 1991). Also, if a loud bass drum sound were always followed by a softer hi-hat sound and vice versa, a soft sound may be statistically more expected after a bass drum sound than after a hi-hat sound, making the comparison of responses to intensity decrements on the beat and offbeat biased. To avoid both acoustic and statistical effects of contextual differences, we frequently introduced bass drum sounds offbeat. This allowed deviants on the beat to not only be identical in sound to deviants offbeat, but also, like the deviants offbeat, to be preceded and followed by bass drum sounds.

While the bass drum sounds offbeat ensured that the transitional probabilities of consecutive sounds were the same for both deviants, louder sounds were statistically still more probable in odd positions (on the beat) and softer sounds in even positions (offbeat). Learning of this statistical regularity in the order of sounds may lead to larger ERP responses to intensity decrements in odd than in even positions regardless of beat perception. To disentangle beat perception from such an effect of sequential learning, we contrasted the responses to deviants in regular sequences, in which all inter-onset

intervals were the same, with responses to deviants in jittered sequences, in which the inter-onset intervals were irregular. The statistical regularity in terms of the order of the different sounds was identical in the regular and jittered conditions. However, beat perception was only possible in the regular condition, but not in the jittered condition. We expected sequential learning of the pattern of alternating loud bass drum and softer hi-hat sounds to lead to larger ERP responses to deviants in odd than in even positions regardless of the temporal regularity of the sequence. If beat perception were present, we would expect this difference to be more pronounced in the regular than in the jittered condition, as beat perception would make the expectation for a loud event on the beat (in an odd position) even stronger. Thus, both in attended and unattended conditions, if a beat were perceived we would expect an interaction between the regularity of the sequence and the position of the deviant.

People vary widely in their ability to perceive a beat (Grahn & Schuit, 2012) and while this ability is highly correlated with musical training, it is possible for non-musicians to be extremely apt at hearing a beat in music. Previously, only the effect of musical training on beat perception was examined (Bouwer et al., 2014). However, there might be differences in beat perception abilities independent of musical training, with both musicians and non-musicians varying in how sensitive they are to a beat. Recently, a test battery has become available to get an estimate of musical abilities in the general population (Goldsmith Musical Sophistication Index, or Gold-MSI; Müllensiefen, Gingras, Musil, & Stewart, 2014). To separate the effects of formal instruction from those caused by a predisposition for beat perception, here we correlated beat perception as measured with ERPs in attended and unattended conditions with scores on both musical training and beat perception ability as measured with the Gold-MSI.

6.2 Methods

6.2.1 Participants

Thirty-four participants (23 women) took part in the experiment. They were on average 25.6 years old (SD 5.2 years, range 19–45 years). Their musical training ranged from no formal lessons at all to training as a professional musician. On average, they had 9.7 years of instrumental lessons (SD 9.6 years, range 0–34 years). None of the participants reported a history of neurological or hearing disorders. All participants provided written informed consent prior to the study. The study was approved by the Ethics Committee of the Faculty of Humanities at the University of Amsterdam.

6.2.2 Materials

Goldsmith Musical Sophistication Index
To assess the overall musical training received by our participants, we used the Gold-MSI questionnaire (Müllensiefen et al., 2014). This questionnaire is designed to index musical sophistication in the general population and contains several subscales, including a subscale for musical training. In addition to instrumental lessons, this subscale also takes into account theory lessons, amount of practice, and number of instruments played. While highly correlated with the absolute years of music lessons received, the Gold-MSI provides us with a more nuanced measure of musical training. Both the

original questionnaire and a Dutch translation were used, to accommodate both Dutch participants and those who did not speak Dutch. For each participant we obtained a score for the musical training subscale. For details concerning the questionnaire and data norms, we refer to Müllensiefen et al. (2014).

Beat Alignment Test

To assess beat perception abilities, we used the beat alignment perception test (BAT) as implemented by Müllensiefen et al. (2014) and conceived by Iversen and Patel (2008). In this test, participants are required to listen to clips of music with overlaid metronome beeps. The metronome is either on the beat, has a slightly different tempo, or is shifted in phase. Participants are asked to judge whether the metronome is on the beat or not. The test contains 17 items and 3 practice items, with varying musical genres. For each participant, an accuracy score was calculated. Accuracy scores of 0.5 or lower show performance at chance level and were replaced by a value of 0.5, as performance below chance is not informative. For details of the music used in the test, see Müllensiefen et al. (2014).

Stimuli

Rhythmic sequences were created using two standard sounds. The first was a combination of simultaneously sounding bass drum and hi-hat sounds (for simplicity we will refer to these as bass drum sounds), and the second consisted of only a hi-hat sound. Both sounds were created using QuickTime's drum timbres (Apple Inc.). Bass drum sounds were longer (110 vs. 70 ms) and louder (16.6 dB difference in volume) than hi-hat sounds and as such were expected to be perceived as more salient than hi-hat sounds. Additional bass drum sounds attenuated with 25 dB (using Praat software; http://www.praat.org) were used as deviants. Four different two-tone configurations were constructed from these three sounds (see Figure 6.1). The majority of the patterns (60%) consisted of a bass drum sound followed by a hi-hat sound (standard pattern S1; see Figure 6.1A). A second pattern was constructed from two consecutive bass drum sounds (standard pattern S2, 30% of all patterns, see Figure 6.1A). Two deviant patterns were used; one consisting of a deviant sound followed by a bass drum sound (deviant pattern D1; 5% of all patterns), and one with a bass drum sound followed by a deviant sound (deviant pattern D2; 5% of all patterns, see Figure 6.1A).

The four patterns were concatenated to create continuous sequences for both the regular and jittered conditions (Figure 6.1B). In the regular condition, all single tones were presented with an inter-onset interval of 225 ms. In this condition, the alternating salient bass drum sounds and less salient hi-hat sounds as occurring in pattern S1 were expected to induce a beat with an inter-beat interval of 450 ms, within the optimal range for beat perception in humans (Drake, Jones, et al., 2000; London, 2012). In the regular condition, all sounds in the first position of a pattern, including deviant $D1_r$, can be considered on the beat, while all sounds in the second position, including deviant $D2_r$, are offbeat. In the jittered condition, the inter-onset intervals in the standard patterns were randomly distributed between 150 and 300 ms (flat distribution), which made beat perception impossible. The inter-onset interval before and following a deviant tone was kept constant at 225 ms. Note that we will refer to the deviants in the jittered context as on the beat ($D1_j$) and offbeat ($D2_j$), even though no beat can be heard
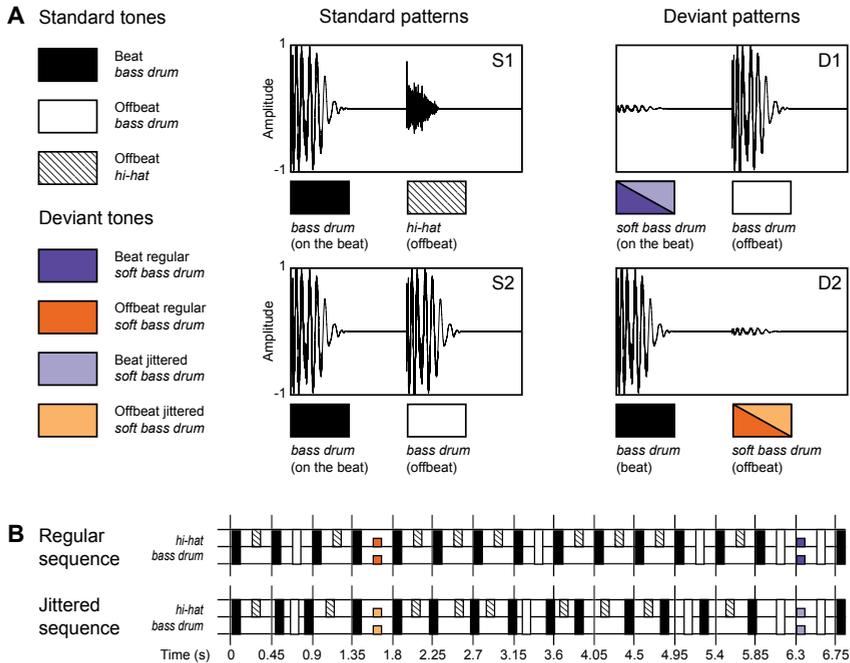
Figure 6.1 Schematic overview of the stimuli. A) Three different sounds were used to create two standard and two deviant patterns. The bass drum sound could occur in two different positions, both on the beat and offbeat. The hi-hat sound only occurred offbeat. An attenuated bass drum sound was used as deviant sound in two different positions, both on the beat and offbeat, and in two conditions, regular and jittered. B) Patterns were concatenated into sequences. In the regular sequence, all inter-onset intervals were equal at 225 ms. In the jittered sequence, inter-onset intervals ranged from 150 to 300 ms. The inter-onset intervals before and after the deviant sounds were always fixed at 225 ms and deviants were always preceded and followed by a bass drum sound. Thus, acoustically, all four deviants and their contexts were identical.

in this condition, to clarify their relationship with the deviants in the regular context (D1$_r$ and D2$_r$).

In both the regular and the jittered condition, the concatenation of patterns was semi-randomized with four constraints on the randomization. First, to optimize beat perception in the regular condition, pattern S2, which contained a bass drum on the offbeat and did not contribute to the perception of the metrical hierarchy, was never presented more than once consecutively. Second, a maximum of four consecutive S1 patterns was allowed. Third, a deviant on the beat (D1) always followed a bass drum sound offbeat (S2). Finally, there were always at least five standard patterns between two deviant patterns. Note that for all four conditions of interest (two types of regularity and two metrical positions) the deviants (D1$_r$, D1$_j$, D2$_r$, D2$_j$) were preceded and followed by a bass drum sound with inter-onset intervals of 225 ms, creating identical

acoustic contexts. For schematic examples of both the regular and jittered sequences, see Figure 6.1B and Supplementary Sound 1 (regular) and 2 (jittered)[5].

6.2.3 Procedure

Stimuli were presented in five-minute blocks consisting of five sequences of 54 seconds (120 patterns per sequence; 600 patterns per block). Regular and jittered blocks were presented in semi-random order, with a maximum of two blocks from the same condition following each other. In the unattended condition, twelve blocks were presented, for a total of 7200 patterns, of which 720 were deviant patterns (180 for each condition). In the attended condition, ten blocks were presented, for a total of 6000 patterns, with 600 deviant patterns (150 for each condition). To ensure attention to the rhythms, participants were asked to detect target tones that were presented unexpectedly early. By using temporal perturbations as targets we aimed to draw attention to the temporal structure of the rhythm, while avoiding noise from manual responses to the deviants. In the regular sequences, the inter-onset interval before a target tone was shortened with 40 ms. In the jittered sequences, the inter-onset interval before a target tone was set to 110 ms. In both conditions, the inter-onset interval after a target tone was lengthened with 40 ms. Each sequence in the attended condition could contain up to 2 target tones. Target tones and five patterns following target tones have been excluded from further analysis.

Participants were tested individually in a soundproof, electrically shielded booth at the University of Amsterdam. After providing consent, participants first completed the unattended EEG experiment and subsequently the attended EEG experiment. In the unattended condition, participants were instructed to ignore the rhythms and focus on a self-selected, muted and subtitled movie. In the attended condition, they were asked to focus on the rhythm and press a response button whenever a tone was unexpectedly early. Before the start of the attended condition, participants were presented with a practice block to get familiarized with the task. Participants could take breaks between blocks as needed. Rhythms were presented through two custom-made speakers at 60 dB SPL using Presentation® software (Version 17.4, http://www.neurobs.com). After the EEG experiment, participants performed the BAT perception task and filled out the questionnaire from the Gold-MSI to assess their beat perception skills and general musical sophistication (Müllensiefen et al., 2014). The entire session lasted on average 3.5 hours.

6.2.4 EEG recording

EEG was recorded at a sampling rate of 8 kHz, using a 64 channel Biosemi Active-Two reference-free EEG system (Biosemi, Amsterdam, The Netherlands). Electrodes were positioned according to the 10/20 system and additional electrodes were placed

---

[5] Supplementary Sounds for this chapter are available online at http://www.fleurbouwer.nl/publications.

at left and right mastoids, on the nose, above and below the right eye, and to the left and right of the eyes.

6.2.5   EEG analysis

Matlab (Mathworks, Inc.) and EEGLAB (Delorme & Makeig, 2004) were used for data preprocessing. EEG data was offline re-referenced to linked mastoids, down-sampled to 512 Hz, and filtered using 0.5 Hz high-pass and 20 Hz low-pass linear finite impulse response filters. For 4 participants, one or two bad channels were removed and replaced by values interpolated from the surrounding channels. Independent component analysis was used to remove eye-blinks. Epochs of 650 ms, starting 150 ms before and aligned to the onset of the deviant sound were extracted for the four deviant patterns ($D1_r$, $D1_j$, $D2_r$, $D2_j$). In addition, epochs of the same length were extracted for bass drum sounds from the standards in the regular condition, both on the beat (from S1, but only if preceded by S2) and offbeat (from S2). The acoustic context preceding all tones used for analysis, deviants and standards, was identical (a bass drum sound 225 ms before the onset of the epoch). Epochs with an amplitude change of more than 150 µV in a sliding 500 ms window were rejected from further analysis. Epochs were baseline corrected using the average voltage of the 150 ms prior to the onset of the tone and averaged to obtain ERPs for each condition and participant. We obtained difference waves by subtracting the ERP responses to the bass drum sounds from the standard patterns from the ERP responses to the deviant tones at the same position (beat or offbeat). Finally, we averaged over participants to obtain grand average ERPs and difference waves.

Both in the attended and the unattended condition, a negative deflection peaking between 100 and 200 ms after the onset of the deviants was visible in the grand average difference waves (Figure 6.2 and Figure 6.3), consistent with the latency of an N2b and an MMN respectively. Scalp distributions ranged from fronto-central for regular deviants on the beat ($D1_r$) to more posterior for jittered deviants offbeat ($D2_j$). To assess possible differences in scalp distribution, we performed the analysis for the two early components on electrodes FCz, Cz, and CPz. We defined the amplitude of the MMN and N2b as the average amplitude from a 60 ms window centered around the average peak latency across conditions on Cz. The MMN peaked on average at 130 ms and the N2b peaked on average at 155 ms. Amplitudes were thus defined as the average amplitude of the difference waves in a 100–160 ms time window for MMN and a 125–185 ms time window for N2b.

Both in attended and unattended conditions, a positive deflection followed the negative component in response to the deviants. In all conditions, this response was maximal over FCz, consistent with the scalp distribution of a P3a elicited by the novelty of a stimulus (Polich, 2007). The deviants were not used as targets in the attended condition and therefore not task-relevant, which explains why a P3a was observed and not a P3b. While for regular deviants on the beat ($D1_r$) a clear peak could be observed for the P3a both in the attended (at 241 ms) and in the unattended condition (at 225 ms), for the other deviants the peak was less pronounced. This was caused by overlap with the P1 response elicited by the next sound, which was presented at 225 ms after the onset of each deviant. This overlap prevented us from reliably estimating the peak latency of

the P3a. To avoid contamination of the subsequent sound as much as possible in the analysis of the amplitudes, we defined the amplitudes for the P3a as the average amplitude from the difference waves in a 60 ms window encompassing mostly the earlier portion of the P3a. To avoid overlap with the MMN and N2b components, we chose windows for the P3a starting 20 ms after the end of the windows used for the previous components in both the unattended (180–240 ms) and attended (205–265 ms) conditions. As the P3a was maximal over FCz for all conditions, we only included this fronto-central electrode in the analysis.

6.2.6 Statistical analysis

For both attended and unattended conditions, the amplitudes extracted from the difference waves were entered into repeated measures ANOVAs with within subject factors position (on the beat or offbeat) and regularity (regular or jittered). For the MMN and N2b, electrode (FCz, Cz, or CPz) was used as an additional factor. To correlate beat perception as measured with ERPs with measures of musical ability, we quantified beat perception as the magnitude of the interaction between position and regularity. For each participant, this measure was obtained by subtracting the difference between the responses to $D1_j$ and $D2_j$, which reflected only sequential learning, from the difference between the responses to $D1_r$ and $D2_r$, which reflected both sequential learning and beat perception. For all ERP components of interest, partial correlations were used to examine the association between beat perception and scores on the musical ability tests. To account for the possible correlation between scores on the BAT and musical training scores (Müllensiefen et al., 2014), each musical ability measure was correlated with beat perception while controlling for the other measure. All statistical analyses were conducted in SPSS (Version 22). Greenhouse-Geisser corrections were used when the assumption of sphericity was violated.

6.3 Results

6.3.1 Musical abilities

On average, participants scored 27.8 (SD 14.4) on the musical training subscale, which is slightly higher than the average score of 26.52 as reported in Müllensiefen et al. (2014). Also, the average accuracy on the BAT perception test was 0.79 (SD 0.17), while the average reported by Müllensiefen et al. (2014) was 0.70. The slightly higher scores in our sample as compared to the norm data is not surprising, as we specifically also included professional musicians in our sample to obtain a large spread in musical abilities. Scores on the musical training subscale correlated with the accuracy on the BAT ($r = 0.50$, $p = 0.003$), similar to Müllensiefen et al. (2014).

6.3.2 ERPs

Table 6.1 shows average amplitudes for all ERP components of interest. ERPs, difference waves and average amplitudes on electrode FCz for all deviants are depicted in Figure 2 (attended) and Figure 3 (unattended). In the N2b window (attended), there was a significant three-way interaction between electrode, position and regularity ($F_{2,66} = 15.3$, $p < 0.0005$, $\eta^2 = 0.32$). Resolving this interaction by electrode showed

Table 6.1 Mean average amplitudes (μV) for all components on FCz. Standard deviations in brackets.

Condition	Attended		Unattended	
	N2b	P3a	MMN	P3a
Beat regular	–3.02 (3.02)	6.87 (5.15)	–2.12 (1.67)	3.88 (1.97)
Offbeat regular	–0.20 (1.56)	4.07 (2.32)	–1.02 (1.59)	2.46 (1.67)
Beat jittered	0.12 (1.84)	5.10 (1.78)	–0.83 (1.63)	3.24 (1.61)
Offbeat jittered	0.49 (2.03)	3.51 (1.89)	–0.57 (1.64)	2.68 (1.99)

that the interaction between position and regularity was significant on FCz ($F_{1,33} = 29.7$, $p < 0.0005$, $\eta^2 = 0.47$), Cz ($F_{1,33} = 17.4$, $p < 0.0005$, $\eta^2 = 0.35$), as well as CPz ($F_{1,33} = 13.2$, $p = 0.001$, $\eta^2 = 0.29$). The three-way interaction was due to the effect size for the two-way interaction being bigger on FCz than on Cz and bigger on Cz than on CPz. The interaction between position and regularity was also significant in the P3a window in the attended condition ($F_{1,33} = 4.3$, $p = 0.046$, $\eta^2 = 0.12$) and in the MMN and P3a windows in the unattended condition ($F_{1,33} = 11.5$, $p = 0.002$, $\eta^2 = 0.26$ and $F_{1,33} = 9.1$, $p = 0.005$, $\eta^2 = 0.22$). The three-way interaction between position, regularity and electrode did not reach significance for the MMN ($F_{2,66} = 0.764$, $p = 0.44$, $\eta^2 = 0.23$), showing that the interaction between position and regularity was equally large on all three electrodes. For all components, both in the attended and the unattended condition, the interaction was in the predicted direction (see Figures 6.2 and 6.3), with a significantly larger difference between the responses to deviants on the beat and offbeat in the regular (D1$_r$ and D2$_r$) than in the jittered condition (D1$_r$ and D2$_r$). This suggests that a beat was perceived, both with and without attention directed at the rhythm.

An analysis of the simple effects of position showed that the difference between the responses to deviants on the beat and offbeat was not only significant in the regular condition ($p < 0.0005$ for all ERP components), but also in the jittered condition. Responses to D1$_j$ were larger than to D2$_j$ in the attended condition in the N2b window on both Cz ($p = 0.020$) and CPz ($p = 0.045$) and in the P3a window ($p < 0.0005$). In the unattended condition, the responses to the jittered deviants differed significantly only in the P3a window ($p = 0.015$) but not in the MMN window ($p = 0.65$). These results suggest that participants could detect the statistical regularity in the order of the sounds in the jittered sequences, both when actively listening to the rhythms and when directing attention elsewhere. The simple effect of regularity was not only significant on the beat ($p < 0.019$ for all components) but also offbeat. ERP responses to D2$_r$ were larger than responses to D2$_j$ in the N2b window (attended) on FCz ($p = 0.012$) and Cz ($p = 0.020$) and in the MMN window (unattended) on FCz ($p = 0.044$). This suggests that the isochronicity of the regular sequence enhanced detection of the deviants, even in the offbeat position, in line with previous findings by Schwartze et al. (2011). Responses to D2$_r$ and D2$_j$ did not differ in the P3a windows (both in attended and unattended conditions $p > 0.17$).
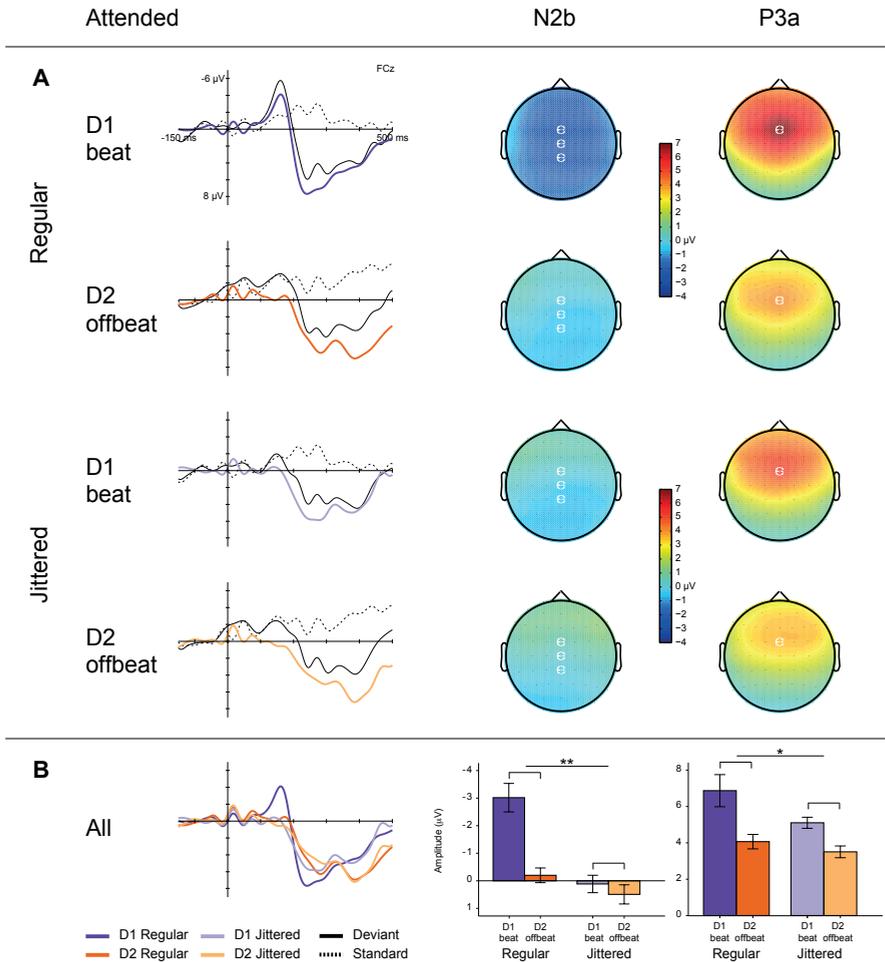
Figure 6.2 ERP responses in the attended condition. A) ERP responses on electrode FCz to standard and deviant tones in all four conditions, their derived difference waves and scalp distributions for the N2b and P3a components, averaged over the windows used for the analysis (125–185 and 205–265 ms respectively). The electrodes used for the analysis are indicated in white. B) Differences waves for all four conditions and the average amplitudes in the time windows used for analysis for both N2b and P3a. **Significant interaction at $p < 0.0005$. *Significant interaction at $p = 0.046$. Note that significance of simple effects is not displayed. Error bars represent one standard error of the mean.

Table 6.2 Partial correlations between beat perception as measured with ERPs on electrode FCz and musical training (while controlling for BAT scores) and BAT scores (while controlling for musical training). Note that correlations between a larger effect in ERPs and higher scores on the Gold-MSI tests are negative for N2b and MMN and positive for P3a, due to the polarity of the components. These correlations are indicated in bold. **Significant at $p < 0.02$.

Musical ability	Attended		Unattended	
	N2b	P3a	MMN	P3a
Musical training	**−0.42****	−0.16	0.00	−0.27
BAT scores	0.21	**0.41****	**−0.18**	**0.17**

6.3.3 Correlations between ERPs and musical abilities

The interaction between position and regularity was maximal over FCz for all components. Thus, partial correlations between beat perception and musical abilities were calculated for this electrode (see Table 6.2). In the attended condition, beat perception as observed in the P3a correlated significantly with the scores on the BAT when controlling for musical training ($r = 0.409$, $p = 0.018$). In the N2b window, beat perception correlated with the scores on the musical training questionnaire when controlling for the scores on the BAT ($r = –0.420$, $p = 0.015$). Interestingly, neither musical training, nor beat perception abilities correlated significantly with beat perception in the unattended condition (see Table 6.2).

6.4 Discussion

In the current research we examined beat perception with and without attention, while disentangling beat perception from sequential learning. The effect of metrical position on the ERP responses elicited by deviants was much larger in the regular condition, in which both beat perception and sequential learning were possible, than in the jittered condition, in which only sequential learning was possible. This effect was present both when participants attended to the rhythm and when their attention was directed away from the rhythm, suggesting that participants perceived a beat in both conditions. Previously, beat perception was found to be possible without attention directed at the rhythm with ecologically valid stimuli or real music (Bolger et al., 2013; Bouwer et al., 2014). While these previous results may have been confounded with effects of sequential learning, here we show that even when controlling for transitional probabilities and pattern learning, beat perception is possible when attention is directed away from the rhythm.
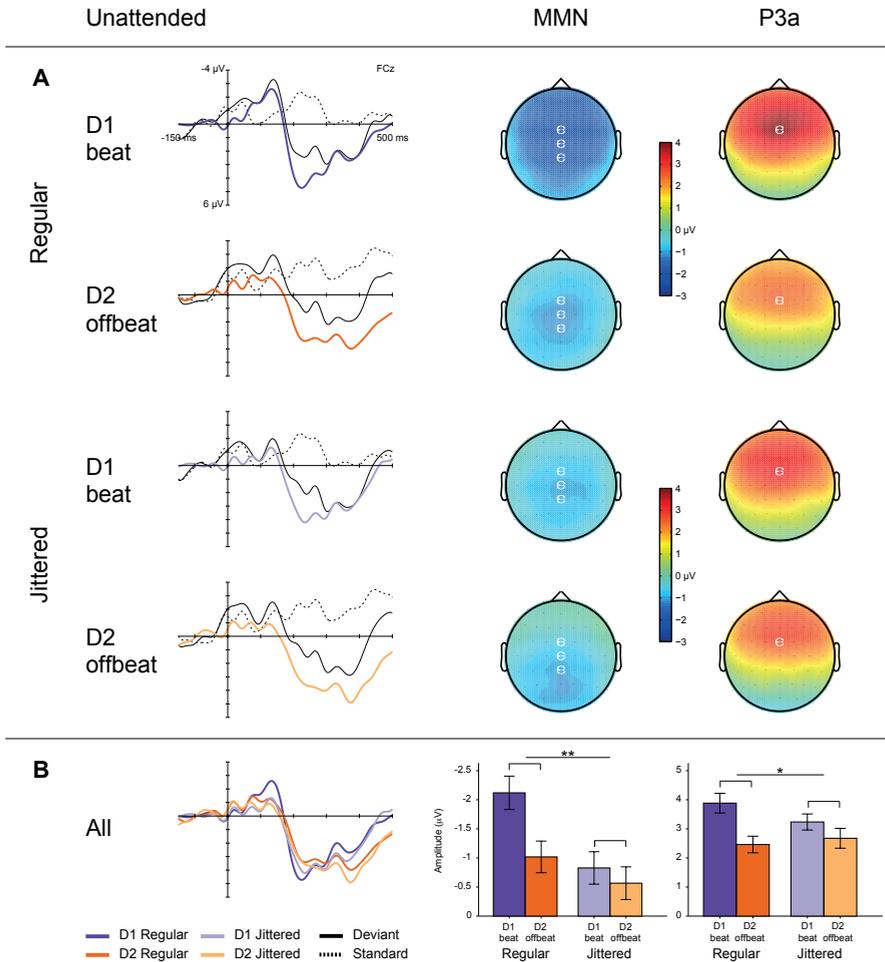
Figure 6.3 ERP responses in the unattended condition. A) ERP responses on electrode FCz to standard and deviant tones in all four conditions, their derived difference waves and scalp distributions for the MMN and P3a components, averaged over the windows used for the analysis (100–160 and 180–240 ms respectively). The electrodes used for the analysis are indicated in white. B) Differences waves for all four conditions and the average amplitudes in the time windows used for analysis for both MMN and P3a. **Significant interaction at $p < 0.0005$. *Significant interaction at $p = 0.002$. Note that significance of simple effects is not displayed. Error bars represent one standard error of the mean.

The effect of metrical position in the jittered condition, while much smaller than in the regular condition, was significant for both N2b (attended) and P3a (attended and unattended) responses, with larger responses on the beat (in odd positions) than offbeat (in even positions). This effect can be explained by assuming that participants learned that the probability of a soft sound was smaller in odd (on the beat) than in even (offbeat) positions, as hi-hat sounds were softer than bass drum sounds and only occurred in even positions. This statistical regularity in the order of sounds would have made intensity decrements less expected in odd positions (on the beat) than in even positions (offbeat), as evidenced by larger ERP responses to deviants in odd than even positions. This shows that sequential learning can affect responses to rhythm, even when the rhythm is irregular and when attention is not directed at the rhythm. We thus replicate previous findings showing that humans have remarkable abilities to extract statistical regularities from auditory sequences, and that sequential learning can occur implicitly, not only when there is no intention to learn, but also when there is no intention to listen (Daltrozzo & Conway, 2014; Van Zuijen et al., 2006). This finding stresses the importance of controlling for statistical regularity in the order of sounds when testing beat perception using ERPs. It is not inconceivable, especially when listening to natural music, that a large part of predicting rhythmic events may be the result of learning patterns. For the perception of pitch and melody in music, models of statistical learning have had considerable success in explaining human behavior (Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). Extending existing models of beat perception with a statistical component may be a promising avenue for future research to differentiate between various processes that contribute to rhythm perception, including beat perception and sequential learning.

The presence of statistical regularity in the order of sounds may aid beat perception by making accents more salient and more predictable. This may explain why real music is more effective in inducing a beat than abstract stimuli (Bolger et al, 2013). Similarly, it could be argued that temporal regularity may aid sequential learning. Indeed, it has been shown that sequential learning can benefit from regularity in grouping structure (Hoch, Tyler, & Tillmann, 2012) and metrical regularity (i.e., beat perception) in non-isochronous rhythms (Selchenkova, Jones, & Tillmann, 2014; Selchenkova, François, et al., 2014). In the current study, sequential learning of the order of sounds may have benefitted from the lower temporal variability in the regular sequences than in the jittered sequences, and this may partly have caused the interaction between the regularity of the sequence and the metrical position of the deviant. To date, it remains unknown whether differences in temporal variability, as in the current study, affect sequential learning similarly to grouping structure (Hoch et al., 2012) and metrical structure (Selchenkova, Jones, et al., 2014; Selchenkova, François, et al., 2014).

In addition to support for the presence of beat perception and sequential learning, we also found better deviant detection in the regular than in the jittered condition in offbeat positions, both in attended (N2b) and unattended (MMN) conditions. An advantage in the detection of deviants in regular as compared to in jittered sequences is in line with previous findings (Schwartze et al., 2011; Takegata & Morotomi, 1999). This advantage may be due to easier prediction of absolute time intervals in the regular than in the jittered sequences, as the former are less variable than the latter. Alternatively,

it may be due to the recruitment of beat-based timing mechanisms during the perception of the regular but not the jittered sequences. On the basis of the current experiment, we cannot rule out either explanation.

The perception of a beat with multiple hierarchical levels may be a somewhat different process from the perception of regularity at one level (Fitch, 2013; Tierney & Kraus, 2014), even when both rely on beat-based timing. Thus, listening to isochronous sequences, as often used in beat perception research (Cirelli et al., 2014; Fujioka et al., 2012), may not only rely partly on absolute interval perception instead of beat-based perception, it may also tap into different beat-based processes than beat perception in real music, as isochronous sequences only contain one level of hierarchy. The view that perception of isochronous sequences differs from beat perception is supported by findings showing that a small portion of the population is unable to synchronize to music, while they can synchronize to a metronome (Sowiński & Dalla Bella, 2013). Understanding the relationship between absolute interval perception, beat-based perception, hierarchical perception of a metrical structure, and sequential learning will be an interesting challenge for future research.

While the influence of regularity on the detection of deviants in offbeat positions was visible in the MMN and N2b responses, it was absent in the P3a responses. This may have been due to a suboptimal estimate of the P3a responses caused by overlap with the responses elicited by subsequent sounds. An interesting alternative interpretation may be that the responses to the offbeat deviants in the regular condition were in fact actively suppressed. If beat perception in the current experiment indeed relied on entrainment of multiple oscillators (Large & Jones, 1999; Large, 2008), not only may the responses on the beat have benefitted from peaks in attentional resources, the responses offbeat may have suffered from troughs in attentional resources. Such suppression may provide an interesting way for future research to separate beat perception, which predicts suppression of responses offbeat, from predictions through absolute interval perception and enhanced sequential learning in regular compared to irregular sequences, neither of which would lead to such suppression.

Beat perception as indexed by the N2b response in the attended condition correlated with the responses on the musical training subscale of the Gold-MSI. Beat perception as indexed by the P3a response in the attended condition correlated with beat perception abilities as measured by the BAT perception task. Thus, confirming previous research (Grahn & Schuit, 2012), both beat perception abilities and musical training explained unique variance in the responses to metrical rhythm. However, in the current study, neither correlated significantly with beat perception in the unattended condition, which is in line with previous research showing no difference in beat perception without attention between musicians and non-musicians (Bouwer et al., 2014). Beat perception in unattended conditions may rely on a mechanism like neural entrainment, which has been suggested to be inherent in the structure of the brain (Large, 2008) and to be independent of attention (Escoffier et al., 2015). While we cannot directly compare the different ERP components measured in this study, the effect size for the interaction between position and regularity was much larger for the N2b, in the attended condition, than for the MMN, in the unattended condition. This could be interpreted as

evidence showing that entrainment can be enhanced by attention. Alternatively, entrainment could be accompanied by additional mechanisms contributing to beat perception that do depend on attention and training. Previously, we have shown that beat perception consists of multiple mechanisms that together shape our perception of metrical rhythm (Bouwer & Honing, 2015). When not attending to the rhythm, participants may have relied only on entrainment, while when attending to the rhythm, they may have used additional mechanisms, which may be enhanced by musical abilities, to induce a beat.

6.5   Conclusion

In the current experiment, while controlling for sequential learning, we showed that beat perception is possible when attention is not directed at a rhythm. In addition, we showed that musical abilities, trained and untrained, are associated with beat perception, but only when attention is directed at the rhythm. Our results stress the importance of carefully defining beat perception, not only as a monolithic cognitive mechanism, but also in terms of the multiple underlying processes that together shape our perception of metrical rhythm. Which subcomponents of beat perception listeners recruit could well depend on the acoustical structure of the music, the resources a listener can devote to beat perception, and the musical abilities of the listener. Decomposing beat perception may be crucial in answering questions regarding the origins (Honing et al., 2015), mechanisms (Merchant et al., 2015), and possible applications (Nombela, Hughes, Owen, & Grahn, 2013) of this unique human ability.

Chapter 7

Discussion

Beat perception is the process that allows us to make and experience music together. While this ability is fundamental to musical behavior (Honing, 2012) and does not require formal training (Merchant et al., 2015), it is not clear what the necessary ingredients are for a listener to pick up on the beat in music. In this dissertation, I examined the influence of characteristics of the context, the listener, and the rhythm on beat perception, using both behavioral and electrophysiological methods. In addition, I addressed several issues pertaining to stimulus design, and I examined the cognitive processes underlying beat perception. In this discussion, I will first summarize and discuss the findings regarding the influence of attention, musical abilities, and properties of the rhythm on beat perception. Subsequently, I will discuss the broader implications of the work presented here, focusing on both the design of beat perception experiments and the mechanisms underlying beat perception.

7.1   Characteristics of the context: attention

Beat perception seems a trivial ability that even has been observed in small infants (Hannon & Johnson, 2005; Zentner & Eerola, 2010). This raises the question of whether beat perception is possible when attention is not directed at a rhythm. In addition to answering fundamental questions about the mechanisms underlying beat perception, the relationship between beat perception and attention also has implications for practical applications of rhythm in clinical settings. A regular beat in music can improve gait in patients with Parkinson's disease, possibly by facilitating the impaired basal ganglia (Nombela et al., 2013). If the perception of a beat depends on attention and requires substantial cognitive resources, this has implications for using music in rehabilitation in everyday life, as it can then not be used while performing other tasks.

Several studies have shown that the detection of regularity (Fujioka et al., 2012) and beat perception (Geiser et al., 2010; Ladinig et al., 2009, 2011) are indeed possible when attention is directed away from rhythm. Others, however, have shown that attention needs to be directed at the rhythm to detect regularity (Schwartze et al., 2011) or perceive a beat (Chapin et al., 2010; Geiser et al., 2009). Throughout this dissertation, we manipulated attention in various ways while measuring responses to rhythms containing a beat. In Chapter 4, we performed two separate experiments: in Experiment 1, attention was directed toward the rhythm, and we measured reaction times and hit rates in response to unexpected deviants in a speeded detection task. In Experiment 2, EEG was used to measure ERP responses to the same deviants without attention directed

toward the rhythm. In Chapter 5, we measured ERP responses while attention was directed toward a subtitled movie and the rhythm was ignored. In Chapter 6, we measured ERP responses both when participants attended to the rhythm and when attention was directed away.

The results from the two experiments in Chapter 4 cannot be compared to each other directly, because of the different methods. Qualitatively, in Experiment 1, in which the rhythm was attended, clear differences were found between responses to deviants in different metrical positions. In contrast, in Experiment 2, in which the rhythm was not attended, differences between conditions did not reach significance for the MMN responses and were very small for the P1 responses. We manipulated attention in Experiment 2 by having participants watch a silent movie. This is a manipulation that is often used when the MMN is measured as an index of pre-attentive processing. However, watching a movie does not require highly focused attention (Haroush et al., 2010) and allows for momentary shifts in attention between the movie and the auditory stimuli. Considering the weak results in Experiment 2 in Chapter 4, even with a rather lenient manipulation of attention and having only highly trained musicians participate, it seems that attention may be necessary to perceive a beat.

Contrary to the findings of Chapter 4, in Chapter 5 and Chapter 6, we did find strong effects of metrical position on MMN responses to unexpected rhythmic events when attention was directed away from the rhythm, even in musical novices. In the experiments presented in Chapter 4, we used very sparse stimuli, which did not contain any accents indicating the beat. In the experiments presented in Chapter 5 and Chapter 6, we used stimuli in which accents were present simultaneously in intensity, timbre and loudness. Such acoustic accents provide exogenous cues to the listener about the position of the beat. Without accents, beat perception may be a more endogenously driven process (Chapin et al., 2010; Grahn & Rowe, 2009). Thus, clear accents may be necessary for beat perception when attention is not directed at a rhythm. In the presence of these exogenous cues, as in Chapter 5 and Chapter 6, beat perception seems possible with attention directed away from rhythm.

The observation that beat perception is possible with attention directed away from rhythm does not necessarily mean that attention does not modulate beat perception. In Chapter 5, we only measured responses without attention directed to the rhythm, and we thus cannot draw any conclusions about a possible modulatory influence of attention on beat perception. In Chapter 6, we measured ERP responses both with and without attention, which potentially would allow for a comparison. Several issues however make a direct comparison of the results in the different attention conditions difficult to interpret. First, in the attended condition, an N2b was elicited, while in the unattended condition, an MMN was observed. These components differ in latency and scalp distribution, and while the MMN originates from auditory cortex (Näätänen et al., 2007), the N2b has been associated with sources both in auditory cortex and anterior cingulate cortex (Folstein & Van Petten, 2008; S. H. Patel & Azzam, 2005). These components thus cannot be directly compared. Instead, we could potentially compare the effect sizes for the experimental factors in the different attention conditions. In the experiment reported in Chapter 6, we indexed beat perception as the interaction between the

position of the deviant and the regularity of the sequence. For the early negative components, the effect size of this interaction was larger in the attended (N2b) than in the unattended (MMN) condition, which may indicate that attention indeed modulates the effects of beat perception. However, such a comparison of effect sizes may not be a fair comparison either. The tasks used in the two attention conditions were very different, with participants watching a silent movie in the unattended condition and performing a rhythm-related task in the attended condition. The differences between the tasks went beyond just a difference in the direction of attention and included differences in visual input (a movie or a fixation cross) and differences in working memory load (a movie or a detection task in the auditory domain). While the MMN is known to occur without attention (Näätänen et al., 2007), a concurrent task can influence MMN amplitude. When attentional load in the auditory domain is high, the detection of regularity in an auditory sequence is more difficult (Chait, Ruff, Griffiths, & McAlpine, 2012), and the formation of a regularity representation, which is crucial for the MMN to be elicited (Sussman, 2007), may be affected by this. In addition, even a visual task can affect MMN amplitude (Haroush et al., 2010). To examine the presence of beat perception, we always compared responses to different rhythmic events within an attention condition, and these should be equally affected by the attention manipulation. However, the differences between the attention conditions that were not related to the direction of attention may have affected the ERP responses. These differences therefore prevent a fair comparison of the strength of beat perception in attended and unattended conditions.

A final issue in the comparison of the attention conditions in Chapter 6 lies in the method we used to index beat perception. Sequential learning could possibly have a confounding effect on the difference between ERP responses to deviants on the beat and off the beat. To control for this, we measured responses to deviants in both regular sequences, in which a beat could be perceived, and jittered sequences, in which the order of sounds in the sequence was preserved, but beat perception was impossible. The presence of beat perception was deduced from a larger difference between responses to deviants on the beat and off the beat in the regular than in the jittered condition. Beat perception was thus indexed as the size of the interaction between the position of the deviant and the regularity of the sequence. The way we indexed beat perception may further bias a comparison of the attention conditions. While sequential learning can occur without attention directed at a sequence (Van Zuijen et al., 2006), the dependence of sequential learning on attention is still debated (Daltrozzo & Conway, 2014). In Chapter 6, more sequential learning of the order of sounds in the rhythmic sequences may have occurred in the attended than in the unattended condition. In the experiment presented in Chapter 6, sequential learning caused differences between responses on the beat and off the beat, both in the regular and the jittered condition, and these differences may thus have been generally larger in the attended than in the unattended condition. A larger difference between responses to deviants on the beat and off the beat in the jittered sequences in the attended condition may have partly obscured the interaction between the regularity of the sequence and the position of the deviant, which is how we indexed beat perception. Enhanced sequential learning in the attended condition may therefore explain why the effect size for beat perception

as measured with the P3a, which showed especially strong effects of sequential learning, was smaller in the attended than in the unattended condition.

To reach conclusions about the possible modulation of beat perception by attention, a better approach would be to measure beat perception while manipulating the attentional load of a second concurrent task, which apart from the difficulty of the task would have to be identical between different attention conditions. An example would be an N-back task with varying levels of difficulty (cf. Pecenka, Engel, & Keller, 2013). Such a task does not explicitly change the direction of attention, but it does change the amount of attentional resources available for processing the rhythm. The direction of attention and the amount of attentional resources can be viewed as two sides of the same coin, with both reflecting limited processing resources. While attentional resources may be related to working memory and may reflect the focusing of attention on maintaining internal representations, selective attention may be seen as focusing of resources on external stimuli (Kiyonaga & Egner, 2013). In examining the influence of attention on beat perception, a manipulation of available attentional resources and a manipulation of the direction of attention may thus essentially be testing the same thing.

In summary, while we can conclude that beat perception is possible when attention is directed away from a rhythm, based on the studies in this dissertation, we cannot draw definite conclusions about a possible modulatory effect of attention on beat perception: this will have to be examined in future studies.

7.2   Characteristics of the listener: musical abilities

Formal training is not a necessity to be able to perceive a beat in music (Merchant et al., 2015), but musical training can enhance beat perception (Geiser et al., 2010; Grahn & Rowe, 2009; Vuust et al., 2005). In Chapter 3 we showed that musical training enhances the sensitivity of the listener to the accent structure in a rhythm. Better sensitivity to the accent structure makes it easier for musically trained listeners to differentiate between rhythms that contain a clear beat and rhythms that do not. Interestingly, it did not matter whether accents were only present in the temporal structure of the rhythm or were clearly marked in intensity differences. Musical training enhanced sensitivity to the accent structure for both types of accents.

Contrary to the results from Chapter 3, in Chapter 5 we did not find any differences between musicians and non-musicians in ERP responses to unexpected omissions on the beat and off the beat. This discrepancy between the results from Chapter 3 and Chapter 5 can be explained by various differences between the experiments in the stimuli and methods. First, in Chapter 5 we used stimuli presented in a continuous stream. The rhythm was mostly strictly metric (i.e., did not contain any syncopation), with accents always present on the beat and never off the beat. Sounds on the beat were not only louder than sounds off the beat, but also longer and of a different timbre. Thus, multiple simultaneous cues indicated the beat. The length of the rhythms and the richness of the sounds may have made beat perception especially easy, leading to ceiling effects that obscured possible differences between musicians and non-musicians in Chapter 5. In Chapter 3 we used short rhythms that only contained accents either in

the temporal grouping of the tones, or in intensity, but never in combination. This may have made the perception of a beat sufficiently hard to prevent a ceiling effect as may have been present in Chapter 5, making it possible in Chapter 3 to detect individual differences in beat perception caused by musical training.

Furthermore, in Chapter 5 we used EEG to measure the responses to unexpected omissions, while in Chapter 3, behavioral methods were used to probe beat perception. In the latter chapter, participants were asked to provide explicit ratings of beat presence, which may have required explicit knowledge that depends on formal musical training. Such explicit knowledge is not necessarily related to implicit processing of musical structure (Bigand & Poulin-Charronnat, 2006), which may depend less on formal training. In line with this, in Chapter 3 participants were attending to the rhythms, while in Chapter 5, they were not. Possibly, processes that are both attention-dependent, and that relate to explicit reporting of the perceived beat, can be trained. This view is corroborated by the results of the experiment presented in Chapter 6, in which we examined ERP responses to unexpected soft sounds on the beat and off the beat in regular and irregular rhythms. In the attended condition, musical training correlated with the effect of beat presence on the amplitude of the N2b. However, when people's attention was directed away from the rhythm, we found no correlation between musical training and the effect of beat presence on the amplitude of the MMN.

Two additional issues concerning the influence of musical training on beat perception must be addressed. First, it is important to note that musical training can be defined in multiple ways. Often, as we have done in Chapter 5, a comparison is made between two groups, with participants having had either no musical training at all (non-musicians), or many years of formal training (musicians). Many different operationalizations of "musician" have been used, ranging from anyone with more than 5 years of musical training (Grahn & Brett, 2007; Grahn & Rowe, 2009), to only trained percussionists (Cameron & Grahn, 2014; Geiser et al., 2010), to only students or graduates majoring in music (Repp, 2010; Vuust et al., 2005), to anyone with 2 to 12 years of musical training (Drake, Penel, et al., 2000). Obviously, the differences in the definition of what constitutes a musician make it very hard to compare different studies directly. In Chapter 5, we included only professional musicians and music students as musicians, and non-musicians were participants with less than 2 years of musical training who were not actively playing an instrument at the time of the experiment. In Chapter 3 and Chapter 6, we used an alternative approach and included musical training as a covariate in the analyses. This may be a better solution when examining the effects of musical training, as it does not require an arbitrary cut-off point. However, some problems may also be associated with using musical training as a covariate. Conceptually, it is hard to imagine that musical training contributes to beat perception in a continuous way, with 10 years of musical training enhancing beat perception more than 9 years of musical training. It is much more likely that some changes due to musical training are qualitative and level off with more years of expertise. While the results of Chapter 3 and Chapter 6 show that musical training at least has some effect on beat perception, the exact nature of this effect remains to be specified.

In addition to the question of how to deal with musical training, it must also be noted that musical abilities are not limited to formal training. Large individual difference in beat perception abilities exist regardless of formal training (Grahn & McAuley, 2009; Grahn & Schuit, 2012; Müllensiefen et al., 2014). In addition, mere exposure to music can influence the ability of listeners to recognize timing in rhythm (Honing & Ladinig, 2009) and can provide listeners with implicit knowledge about musical structure (Bigand & Poulin-Charronnat, 2006). In Chapter 6, we used the recently published Gold-MSI (Müllensiefen et al., 2014) to allow for a broader definition of musical abilities than musical training alone. The Gold-MSI consists of a questionnaire and listening tests designed to measure musical abilities in the general population. We used two subtests of the Gold-MSI to measure both beat perception abilities and musical training. The Gold-MSI provides a more nuanced measure of musical training than the number of years of formal training alone, which also includes the time spent practicing and the years spent on music theory lessons. Using the Gold-MSI, we were able to show that innate beat perception abilities contribute to individual differences in responses to rhythm independently from musical training (Chapter 6). A better estimate of musical abilities of participants, both innate and passively or actively learned, is important in examining individual differences in beat perception. In addition to the Gold-MSI, several tests of musical abilities recently became available (Dauvergne, Begel, Benoit, Kotz, & Dalla Bella, 2015; Peretz, Champod, & Hyde, 2003). While care must be taken to choose valid tests (Tranchant & Vuvan, 2015), the use of these tests will be valuable in further examining the relationship between musical abilities and beat perception.

7.3   Characteristics of the rhythm: accents

As I described above, not all listeners are equally apt at picking up a beat in music, and the context of a listening experience can possibly influence whether we perceive a beat by affected how much resources are available for beat perception. Of course, a final aspect that can influence whether or not we perceive a beat in music is the rhythm itself. When accents occur in a rhythm with a regular temporal spacing between them, we can infer a beat from the structure of the accents (Lerdahl & Jackendoff, 1983a). Exogenous cues do not necessarily have to be present in a rhythm for a listener to perceive a beat. We can also perceive a beat in rhythm when no accents are present at all, as when we hear an isochronous sequence as alternating strong and weak tones (Abecasis et al., 2005; Brochard et al., 2003; Potter et al., 2009). Finally, a beat can also be imposed on rhythm through processes under cognitive control (Iversen et al., 2009; Nozaradan et al., 2011; Schaefer et al., 2010).

In Chapter 3 we examined whether the type of accent that is used influences how easy it is to map the structure of accents to a perceived beat. We compared responses to rhythms varying in metrical complexity, in which the beat was either indicated by intensity accents or by temporal accents. Previously, Grahn and Rowe (2009) showed that the perception of a beat in temporal rhythms activated the supplementary motor area (SMA) and putamen more than perception of a beat in intensity rhythms, indicating that beat perception through temporal accents may require more internal effort than beat perception through intensity accents. Consistent with this, connectivity between

auditory and motor areas was larger for temporal rhythms containing a beat than for intensity rhythms containing a beat, though only in musicians (Grahn & Rowe, 2009). Based on these results, we anticipated that intensity accents would be easier to extract from a rhythm than temporal accents and would thus have a larger effect on the perceived metrical structure, especially for musical novices. We found that musical training enhanced sensitivity to the structure of accents in general, regardless of the type of accent. Contrary to what we expected, both musical novices and musical experts differentiated more between rhythms with various degrees of syncopation when temporal accents were used than when intensity accents were used.

As described earlier in this discussion, attention may be necessary for beat perception when no accents are present and the listener necessarily needs to rely on endogenous processes, as in the experiment reported in Chapter 4. When clear accents in intensity, timbre and duration are present, as in the experiments reported in Chapter 5 and Chapter 6, beat perception most likely is driven by these exogenous cues, and this may make it possible for a beat to be perceived with less attentional resources available. Previously, it was shown that attention may also be necessary for beat perception when only temporal accents are present, and when rhythms are highly complex (Chapin et al., 2010). Together, these results suggest that beat perception is easier and requires fewer resources when acoustic rather than temporal accents are used to indicate the beat. This contradicts the findings of Chapter 3, where we found higher sensitivity to the structure of temporal than intensity accents. However, in the intensity rhythms, all subdivisions were marked with a sound. The presence of regular marked subdivisions may have made the mismatch between the accent structure and the beat in the more syncopated rhythms less disturbing in the intensity than in the temporal rhythms, leading to an underestimation of the effects of counterevidence for the intensity rhythms.

Generally, knowledge is lacking about the relationship between different types of accents and beat perception. Most research looking at the neural mechanisms underlying beat perception has only used temporal accents (Bengtsson et al., 2009; Chen et al., 2008a; Grahn & Brett, 2007; Grahn & Rowe, 2013). The involvement of the basal ganglia and the SMA in beat perception has been repeatedly shown, but these areas are less responsive during beat perception with intensity accents (Grahn & Rowe, 2009). Thus, the processes underlying beat perception with different types of accents may be qualitatively different. Consistent with these previous observations, in Chapter 3 we show that people differentiate between strictly metrical rhythms (i.e., rhythms with no syncopation at all) and rhythms with some degree of syncopation equally well when intensity accents and temporal accents are used. However, they differentiate more between rhythms with varying degrees of syncopation when temporal accents are used than when intensity accents are used, again showing qualitative differences between the perception of a beat with different types of accents. Although it is generally agreed that a beat can be inferred from intensity accents (Jackendoff & Lerdahl, 2006; Large, 2000; Lerdahl & Jackendoff, 1983a), the underlying (neural) mechanisms are not very well understood. Previous research may therefore have neglected beat perception in its most ecologically valid form, with intensity differences to indicate the beat, as is very common in real music.

7.4 Design issues

In Chapters 4, 5 and 6 we took a specific approach to testing beat perception with EEG. Instead of measuring beat perception directly, we used ERPs to probe the influence of beat perception on processing of auditory events in rhythm. In Chapter 4, we showed that the P1 response was larger for events on the beat than for events off the beat. This can be interpreted as evidence for heightened sensitivity on the beat, possibly due to peaks in attentional energy (Large & Jones, 1999; Large, 2008). In Chapter 5 and Chapter 6, we showed that the MMN, N2b and P3a responses to unexpected omissions and intensity decrements were larger on the beat than off the beat. This can be interpreted as evidence for heightened sensitivity on the beat, leading to larger responses on the beat than off the beat. It could also be viewed as evidence for strong predictions that an event will occur on the beat. An omission or intensity decrement on the beat is a larger violation of these predictions than an omission or decrement off the beat and as such elicits a larger response.

The use of our approach to probing beat perception has two clear advantages. First, we looked at well-studied ERP components with a long history (see for example Näätänen et al., 2007; Näätänen & Picton, 1987; Polich, 2007; Winkler, 2007). Second, we measured responses in real time, enabling us to look at the effects of beat perception in the temporal domain. Research looking at oscillatory responses to a regular beat similarly has the advantage of tracking responses to rhythm in real time. It could be argued that measurements of oscillations probe the processes underlying beat perception directly, instead of only looking at the effects of beat perception on auditory processing of rhythmic events. However, results concerning oscillatory responses to a regular beat have been mixed, with effects in the gamma range (Fujioka et al., 2009; Zanto et al., 2005), the beta range (Cirelli et al., 2014; Fujioka et al., 2015; Fujioka et al., 2012; Iversen et al., 2009; Te Woerd et al., 2014) and the delta range (Nozaradan et al., 2011, 2012; Nozaradan, 2014). In most of these studies, isochronous stimuli were used. In two cases stimuli consisted of a rhythm that contained some accent structure, but only in the temporal domain (Iversen et al., 2009; Nozaradan et al., 2012). Moreover, participants were presented with many repetitions of the same rhythmic pattern (Iversen et al., 2009; Nozaradan et al., 2012) and were asked to consciously impose a beat on the stimuli (Iversen et al., 2009). Thus, results from EEG studies measuring oscillatory responses to a regular beat do not provide a clear picture of the underlying processes of beat perception and are based on stimuli and tasks that arguably, are not ecologically relevant. In Chapter 5 and Chapter 6, we show how ERPs can be used to reliably elicit responses to rhythmic events in more complex rhythms that contain multiple hierarchical levels and various types of accents, and as such, more resemble real music.

The approach taken in this dissertation to measuring beat perception with ERPs also has disadvantages. Beat perception relies on a network of motor areas in the brain, and specifically the basal ganglia (Grahn, 2009). Because of the poor spatial resolution of EEG and the lack of visibility of deeper sources in the brain, we cannot look directly at activity in the basal ganglia. The measurement of ERPs in response to auditory events in rhythm thus cannot show activity that directly relates to all of the processes underlying beat perception. Additionally, attentional processes, which, according to

DAT, are involved in beat perception, may rely on synchronization of oscillatory activity across brain regions (Womelsdorf & Fries, 2007). With ERPs, oscillatory activity is not examined. The use of fMRI and measurements of oscillatory activity with EEG may thus be better suited to draw conclusions about the mechanisms underlying beat perception.

When measuring beat perception with ERPs, two additional design issues are important to consider. First, ERPs are extremely susceptible to sound differences, both in the event of interest and in its acoustic context (Woldorff & Hillyard, 1991; Woodman, 2010). When designing stimuli that can induce a beat, one strategy is to introduce accents on the beat. However, this often means that the sounds on the beat will necessarily be different from the sounds off the beat and that ERP responses to events on and off the beat cannot be directly compared. To circumvent this problem, accented events could be introduced sporadically off the beat, which usually does not intervene with beat perception. In Chapter 5, the introduction of a bass drum sound off the beat in position 6 of the pattern allowed us to compare the omission of identical sounds. However, the omissions on the beat always followed a hi-hat sound and the omissions off the beat always followed a bass drum sound. We used difference waves to subtract activity related to the acoustic context from the ERPs in response to the deviants, but this may not have completely eliminated the effects of the differences in context. ERPs are not only influenced by sounds themselves, but also by the offset of a sound (Hari et al., 1987). Thus, the response to an omission following a bass drum sound may have differed from the response to an omission following a hi-hat sound due to differences in offset response, rather than differences in metrical position. In Chapter 6, we show how to control for the acoustic context of the deviants by frequently introducing accented events off the beat that were used to compare responses to the different positions with not only identical sounds, but also identical acoustic contexts.

A second design issue that must be considered is that processes other than beat perception may cause differences between the responses to events in different metrical positions. In Chapter 4, we argue that beat perception affects processing of auditory events through the generation of predictions, and we measure the responses to violations of these predictions. Importantly, various other processes may be used by the auditory system to generate predictions about upcoming events. In Chapter 5, we looked at the MMN responses to omissions not only in the deviant patterns, but also in the standard patterns. We showed that the MMN in response to an omission on the second sixteenth of the standard pattern (S2, see Chapter 5) was larger than the response to an omission on the fourth or eighth sixteenth of the pattern (S3 and S4). These positions are all metrically equally weak. The differences can be explained by looking at the statistical properties of the patterns used. The omissions in patterns S3 and S4, which comprised 45 percent of the total patterns, always followed a snare drum sound. The omissions in pattern S2, comprising 22.5 percent of the total patterns, followed a bass drum sound, similar to omissions in pattern D3 (3.3 percent of the total patterns). Thus, the probability of an omission was higher after snare drum sounds than after bass drum sounds, and an omission was therefore more predictable following the former than the latter. These differences in transitional probabilities may have led to a smaller MMN response to omissions in S3 and S4 than to omissions in S2.

In addition, in Chapter 5 we showed that the MMN responses to omissions in standard pattern S2 were smaller than the MMN responses to omissions in deviant pattern D3, even though in both cases, the omission occurred on the second sixteenth of a beat and followed a bass drum sound. These differences cannot be explained by differences in transitional probabilities or acoustic context. However, in pattern D3, a bass drum sound was omitted, while in pattern S2, a hi-hat sound was omitted. A listener would only be capable of detecting this difference if the complete eight-tone pattern was represented in memory. The fact that we found a difference in responses to the omissions in patterns S2 and D3 thus shows that pattern learning can influence the predictions made by the auditory system, even in patterns as long as eight tones,. This has implications for the interpretation of two previous studies that used similar stimuli (Ladinig et al., 2009, 2011; Winkler et al., 2009). The results from these studies were interpreted as evidence showing beat perception in newborns (Winkler et al., 2009) and untrained adults (Ladinig et al., 2009, 2011), but can alternatively be explained by assuming that participants learned the eight-tone pattern and generated predictions based on their learning. The findings from Chapter 5 thus show the importance of controlling for learning of transitional probabilities and repeating patterns when comparing ERP responses to rhythmic events in different metrical positions.

To avoid differences in predictions based on acoustic variation within a rhythmic sequence, a possible strategy is to abolish sound differences altogether. Many studies have used isochronous sequences to study beat perception, completely ruling out any influence of pitch, timbre, intensity, or rhythmic variability (Fujioka et al., 2012; Hattori, Tomonaga, & Matsuzawa, 2013; Large & Gray, 2015; Nozaradan et al., 2011; Schwartze et al., 2011, 2013; Zarco et al., 2009). For example, the prediction of onsets in an isochronous sequence has been compared to the prediction of onsets in an irregular sequence, and the difference in processing has been interpreted as a hallmark of beat-based perception, or relative timing (Fujioka et al., 2012; Schwartze et al., 2013). However, prediction of an onset in a rhythmic sequence can be achieved by using absolute timing. Humans can reproduce rhythms that do not contain a beat (Cameron & Grahn, 2014), and this can only be achieved by predicting absolute temporal intervals. Moreover, the accuracy of the reproduction of a nonmetric rhythm is negatively associated with the variability of the rhythmic intervals (Cameron & Grahn, 2014). Thus, when comparing responses to regular and irregular sequences, it cannot be ruled out that both rely on interval-based perception. In regular sequences, a listener could learn the absolute length of the repeating interval and predict that subsequent intervals are equally long. In irregular sequences, learning of absolute intervals would be more difficult or even impossible, as the intervals are more variable than in regular sequences, or even random. Differences in responses to regular and irregular sequences could therefore be caused by differences in temporal variability that affect interval timing, instead of differences in beat-based processing. Thus, while isochronous sequences may eliminate the confounding effects of acoustic differences between metrical positions and sequential learning, they may not be optimal in testing beat perception.

In summary, ERPs provide a well-studied and reliable tool to study beat perception, but careful design is needed to ensure that differences between ERP responses to

events in different metrical positions are not caused by acoustic and statistical properties of the stimuli. In Chapter 6, we introduce a possible way of controlling for pattern learning, while still using sound differences to induce a beat with multiple metrical levels. We compared the responses to deviants in different metrical positions in temporally regular and jittered sequences. In the regular sequences, both beat perception and sequential learning were possible. In the jittered sequences, beat perception was impossible, while the statistical properties of the order of sounds in the sequences were preserved. We found that the difference between the responses to deviants on the beat and off the beat was larger for the regular than the jittered sequences, showing that beat perception affected the responses in addition to sequential learning. The fact that the position of the deviant influenced the magnitude of the ERP responses even in the jittered condition and even when attention was directed away from the rhythms shows that people are remarkably apt at learning the order of sounds. These findings further stress the importance of controlling for the predictions that are made based on sequential learning.

7.5   A look ahead: Decomposing beat perception

Beat perception is usually studied as a single process, underpinned by a dedicated neural system. In Chapter 4, we aimed to differentiate between temporal fluctuations in attention, as described by Dynamic Attending Theory (Large & Jones, 1999) and temporal predictions, as described by Predictive Coding (Vuust & Witek, 2014). To differentiate between these theories, we examined the responses to two types of unexpected events in the rhythm, intensity increments and decrements. In Experiment 1, we found that increments were detected better off the beat than on the beat when they were the size of a subjective accent. This is consistent with predictive coding, which predicts enhanced detection of deviants that violate predictions, such as an increment off the beat. However, increments were detected better on the beat than off the beat when they were larger, in line with DAT, which predicts better processing of any event that coincides with heightened attention on the beat. These results suggest that both fluctuations in attention and temporal predictions play a role in the perception of metrical rhythm and thus, that beat perception consists of several distinct underlying mechanisms.

To understand how humans perceive a beat, it is crucial to understand the different components that are part of or contribute to beat perception. Below, I will discuss four processes that may contribute to beat perception: temporal entrainment, temporal prediction, accent extraction, and sequential learning. One of the challenges for future research will be to tease apart the subcomponents of beat perception, and to arrive at a more complete neurocognitive model of this unique human ability.

The dynamic fluctuations in attention in time when listening to a regular stimulus are thought to result from the entrainment of delta oscillations to an external regularity (Henry & Herrmann, 2014; Schroeder & Lakatos, 2009). Delta oscillations act as an attentional selection mechanism by entraining their phase and consequently neuronal excitability to the regular stimulus. At moments of high excitability, cortical responses to sensory input are amplified (Schroeder & Lakatos, 2009), which is consistent with

the effects of attention on early auditory processing (Lange, 2013; Woldorff et al., 1993).

In previous research, attention has often been confused with prediction, even though they are distinct mechanisms (Schröger, Marzecová, & SanMiguel, 2015). Prediction of time intervals, and specifically the beat-based prediction of multiple identical time intervals, may rely on the basal ganglia (Grahn & Rowe, 2013; Te Woerd et al., 2014). Entrainment of delta oscillations to rhythm involves bilateral auditory cortices (Herrmann, Henry, Haegens, & Obleser, 2016; Nozaradan, Zerouali, Peretz, & Mouraux, 2015). Temporal prediction and entrainment of attention may thus be dissociable in the brain. However, in beat perception, they are likely to occur together in time. The relationship between them may be explained by assuming that entrainment of attention, originating from the auditory cortices, leads to predictions generated by the basal ganglia, which in turn influence processing in auditory cortices. This is consistent with observations of greater connectivity from basal ganglia to auditory cortex when a beat is perceived (Grahn & Rowe, 2009), and it also nicely fits with models that describe beat perception as an interplay between top-down and bottom-up processes (Vuust & Witek, 2014). The entrainment of attention would in this case be driven by bottom-up, exogenous cues, while predictions would have a top-down effect on perception.

Two recent theories of beat perception, the Action Simulation for Auditory Prediction (ASAP) hypothesis (Patel & Iversen, 2014) and the Gradual Audiomotor Evolution hypothesis (Merchant & Honing, 2014) similarly stress the importance of connections between the auditory and motor systems for beat perception. Here I expand on these theories by suggesting the possible mechanisms that may be subserved by these brain systems. Patel and Iversen (2014) state that entrainment is not sufficient for beat perception. Entrainment in auditory cortices, while not sufficient for beat perception, may however be crucial for the basal ganglia to align its predictions to a regular beat. The hypothesized causal relationship between entrainment in auditory cortices and predictions generated by the basal ganglia as I propose here is a hypothesis that can be tested in future research.

Considering entrainment and prediction to be separate processes may explain how beat perception affects ERP responses to auditory events and how musical training and attention mediate beat perception. In Chapter 6, we show that musical abilities only affect ERP responses to rhythm when attention is directed towards the rhythm, suggesting that some subprocess of beat perception depends on both musical abilities and attention. This dependence can be explained by assuming that entrainment may be inherent in the properties of the brain and thus largely independent of both attention and musical abilities, while the generation of precise predictions may be more dependent on learned abilities and top-down processing.

In Chapter 4, we focus on the role of entrainment and prediction in beat perception. These are of course not the only processes that may contribute to the coherent percept of the beat we experience when listening to music. A process that is often ignored in examining beat perception is the extraction of accents from an auditory sequence. In

Chapter 3, we focus on the relationship between the structure of accents in a rhythm and the perceived beat. We show that there are qualitative differences between the inference of a beat from temporal accents and intensity accents. Determining which sounds are most salient in a sequence, and thus which sounds are accented, can be regarded as a general aspect of auditory scene analysis. Auditory scene analysis is not necessarily specific to music perception (Trainor, 2015), but in order to hear a beat, the listener must first successfully process low-level auditory features, like the loudness and timing of sounds. Musical training enhances many aspects of auditory processing (Moreno & Bidelman, 2014). In future research, the possibility that enhanced beat perception in musicians is in fact due to enhanced general processing of sound must not be excluded.

Finally, in Chapter 6, we explicitly aimed to differentiate between the perception of a regular beat and sequential learning. Like beat perception, sequential learning in a rhythm can lead to predictions about the identity of upcoming events. While beat perception and sequential learning can be seen as separate processes, they may interact, with beat perception aiding sequential learning (Hoch et al., 2012; Selchenkova, François, et al., 2014). In addition, regularity in the non-temporal properties of rhythmic sounds (i.e., the statistical regularity in the order of sounds that allows for sequential learning) may aid the perception of a regular beat. Such a facilitating effect of statistical regularity on beat perception may explain why it is easier to hear a beat in real music than in sparse, simplified stimuli (Bolger et al., 2013; Tierney & Kraus, 2014).

If we regard beat perception from a functional point of view, its main function is the generation of precise predictions that allow us to synchronize to music (Honing et al., 2015). As I have argued throughout this dissertation, predictions in the auditory domain can also be generated through sequential learning and interval-based perception. Interval-based perception relies on different brain networks than beat-based perception, with an important role for the cerebellum (Grube et al., 2010; Teki et al., 2011). While beat perception thus seems to be dissociable from interval-based perception, behavioral and electrophysiological indexes of beat perception may also be affected by sequential learning and interval-based perception. To what extent these processes are active may depend on the context (e.g., is a rhythm attended to?), the rhythm itself (e.g., what kind of accents are present to indicate the beat?), and the listener (e.g., is the listener a "good" beat perceiver?). The interplay between different processes that lead to auditory predictions provides a host of topics for future research. An open question is whether perceptual effects similar to those that result from beat-based predictions could be accomplished with sequential learning and interval-based timing alone. The reliance on mechanisms other than beat-based timing in predicting rhythmic events could aid the use of rhythm in the rehabilitation of movement, for example in Parkinson's disease. Interval-based timing combined with sequential learning could either aid the beat-based predictions that are generated by the basal ganglia, or could allow for the generation of predictions while bypassing the impaired basal ganglia altogether (Nombela et al., 2013).

7.6   Conclusion

In this dissertation, I examined how beat perception is influenced by attention, musical abilities, and the type of accent that indicates the beat. In addition, I looked at several processes that contribute to beat perception: temporal entrainment of attention, temporal predictions, and sequential learning. While beat perception is a widespread human ability that does not seem to require much musical experience (Merchant et al., 2015; Winkler et al., 2009), I show that many factors influence how we perceive a beat, including attention, musical abilities, and the type of accent present in rhythm. I also show that many processes that may seem unrelated to beat perception, like sequential learning, contribute to the predictions we make in auditory rhythm. The ability to synchronize to a musical beat may therefore be regarded as the result of a collection of processes that together guide our predictions in music. Each of these processes may require a specific set of skills and experience. Different groups of people may rely on different subsets of processes contributing to beat perception. Thinking about beat perception as a construct of several different components provides a lot of room for future explorations of this intriguing human ability.

References

Abecasis, D., Brochard, R., Del Río, D., Dufour, A., & Ortiz, T. (2009). Brain lateralization of metrical accenting in musicians. *Annals of the New York Academy of Sciences*, *1169*, 74–78. doi:10.1111/j.1749-6632.2009.04766.x

Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception*, *22*(3), 549–562. doi:10.1525/mp.2005.22.3.549

Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. N. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, *82*(5), 356–368. doi:10.1016/0013-4694(92)90005-3

Arnason, U., Gullberg, A., Janke, A., Kullberg, M., Lehman, N., Petrov, E. A., & Väinölä, R. (2006). Pinniped phylogeny and a new hypothesis for their origin and dispersal. *Molecular Phylogenetics and Evolution*, *41*(2), 345–354. doi:10.1016/j.ympev.2006.05.022

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, *41*(3), 254–311. doi:10.1006/cogp.2000.0738

Bauer, A.-K. R., Jaeger, M., Thorne, J. D., Bendixen, A., & Debener, S. (2015). The auditory dynamic attending theory revisited: A closer look at the pitch comparison task. *Brain Research*, *1626*, 198–210. doi:10.1016/j.brainres.2015.04.032

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*(2), 120–131. doi:10.1016/j.ijpsycho.2011.08.003

Bendixen, A., Schröger, E., & Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. *The Journal of Neuroscience*, *29*(26), 8447–8451. doi:10.1523/JNEUROSCI.1493-09.2009

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., … Henrik Ehrsson, H. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, *45*(1), 62–71. doi:10.1016/j.cortex.2008.07.002

Bergeson, T. R., & Trehub, S. E. (2006). Infant's perception of rhythmic patterns. *Music Perception*, *23*(4), 345–360. doi:10.1525/mp.2006.23.4.345

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. doi:10.1016/j.cognition.2005.11.007

Bolger, D., Trost, W. J., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica*, *142*(2), 238–244. doi:10.1016/j.actpsy.2012.11.012

Bouwer, F. L., & Honing, H. (2015). Temporal attending and prediction influence the perception of metrical rhythm: Evidence from reaction times and ERPs. *Frontiers in Psychology*, *6*(July), 1094. doi:10.3389/fpsyg.2015.01094

Bouwer, F. L., Van Zuijen, T. L., & Honing, H. (2014). Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study. *PLoS ONE*, *9*(5), e97467. doi:10.1371/journal.pone.0097467

Bouwer, F. L., Werner, C. M., Knetemann, M., & Honing, H. (2016). Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm. *Neuropsychologia, 85*(May), 80-90. doi:10.1016/j.neuropsychologia.2016.02.018

Brochard, R., Abecasis, D., Potter, D. D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*(4), 362–366. doi:10.1111/1467-9280.24441

Cameron, D. J., & Grahn, J. A. (2014). Enhanced timing abilities in percussionists generalize to rhythms without a musical beat. *Frontiers in Human Neuroscience*, *8*(December), 1003. doi:10.3389/fnhum.2014.01003

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, *42*(12), 1150–1152. doi:10.1111/j.1365-2923.2008.03172.x

Carlyon, R. P. (1988). The development and decline of forward masking. *Hearing Research*, *32*(1), 65–79. doi:10.1016/0378-5955(88)90147-5

Carpenter, S. (2012). Psychology's bold initiative. *Science*, *335*(March), 1558–1560. doi:10.1126/science.335.6076.1558

Chait, M., Ruff, C. C., Griffiths, T. D., & McAlpine, D. (2012). Cortical responses to changes in acoustic regularity are differentially modulated by attentional load. *NeuroImage*, *59*(2), 1932–1941. doi:10.1016/j.neuroimage.2011.09.006

Chapin, H. L., Zanto, T. P., Jantzen, K. J., Kelso, J. A. S., Steinberg, F. L., & Large, E. W. (2010). Neural responses to complex auditory rhythms: The role of attending. *Frontiers in Psychology*, *1*(December), 224. doi:10.3389/fpsyg.2010.00224

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008a). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, *18*(12), 2844–2854. doi:10.1093/cercor/bhn042

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008b). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239. doi:10.1162/jocn.2008.20018

Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *NeuroImage*, *32*(4), 1771–1781. doi:10.1016/j.neuroimage.2006.04.207

Christensen, R. H. B. (2015). Ordinal - Regression models for ordinal data. R package version 2015.6-28. Retrieved from http://www.cran.r-project.org/package=ordinal/

Cirelli, L. K., Bosnyak, D., Manning, F. C., Spinelli, C., Marie, C., Fujioka, T., … Trainor, L. J. (2014). Beat-induced fluctuations in auditory cortical beta-band activity: Using EEG to measure age-related changes. *Frontiers in Psychology*, *5*(July), 742. doi:10.3389/fpsyg.2014.00742

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. doi:10.1017/S0140525X12000477

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 473–500). New York: Academic Press.

Clarke, E. F., & Cook, N. (2004). *Empirical musicology: Aims, methods, prospects.* Oxford, England: Oxford University Press.

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24–30. doi:10.1016/S0168-5597(98)00033-1

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, *127*(2), 412–427. doi:10.1037/a0032345

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago, IL: University of Chicago Press.

Csépe, V., Karmos, G., & Molnár, M. (1987). Evoked potential correlates of stimulus deviance during wakefulness and sleep in cat -animal model of mismatch negativity. *Electroencephalography and Clinical Neurophysiology*, *66*(6), 571–578. doi:10.1016/0013-4694(87)90103-9

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*(June), 437. doi:10.3389/fnhum.2014.00437

Dau, T., Püschel, D., & Kohlrausch, A. (1996). A quantitative model of the "effective" signal processing in the auditory system. II. Simulations and measurements. *The Journal of the Acoustical Society of America*, *99*(6), 3623–3631. doi:10.1121/1.414960

Dauvergne, C., Begel, V., Benoit, C.-E., Kotz, S. A., & Dalla Bella, S. (2015). Battery for the assessment of auditory sensorimotor and timing abilities (BAASTA): A rehabilitation perspective. In *Annals of Physical and Rehabilitation Medicine* (Vol. 58S, p. e76). Elsevier Masson SAS. doi:10.1016/j.rehab.2015.07.177

De Waal, F., & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, *14*(5), 201–207. doi:10.1016/j.tics.2010.03.003

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. doi:10.1016/j.jneumeth.2003.10.009

Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research*, *28*(1), 29–42. doi:10.1076/jnmr.28.1.29.3123

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32*(3), 341–65. doi:10.1068/p3370

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288. doi:10.1016/S0010-0277(00)00106-2

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*(1), 1–23. doi:10.2307/40285899

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 264–280. doi:10.1037/a0013482

Escoffier, N., Herrmann, C. S., & Schirmer, A. (2015). Auditory rhythms entrain visual processes in the human brain: Evidence from evoked oscillations and event-related potentials. *NeuroImage*, *111*, 267–276. doi:10.1016/j.neuroimage.2015.02.024

Escoffier, N., Sheng, D. Y. J., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychologica*, *135*(1), 12–16. doi:10.1016/j.actpsy.2010.04.005

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*(December), 215. doi:10.3389/fnhum.2010.00215

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*(1), 173–215. doi:10.1016/j.cognition.2005.11.009

Fitch, W. T. (2009). Biology of music: Another one bites the dust. *Current Biology*, *19*(10), 403–404. doi:10.1016/j.cub.2009.04.004

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, *7*(October), 68. doi:10.3389/fnsys.2013.00068

Fitzgerald, P., & Picton, T. W. (1983). Event-related potentials recorded during the discrimination of improbable stimuli. *Biological Psychology*, *17*(4), 241–276. doi:10.1016/0301-0511(83)90003-0

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*(1), 152–170. doi:10.1111/j.1469-8986.2007.00602.x

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *Psychology of Music* (1st ed., pp. 149–180). New York, NY: Academic Press.

Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836. doi:10.1098/rstb.2005.1622

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *The Journal of Neuroscience*, *35*(45), 15187–15198. doi:10.1523/JNEUROSCI.2397-15.2015

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, *1169*, 89–92. doi:10.1111/j.1749-6632.2009.04779.x

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic Beta oscillations. *The Journal of Neuroscience*, *32*(5), 1791–1802. doi:10.1523/JNEUROSCI.4107-11.2012

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*(6), 1010–1021. doi:10.1162/0898929041502706

Gardner, H. J., & Martin, M. a. (2007). Analyzing ordinal scales in studies of virtual environments: Likert or lump it ! *Technology*, *16*(4), 439 – 446. doi:10.1162/pres.16.4.439

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. doi:10.1016/j.tics.2011.11.014

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception - training increases hierarchical metre processing. *The European Journal of Neuroscience*, *32*(11), 1979–1985. doi:10.1111/j.1460-9568.2010.07462.x

Geiser, E., Ziegler, E., Jäncke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, *45*(1), 93–102. doi:10.1016/j.cortex.2007.09.010

Gerry, D. W., Faux, A. L., & Trainor, L. J. (2010). Effects of Kindermusik training on infants' rhythmic enculturation. *Developmental Science*, *13*(3), 545–551. doi:10.1111/j.1467-7687.2009.00912.x

Grahn, J. A. (2009a). Neuroscientific investigations of musical rhythm: Recent advances and future challenges. *Contemporary Music Review*, *28*(3), 251–277. doi:10.1080/07494460903404360

Grahn, J. A. (2009b). The role of the basal ganglia in beat perception: Neuroimaging and neuropsychological investigations. *Annals of the New York Academy of Sciences*, *1169*, 35–45. doi:10.1111/j.1749-6632.2009.04553.x

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. *Topics in Cognitive Science*, *4*(4), 1–22. doi:10.1111/j.1756-8765.2012.01213.x

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906. doi:10.1162/jocn.2007.19.5.893

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: Audition primes vision, but not vice versa. *NeuroImage*, *54*(2), 1231–1243. doi:10.1016/j.neuroimage.2010.09.033

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, *47*(4), 1894–1903. doi:10.1016/j.neuroimage.2009.04.039

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience*, *29*(23), 7540–7548. doi:10.1523/JNEUROSCI.2018-08.2009

Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, *23*, 913–921. doi:10.1093/cercor/bhs083

Grahn, J. A., & Schuit, D. (2012). Individual differences in rhythmic ability: Behavioral and neuroimaging investigations. *Psychomusicology: Music, Mind, and Brain*, *22*(2), 105–121. doi:10.1037/a0031188

Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11597–11601. doi:10.1073/pnas.0910473107

Grube, M., & Griffiths, T. D. (2009). Metricality-enhanced temporal encoding and the subjective perception of rhythmic sequences. *Cortex*, *45*(1), 72–79. doi:10.1016/j.cortex.2008.01.006

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, *50*(4), 354–377. doi:10.1016/j.cogpsych.2004.09.003

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 956–974. doi:10.1037/0096-1523.30.5.956

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, *16*(1), 48–55. doi:10.1111/j.0956-7976.2005.00779.x

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(35), 12639–12643. doi:10.1073/pnas.0504254102

Hari, R., Pelizzone, M., Mäkelä, J. P., Hällström, J., Leinonen, L., & Lounasmaa, O. V. (1987). Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology*.

Haroush, K., Hochstein, S., & Deouell, L. Y. (2010). Momentary fluctuations in allocation of attention: Cross-modal effects of visual task load on auditory discrimination. *Journal of Cognitive Neuroscience*, *22*(7), 1440–1451. doi:10.1162/jocn.2009.21284

Hasegawa, A., Okanoya, K., Hasegawa, T., & Seki, Y. (2011). Rhythmic synchronization tapping to an audio–visual metronome in budgerigars. *Scientific Reports*, *1*, 120. doi:10.1038/srep00120

Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Scientific Reports*, *3*, 1566. doi:10.1038/srep01566

Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception*, *2*(1), 62–86. doi:10.1163/22134468-00002011

Herholz, S. C., Lappe, C., & Pantev, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, *10*, 42. doi:10.1186/1471-2202-10-42

Herrmann, B., Henry, M. J., Haegens, S., & Obleser, J. (2016). Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *NeuroImage*, *124*, 487–497. doi:10.1016/j.neuroimage.2015.09.019

Hoch, L., Tyler, M. D., & Tillmann, B. (2012). Regularity of unit length boosts statistical learning in verbal and nonverbal artificial languages. *Psychonomic Bulletin & Review*, 142–147. doi:10.3758/s13423-012-0309-8

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, *1252*(1), 85–91. doi:10.1111/j.1749-6632.2011.06402.x

Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *Psychology of Music* (3rd ed., pp. 369–404). London, UK: Academic Press. doi:10.1016/B978-0-12-381460-9.00009-2

Honing, H., Bouwer, F. L., & Háden, G. P. (2014). Perceiving temporal regularity in music: The role of auditory event-related potentials (ERPs) in probing beat perception. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of Interval Timing* (pp. 305–323). New York: Springer. doi:10.1007/978-1-4939-1782-2_16

Honing, H., & Ladinig, O. (2009). Exposure influences expressive timing judgments in music. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(1), 281–288. doi:10.1037/a0012732

Honing, H., Ladinig, O., Háden, G. P., & Winkler, I. (2009). Is beat induction innate or learned? Probing emergent meter perception in adults and newborns using event-related brain potentials. *Annals of the New York Academy of Sciences*, *1169*, 93–96. doi:10.1111/j.1749-6632.2009.04761.x

Honing, H., Merchant, H., Háden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus monkeys (Macaca mulatta) detect rhythmic groups in music, but not the beat. *PLoS ONE*, *7*(12), e51369. doi:10.1371/journal.pone.0051369

Honing, H., & Ploeger, A. (2012). Cognition and the evolution of music: Pitfalls and prospects. *Topics in Cognitive Science*, *4*(2012), 513–524. doi:10.1111/j.1756-8765.2012.01210.x

Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*, 20140088. doi:10.1098/rstb.2014.0088

Horváth, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, *79*(2), 139–147. doi:10.1016/j.biopsycho.2008.04.001

Iversen, J. R., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. In *10th Intl. Conf. on Music Perception and Cognition (ICMPC)* (p. 465). Sapporo, Japan.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, *1169*, 58–73. doi:10.1111/j.1749-6632.2009.04579.x

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*(1), 33–72. doi:10.1016/j.cognition.2005.11.005

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, *38*(12), 1217–1218. doi:10.1111/j.1365-2929.2004.02012.x

Jaramillo, M., Paavilainen, P., & Näätänen, R. N. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters*, *290*(2), 101–104. doi:10.1121/1.414960

Jones, M. R. (2009). Musical time. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford Handbook of Music Psychology* (pp. 81–92). Oxford: Oxford University Press.

Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491. doi:10.1037/0033-295X.96.3.459

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319. doi:10.1111/1467-9280.00458

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology*, *51*(4), 271–291. doi:10.1037/1196-1961.51.4.271

Karns, C. M., & Knight, R. T. (2009). Intermodal auditory, visual, and tactile attention modulates early stages of neural processing. *Journal of Cognitive Neuroscience*, *21*(4), 669–683. doi:10.1162/jocn.2009.21037

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, *20*(2), 228–242. doi:10.3758/s13423-012-0359-y

Kok, P., Rahnev, D. A., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*(9), 2197–2206. doi:10.1093/cercor/bhr310

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology*, *74*(1), 1–19. doi:10.1016/j.biopsycho.2006.06.001

Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting cortical and basal ganglia networks underlying finding and tapping to the musical beat. *Journal of Cognitive Neuroscience*, *25*(3), 401–420. doi:10.1162/jocn_a_00325

Kung, S.-J., Tzeng, O. J. L., Hung, D. L., & Wu, D. H. (2011). Dynamic allocation of attention to metrical and grouping accents in rhythmic sequences. *Experimental Brain Research*, *210*(2), 269–282. doi:10.1007/s00221-011-2630-2

Ladinig, O., Honing, H., Háden, G. P., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception*, *26*(4), 377–386. doi:10.1525/mp.2009.26.4.377

Ladinig, O., Honing, H., Háden, G. P., & Winkler, I. (2011). Erratum to Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception*, *26*(4), 444. doi:10.1525/mp.2011.28.4.444

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, *69*(1), 127–137. doi:10.1016/j.bandc.2008.06.004

Lange, K. (2013). The ups and downs of temporal orienting: A review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Frontiers in Human Neuroscience*, *7*(June), 263. doi:10.3389/fnhum.2013.00263

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, *19*(4), 527–566. doi:10.1016/S0167-9457(00)00026-9

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of time* (pp. 189–231). Bingley, UK: Emerald Group Publishing. doi:10.1016/B978-0-08046-977-5.00006-5

Large, E. W., & Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo (Pan paniscus). *Journal of Comparative Psychology*, *129*(4), 317–328. doi:10.1037/com0000011

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, *9*(November), 159. doi:10.3389/fnsys.2015.00159

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. doi:10.1037/0033-295X.106.1.119

Leow, L., & Grahn, J. A. (2014). Neural mechanisms of rhythm perception: Present findings and future directions. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of Interval Timing* (pp. 325–338). New York: Springer. doi:10.1007/978-1-4939-1782-2

Lerdahl, F., & Jackendoff, R. (1983a). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Lerdahl, F., & Jackendoff, R. (1983b). An overview of hierarchical structure in music. *Music Perception*, *1*(2), 229–252. doi:10.2307/40285257

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception*, *19*(4), 529–550. doi:10.1525/mp.2002.19.4.529

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. (2nd ed.). Oxford: Oxford University Press.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, *1*(4), 424–441.

Luck, S. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Madison, G. S., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1578–1594. doi:10.1037/a0024323

McAuley, J. D., & Fromboluti, E. K. (2014). Attentional entrainment and perceived event duration. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*, 20130401. doi:10.1098/rstb.2013.0401

Mento, G. (2013). The passive CNV: Carving out the contribution of task-related processes to expectancy. *Frontiers in Human Neuroscience*, *7*(December), 827. doi:10.3389/fnhum.2013.00827

Merchant, H., Grahn, J. A., Trainor, L. J., Rohrmeier, M. A., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*, 20140093. doi:10.1098/rstb.2014.0093

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*(January), 274. doi:10.3389/fnins.2013.00274

Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, *45*(4), 224–240. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=10272

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research*, *308*, 84–97. doi:10.1016/j.heares.2013.09.012

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS ONE*, *9*(2), e89642. doi:10.1371/journal.pone.0089642

Näätänen, R. N., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. doi:10.1016/j.clinph.2007.04.026

Näätänen, R. N., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375–425. doi:10.1111/j.1469-8986.1987.tb00311.x

Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience and Biobehavioral Reviews*, *37*, 2564–2570. doi:10.1016/j.neubiorev.2013.08.003

Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. N. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, *20*(1), 26–36. doi:10.1016/j.cogbrainres.2003.12.011

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1658). doi:10.1098/rstb.2013.0393

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234–10240. doi:10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience*, *32*(49), 17572–17581. doi:10.1523/JNEUROSCI.3203-12.2012

Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex*, *25*(3), 736–747. doi:10.1093/cercor/bht261

Paavilainen, P., Arajärvi, P., & Takegata, R. (2007). Preattentive detection of nonsalient contingencies between auditory features. *NeuroReport*, *18*(2), 159–63. doi:10.1097/WNR.0b013e328010e2ac

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*(4), 409–464.

Pashler, H., & Wagenmakers, E.-J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*(6), 528–530. doi:10.1177/1745691612465253

Patel, A. D. (2006). Musical rhythm, lingistic rhythm, and human evolution. *Music Perception*, *24*(1), 99–104. doi:10.1525/mp.2006.24.1.99

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

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*(May), 57. doi:10.3389/fnsys.2014.00057

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*(10), 827–830. doi:10.1016/j.cub.2009.03.038

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the Event-Related Potential. *International Journal of Medical Sciences*, *2*(4), 147–154. doi:10.7150/ijms.2.147

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 302–313. doi:10.1016/j.neuroimage.2009.12.019

Pecenka, N., Engel, A., & Keller, P. E. (2013). Neural correlates of auditory temporal predictions during sensorimotor synchronization. *Frontiers in Human Neuroscience*, *7*(August), 380. doi:10.3389/fnhum.2013.00380

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*, 58–75. doi:10.1196/annals.1284.006

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2011). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception*, *28*(1), 3–14. doi:10.1525/mp.2010.28.1.3.The

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430. doi:10.1126/science.1110922

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II: effects of attention. *Electroencephalography and Clinical Neurophysiology*, *36*, 191–199. doi:10.1016/0013-4694(74)90156-4

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography and Clinical Neurophysiology*, *36*, 179–190. doi:10.1016/0013-4694(74)90155-2

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. doi:10.1016/j.clinph.2007.04.019

Potter, D. D., Fenwick, M., Abecasis, D., & Brochard, R. (2009). Perceiving rhythm where none exists: Event-related potential (ERP) correlates of subjective accenting. *Cortex*, *45*(1), 103–109. doi:10.1016/j.cortex.2008.01.004

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*(4), 411–440. doi:10.2307/40285311

Povel, D.-J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, *30*(6), 565–572. doi:10.3758/BF03202011

Praamstra, P., Kourtis, D., Kwok, H. F., & Oostenveld, R. (2006). Neurophysiology of implicit timing in serial choice reaction-time performance. *The Journal of Neuroscience*, *26*(20), 5448–5455. doi:10.1523/JNEUROSCI.0440-06.2006

Quené, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica*, *62*(1), 1–13. doi:10.1159/000087222

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992. doi:10.3758/BF03206433

Repp, B. H. (2010). Do metrical accents create illusory phenomenal accents? *Attention, Perception & Psychophysics*, *72*(5), 1390–1403. doi:10.3758/APP

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1085–1099. doi:10.1037/0096-1523.28.5.1085

Rimmele, J., Jolsvai, H., & Sussman, E. S. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. *Journal of Cognitive Neuroscience*, *23*(5), 1136–1147. doi:10.1162/jocn.2010.21437

Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research*, *1077*(1), 135–143. doi:10.1016/j.brainres.2006.01.043

Sabri, M., & Campbell, K. B. (2001). Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. *Cognitive Brain Research*, *12*, 171–180. doi:10.1016/S0926-6410(01)00026-X

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. doi:10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. doi:10.1016/S0010-0277(98)00075-4

Sanabria, D., & Correa, Á. (2013). Electrophysiological evidence of temporal preparation driven by rhythms in audition. *Biological Psychology*, *92*(2), 98–105. doi:10.1016/j.biopsycho.2012.11.012

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, *19*(10), 831–836. doi:10.1016/j.cub.2009.03.061

Schaefer, R. S., & Overy, K. (2015). Motor responses to a steady beat. *Annals of the New York Academy of Sciences*, *1337*(1), 40–44. doi:10.1111/nyas.12717

Schaefer, R. S., Vlek, R. J., & Desain, P. (2010). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological Research*, 95–106. doi:10.1007/s00426-010-0293-4

Schafer, E. W. P., Amochaev, a., & Russell, M. J. (1981). Knowledge of stimulus timing attenuates human evoked cortical potentials. *Electroencephalography and Clinical Neurophysiology*, *52*(1), 9–17. doi:10.1016/0013-4694(81)90183-8

Schmidt-Kassow, M., & Kotz, S. A. (2009). Attention and perceptual regularity in speech. *NeuroReport*, *20*, 1643–1647. doi:10.1097/WNR.0b013e328333b0c6

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. doi:10.1016/j.tins.2008.09.012

Schröger, E., Bendixen, A., Trujillo-Barreto, N. J., & Roeber, U. (2007). Processing of abstract rule violations in audition. *PLoS ONE*, *2*(11). doi:10.1371/journal.pone.0001131

Schröger, E., Kotz, S. A., & SanMiguel, I. (2015). Bridging prediction and attention in current research on perception and action. *Brain Research*, *1626*, 1–13. doi:10.1016/j.brainres.2015.08.037

Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *The European Journal of Neuroscience*, *41*(5), 641–664. doi:10.1111/ejn.12816

Schröger, E., & Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. *Neuroscience Letters*, *193*, 185–188. doi:10.1016/0304-3940(95)11696-T

Schwartze, M., Farrugia, N., & Kotz, S. A. (2013). Dissociation of formal and temporal predictability in early auditory evoked potentials. *Neuropsychologia*, *51*(2), 320–325. doi:10.1016/j.neuropsychologia.2012.09.037

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, *87*(1), 146–151. doi:10.1016/j.biopsycho.2011.02.021

Selchenkova, T., François, C., Schön, D., Corneyllie, A., Perrin, F., & Tillmann, B. (2014). Metrical presentation boosts implicit learning of artificial grammar. *PloS One*, *9*(11), e112233. doi:10.1371/journal.pone.0112233

Selchenkova, T., Jones, M. R., & Tillmann, B. (2014). The influence of temporal regularities on the implicit learning of pitch structures. *The Quarterly Journal of Experimental Psychology*, *67*(12), 2360–2380. doi:10.1080/17470218.2014.929155

Shmulevich, I., & Povel, D.-J. (2000). Measures of temporal pattern complexity. *Journal of New Music Research*, *29*(1), 61–69. doi:10.1076/0929-8215(200003)29:01;1-P;FT061

Smith, L. M., & Honing, H. (2008). Time–frequency representation of musical rhythm by continuous wavelets. *Journal of Mathematics and Music*, *2*(2), 81–97. doi:10.1080/17459730802305336

Snyder, J. S., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception*, *18*(4), 455–489. doi:10.1525/mp.2001.18.4.455

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117–126. doi:10.1016/j.cogbrainres.2004.12.014

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia*, *51*(10), 1952–1963. doi:10.1016/j.neuropsychologia.2013.06.027

Sussman, E. S. (2007). A new view on the MMN and attention debate. *Journal of Psychophysiology*, *21*(3), 164–175. doi:10.1027/0269-8803.21.34.164

Takegata, R., & Morotomi, T. (1999). Integrated neural representation of sound and temporal features in human auditory sensory memory: An event-related potential study. *Neuroscience Letters*, *274*(3), 207–210. doi:10.1016/S0304-3940(99)00711-9

Te Woerd, E. S., Oostenveld, R., de Lange, F. P., & Praamstra, P. (2014). A shift from prospective to reactive modulation of beta-band oscillations in Parkinson's disease. *NeuroImage*, *100*, 507–519. doi:10.1016/j.neuroimage.2014.06.039

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience*, *31*(10), 3805–3812. doi:10.1523/JNEUROSCI.5561-10.2011

Temperley, D. (2013). Computational models of music cognition. In D. Deutsch (Ed.), *The Psychology of Music* (3rd ed., pp. 327–368). London, UK: Academic Press. doi:10.1016/B978-0-12-381460-9.00008-0

Tierney, A. T., & Kraus, N. (2013). Neural responses to sounds presented on and off the beat of ecologically valid music. *Frontiers in Systems Neuroscience*, *7*(May), 14. doi:10.3389/fnsys.2013.00014

Tierney, A. T., & Kraus, N. (2014). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, *27*(2), 400–408. doi:10.1162/jocn_a_00704

Tiitinen, H., May, P. J. C., Reinikainen, K., & Näätänen, R. N. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, *372*, 90–92. doi:10.1038/372090a0

Tillmann, B., & McAdams, S. (2004). Implicit learning of musical timbre sequences: statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1131–1142. doi:10.1037/0278-7393.30.5.1131

Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*, 20140089. doi:10.1098/rstb.2014.0089

Tranchant, P., & Vuvan, D. T. (2015). Current conceptual challenges in the study of rhythm processing deficits. *Frontiers in Neuroscience*, *9*(June), 197. doi:10.3389/fnins.2015.00197

Trehub, S. E., Becker, J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140096. doi:10.1098/rstb.2014.0096

Van Zuijen, T. L., Simoens, V. L., Paavilainen, P., Näätänen, R. N., & Tervaniemi, M. (2006). Implicit, intuitive, and explicit knowledge of abstract regularities in a sound sequence: An event-related brain potential study. *Journal of Cognitive Neuroscience*, *18*(8), 1292–1303. doi:10.1162/jocn.2006.18.8.1292

Van Zuijen, T. L., Sussman, E. S., Winkler, I., Näätänen, R. N., & Tervaniemi, M. (2004). Grouping of sequential sounds-an event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience*, *16*(2), 331–338. doi:10.1162/089892904322984607

Van Zuijen, T. L., Sussman, E. S., Winkler, I., Näätänen, R. N., & 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*(2-3), 270–276. doi:10.1016/j.cogbrainres.2004.10.007

Vuust, P., Gebauer, L. K., & Witek, M. a. G. (2014). Neural underpinnings of music: The polyrhythmic brain. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of Interval Timing* (pp. 339–355). New York, NY: Springer. doi:10.1007/978-1-4939-1782-2_18

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music-brain responses to rhythmic incongruity. *Cortex*, *45*(1), 80–92. doi:10.1016/j.cortex.2008.05.014

Vuust, P., Pallesen, K. J., Bailey, C., Van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, *24*(2), 560–564. doi:10.1016/j.neuroimage.2004.08.039

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*(October), 1111. doi:10.3389/fpsyg.2014.01111

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electrical sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384. doi:10.1038/203380a0

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(3), 147–163. doi:10.1027/0269-8803.21.34.147

Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, *83*(2), 132–143. doi:10.1016/j.ijpsycho.2011.10.001

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7), 2468–2471. doi:10.1073/pnas.0809035106

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(18), 8722–8726.

Woldorff, M. G., & Hillyard, S. a. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*(3), 170–191. doi:10.1016/0013-4694(91)90136-R

Womelsdorf, T., & Fries, P. (2007). The role of neuronal synchronization in selective attention. *Current Opinion in Neurobiology*, *17*(2), 154–160. doi:10.1016/j.conb.2007.02.002

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception & Psychophysics*, *72*(8), 2031–2046. doi:10.3758/APP.72.8.2031

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

Zanto, T. P., Large, E. W., Fuchs, A., & Kelso, J. A. S. (2005). Gamma-band responses to perturbed auditory sequences: Evidence for synchronization of perceptual processes. *Music Perception*, *22*(3), 531– 547. doi:10.1525/mp.2005.22.3.531

Zanto, T. P., Snyder, J. S., & Large, E. W. (2006). Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology*, *2*(2), 221–231. doi:10.2478/v10053-008-0057-5

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, *102*(6), 3191–3202. doi:10.1152/jn.00066.2009

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews. Neuroscience*, *8*(7), 547–558. doi:10.1038/nrn2152

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 5768–5773. doi:10.1073/pnas.1000121107

Zwicker, E. (1984). Dependence of post-masking on masker duration and its relation to temporal effects in loudness. *Journal of the Acoustic Society of America*, *75*(1), 219–223. doi:10.1121/1.390398

Samenvatting

Wat hebben we nodig om de maat te horen? De invloed van aandacht, muzikale vaardigheden, en accenten op het waarnemen van metrisch ritme

Overal ter wereld maken mensen samen muziek. Het mechanisme dat ons in staat stelt om te synchroniseren met muziek is *maatgevoel*. In dit proefschrift onderzoek ik welke factoren nodig zijn om de maat te kunnen horen in een ritme. Ik bekijk hoe maatgevoel beïnvloed wordt door aandacht, muzikale vaardigheden, en de accenten in een ritme die de maat aangeven. Daarnaast onderzoek ik welke processen ten grondslag liggen aan maatgevoel en bespreek ik een aantal methodologische zaken die van belang zijn wanneer maatgevoel onderzocht wordt met behulp van *event-related potentials* (ERPs).

De maat kan worden afgeleid uit de structuur van accenten in een ritme. In een online experiment laten we zien dat muziekles de luisteraar gevoeliger maakt voor deze structuur. Daarnaast laten we zien dat luisteraars kwalitatief verschillend reageren op ritmes waarin de maat is aangegeven met accenten in volume en op ritmes waarin de maat is aangegeven met accenten in de groepering van tonen. Wanneer de maat eenmaal afgeleid is uit een ritme, beïnvloedt de waargenomen maat hoe de daarop volgende tonen in een ritme worden gehoord. We laten zien dat deze invloed kan worden beschouwd als het effect van zowel fluctuaties in aandacht in de tijd als temporele voorspellingen. Het eerste verbetert de verwerking van tonen op de tel, terwijl het tweede de verwerking van onverwachte tonen verbetert.

In een aantal experimenten gebruiken we EEG om maatgevoel te onderzoeken wanneer luisteraars hun aandacht niet op een ritme richten. Maatgevoel wordt verondersteld invloed te hebben op de vroege auditieve verwerking van tonen in een ritme. Deze invloed kan worden gekwantificeerd met behulp van ERPs. In een groep professionele musici vinden we aanwijzingen voor maatgevoel, zelfs wanneer hun aandacht niet op een ritme gericht is. We laten zien dat een ERP component die spontaan wordt uitgelokt door geluid, de P1, groter is voor tonen op de tel dan niet op de tel. In twee daaropvolgende EEG experimenten onderzoeken we maatgevoel door ERP responsen te meten die optreden in reactie op de schending van metrische verwachtingen, zowel op de tel als niet op de tel. Het is bekend dat de grootte van een aantal goed bestudeerde ERP componenten, zoals de *mismatch negativity* (MMN) en de P3a, samenhangt met

de grootte van de schending van een auditieve verwachting. We laten zien dat zelfs wanneer de aandacht niet op een ritme gericht is, onverwachte stiltes en afnames in volume grotere ERP responsen uitlokken op de tel dan niet op de tel. Dit ondersteunt de opvatting dat het mogelijk is de maat te horen wanneer een luisteraar niet op een ritme let.

De hier gepresenteerde onderzoeken geven een genuanceerd beeld van de omstandigheden waaronder wij de maat kunnen horen in muziek. Het lijkt mogelijk te zijn de maat te horen wanneer de aandacht niet op een ritme gericht is, maar aandacht kan wel het effect van muziekles op het waarnemen van de maat beïnvloeden. Op soortgelijke wijze kan het soort accent in een ritme beïnvloeden hoe gemakkelijk de maat kan worden gehoord, maar ook dit kan afhangen van de muzikale ervaring van een luisteraar.

Door dit proefschrift heen laat ik zien dat het belangrijk is om de effecten van maatgevoel te onderscheiden van de effecten van andere factoren die de grootte van ERP responsen kunnen beïnvloeden, zoals akoestische variatie in een ritme en verwachtingen die worden veroorzaakt door het leren van statistische regelmaat in een opeenvolging van geluiden. Tot slot opper ik dat maatgevoel kan worden beschouwd als een vaardigheid die bestaat uit verschillende onderliggende mechanismen, die samen ons in staat stellen te synchroniseren met muziek.

Summary

What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

All over the world, people make music together. The process that allows us to synchronize to music is beat perception. In this dissertation, I explore the ingredients that are necessary for us to perceive a beat in musical rhythm. I examine how beat perception is affected by attention, musical abilities, and the type of accent that indicates the beat in a rhythm. Additionally, I examine the processes underlying beat perception, and I address several methodological issues that arise when probing beat perception with event-related potentials (ERPs).

A beat can be inferred from the structure of accents in a rhythm. Using a web-based setup, we show that musical training increases the sensitivity of a listener to the structure of accents in rhythm. We also show that listeners respond qualitatively differently to rhythms in which the beat is indicated by the temporal grouping of events and rhythms in which the beat is indicated by intensity increases. Once a beat is inferred from a rhythm, a perceived beat influences processing of subsequent rhythmic events. Using a speeded detection task, we show that this influence can be characterized both by temporal fluctuations in attention and by temporal predictions. The former enhances processing of events that coincide with the beat, while the latter enhances the detection of events that are unexpected.

In several experiments, we use EEG to examine beat perception when listeners do not attend to a rhythm. Beat perception is thought to influence early auditory processing of rhythmic events. This influence can be indexed using ERPs. In a group of highly trained musicians, we provide evidence for the presence of beat perception when attention is directed away from a rhythm. We show that the P1 response, which is an obligatory ERP response to sound, is larger for events on the beat than for events off the beat. In two subsequent EEG studies, we probe beat perception by measuring ERP responses to prediction violations occurring both on and off the beat in metrical rhythms. Several well-studied ERP responses are known to index the magnitude of prediction violations, including the mismatch negativity (MMN) and the P3a response. We show that even when attention is not directed at a rhythm, unexpected silences and intensity decreases elicit larger ERP responses on the beat than off the beat. This supports the view that beat perception is possible when attention is directed away from a

rhythm. Musical training only enhances the effects of beat perception on ERP responses to unexpected rhythmic events when attention is directed towards the rhythm.

Together, the collection of studies presented here provides a nuanced picture of the conditions under which we are capable of perceiving a beat in music. Beat perception may be possible when attention is directed away from a rhythm, but attention may interact with the effects of musical training on beat perception. Likewise, the type of accents present in a rhythm can influence how easy it is to perceive a beat, but this may depend on musical training.

Throughout this dissertation, I show that when probing beat perception with ERPs it is important to differentiate the effects of beat perception from other factors that may influence the magnitude of ERP responses, such as acoustic variation in a rhythm and predictions that are generated through statistical learning of a sound sequence. Finally, I propose that beat perception can be regarded as an ability that consists of several different mechanisms that together allow us to synchronize to music.

Titles in the ILLC Dissertation Series

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: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder...": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2010-12: **Martin Mose Bentzen**
Stit, Iit, and Deontic Logic for Action Types

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC DS-2011-08: **Stefan Minica**
Dynamic Logic of Questions

ILLC DS-2011-09: **Raul Andres Leal**
Modalities Through the Looking Glass: A study on coalgebraic modal logic and their applications

ILLC DS-2011-10: **Lena Kurzen**
Complexity in Interaction

ILLC DS-2011-11: **Gideon Borensztajn**
The neural basis of structure in language

ILLC DS-2012-01: **Federico Sangati**
Decomposing and Regenerating Syntactic Trees

ILLC DS-2012-02: **Markos Mylonakis**
Learning the Latent Structure of Translation

ILLC DS-2012-03: **Edgar José Andrade Lotero**
Models of Language: Towards a practice-based account of information in natural language

ILLC DS-2012-04: **Yurii Khomskii**
Regularity Properties and Definability in the Real Number Continuum: idealized forcing, polarized partitions, Hausdorff gaps and mad families in the projective hierarchy.

ILLC DS-2012-05: **David García Soriano**
Query-Efficient Computation in Property Testing and Learning Theory

ILLC DS-2012-06: **Dimitris Gakis**
Contextual Metaphilosophy - The Case of Wittgenstein

ILLC DS-2012-07: **Pietro Galliani**
The Dynamics of Imperfect Information

ILLC DS-2012-08: **Umberto Grandi**
Binary Aggregation with Integrity Constraints

ILLC DS-2012-09: **Wesley Halcrow Holliday**
Knowing What Follows: Epistemic Closure and Epistemic Logic

ILLC DS-2012-10: **Jeremy Meyers**
Locations, Bodies, and Sets: A model theoretic investigation into nominalistic mereologies

ILLC DS-2012-11: **Floor Sietsma**
Logics of Communication and Knowledge

ILLC DS-2012-12: **Joris Dormans**
Engineering emergence: applied theory for game design

ILLC DS-2013-01: **Simon Pauw**
Size Matters: Grounding Quantifiers in Spatial Perception

ILLC DS-2013-02: **Virginie Fiutek**
Playing with Knowledge and Belief

ILLC DS-2013-03: **Giannicola Scarpa**
Quantum entanglement in non-local games, graph parameters and zero-error information theory

ILLC DS-2014-01: **Machiel Keestra**
Sculpting the Space of Actions. Explaining Human Action by Integrating Intentions and Mechanisms

ILLC DS-2014-02: **Thomas Icard**
The Algorithmic Mind: A Study of Inference in Action

ILLC DS-2014-03: **Harald A. Bastiaanse**
Very, Many, Small, Penguins

ILLC DS-2014-04: **Ben Rodenhäuser**
A Matter of Trust: Dynamic Attitudes in Epistemic Logic

ILLC DS-2015-01: **María Inés Crespo**
Affecting Meaning. Subjectivity and evaluativity in gradable adjectives.

ILLC DS-2015-02: **Mathias Winther Madsen**
The Kid, the Clerk, and the Gambler - Critical Studies in Statistics and Cognitive Science

ILLC DS-2015-03: **Shengyang Zhong**
Orthogonality and Quantum Geometry: Towards a Relational Reconstruction of Quantum Theory

ILLC DS-2015-04: **Sumit Sourabh**
Correspondence and Canonicity in Non-Classical Logic

ILLC DS-2015-05: **Facundo Carreiro**
Fragments of Fixpoint Logics: Automata and Expressiveness

ILLC DS-2016-01: **Ivano Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC-DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm