



UNIVERSITÀ DEGLI STUDI  
DI TRENTO

CIMeC - Center for Mind/Brain Sciences

CiMeC

Doctoral School in Cognitive and Brain Sciences

PhD Thesis

**MUSICAL EXPERTISE AND RHYTHM PROCESSING**

**Francesca Perna**

Advisor: Prof. **Veronica Mazza**

Co-advisor: Prof. **Francesco Pavani**

XXX Cycle - March 2018





## Acknowledgments

First and foremost I wish to express my sincere gratitude to my advisor, Prof. Veronica Mazza, for her distinguished continuous support, guidance and motivation through the rough road of the PhD, giving me the moral support and the freedom I needed to move on. My sincere gratitude for each moment spent on this project, for believing in this amazing experience, for allowing me to grow as a person and as a research scientist during these years. I would like to thank her for the brilliant comments and suggestions, but also for the hard questions which incited me to widen my research from various perspectives.

Profound gratitude also goes to my co-advisor, Prof. Francesco Pavani, who has been a truly dedicated mentor. I wish to thank Francesco for his precious support, insightful suggestions, and for having nurtured my enthusiasm for the research conducted during this PhD.

Also I wish to thank Prof. Massimiliano Zampini for being part of my Oversight committee during these years and for his professional support even at hardship.

A special acknowledgement goes to my extraordinary coach supporting me during the most difficult days. Without him I would not have been able to survive! Thank you very much for believing in my research ambitions, thank you for all inspiring conversations that raised in me the motivation to pursue working every time I was almost giving it up.

I am particularly thankful to my family for never stopping motivating me, to infuse great optimism and energy even in the darkest days.

I am infinitely thankful to my determination, courage and ambition. For the endless dedication and love for in this research project.



# Contents

<b>Abstract</b> .....	<b>14</b>
<b>Chapter 1. Introduction</b> .....	<b>18</b>
1.1 Rhythm and Meter as Structures in Time.....	18
1.2 Theories of Temporal Structures.....	19
1.3 Electrophysiological Correlates of Regularity Violation.....	22
1.3.1 Mismatch Negativity (MMN).....	22
1.3.2 Middle Latency Responses (MLR).....	24
1.4 Long-term Musical Training: Musicians as a Model to Study Rhythm and Meter Perception.....	26
1.4.1 Rhythm/Meter Processing and Musical Expertise.....	26
1.5 Cueing Effects on Orienting Attention in Time.....	29
1.6 The present research: aims and hypotheses.....	31
<b>Chapter 2. Behavioral Dynamics of Rhythm and Meter Perception: The Effect of Musical Expertise</b> <b>in Deviance Detection</b> .....	<b>36</b>
2.1 Introduction.....	36
2.2 Experiment 1.....	40
2.3 Materials and Methods .....	40
2.3.1 Participants .....	40
2.3.2 Stimuli.....	41
2.3.3 Procedure.....	43
2.4. Results.....	44
2.4.1 Reaction Times.....	44
2.4.2 Sensitivity ( $d'$ ).....	45

2.4.3 Response Criterion (c).....	46
2.5 Discussion.....	47
2.6 Experiment 2.....	48
2.7 Material and Methods.....	48
2.7.1 Participants.....	48
2.7.2 Stimuli.....	48
2.7.3 Procedure.....	50
2.8 Results.....	50
2.8.1 Reaction Times.....	50
2.8.2 Sensitivity (d').....	50
2.8.3 Response Criterion (c).....	51
2.9 Discussion.....	51
3. General Discussion.....	52
<b>Chapter 3. The time course of auditory deviance processing: the effect of long-term musical</b>	
<b>training .....</b>	<b>59</b>
3.1 Introduction.....	59
3.2 Materials and Methods.....	66
3.2.1 Participants .....	66
3.2.2 Stimuli.....	66
3.2.3 Procedure.....	69
3.3 EEG data collection and pre-processing.....	69
3.3.1 MMN.....	70
3.3.2 MLR.....	71
3.4. Results.....	71
3.4.1 Behavioral Results.....	71



3.4.1.1 Reaction Times.....	72
3.4.1.2 Sensitivity ( $d'$ ).....	73
3.4.1.3 <i>Response Criterion (c)</i> .....	74
3.4.1.4 Musicality Tests.....	75
3.4.2 EEG results.....	75
3.4.3 Mismatch Negativity (MMN).....	75
3.4.3.1 MMN amplitude.....	76
3.4.3.2 MMN latency.....	76
3.4.4 Analyses on deviant and standard stimuli responses (100-300 ms interval).....	77
3.4.4.1 Mean amplitude.....	77
3.4.4.2 Mean latency.....	81
3.4.5 Middle Latency Responses.....	81
3.4.5.1 Na amplitude.....	81
3.4.5.2 P0 amplitude.....	83
3.4.5.3 Control Analysis.....	83
3.5 Discussion.....	83
3.5.1 Behavioral results.....	84
3.5.2 Electrophysiological results.....	84
3.6 Conclusion.....	87

**Chapter 4. Orienting Attention in Time: The Effect of Cue Modality and Musical**

Expertise .....	<b>89</b>
4.1 Introduction.....	89
4.2 Materials and Methods.....	95
4.2.1 Participants.....	95
4.2.2 Stimuli.....	96

4.2.3 Procedure.....	97
4.3 EEG Data Collection and Pre-processing.....	98
4.3.1 CNV.....	99
4.3.2 MMN.....	99
4.3.3 P300.....	99
4.3.4 Steady-State Evoked Potentials (SS-EPs) .....	100
4.4. Results.....	101
4.4.1 Behavioral Results.....	101
4.4.1.2 Inverse Efficiency score (IES).....	102
4.4.1.3 Response frequency around target stimulus onset).....	103
4.4.2 EEG results.....	106
4.4.2.1 Contingent Negative Variation (CNV).....	106
4.4.2.2 CNV over centro-frontal and posterior-occipital Regions of Interest (ROI)....	108
4.4.2.3 MMN.....	109
4.4.2.4 P300.....	110
4.4.2.5 SS-EPs.....	111
4.5 Discussion.....	113
4.5.1 Behavioral results.....	114
4.5.2 EEG results .....	116
4.5.2.1 Pre-target stimulus response .....	116
4.5.2.2 Post-target stimulus response.....	117
4.5.2.3 SS-EPs.....	118
4.6 Conclusion.....	119
<b>Chapter 5. Conclusions .....</b>	<b>121</b>
<b>References .....</b>	<b>133</b>
<b>Appendix.....</b>	<b>148</b>

# List of Figures

## Introduction

Figure 1.1 Example of Mismatch Negativity Component.....	24
Figure 1.2 Example of Middle Latency Responses.....	26

## Experiment 1

Figure 2.1 Stimulus Material Experiment 1. Example of Rhythmic module .....	42
Figure 2.2 Stimulus Material Experiment 1. Example of Non-rhythmic module .....	43
Figure 2.3 Results of Experiment 1.....	47

## Experiment 2

Figure 3.1 Stimulus Material Experiment 2. Example of Rhythmic module .....	49
Figure 3.2 Results of Experiment 2.....	51

## Experiment 3

Figure 4.1 Stimulus Material Experiment 3. Example of Rhythmic module.....	67
Figure 4.2 Stimulus Material Experiment 3. Example of Non-rhythmic module .....	68
Figure 4.3 Behavioral Results of Experiment 3.....	75
Figure 4.4.1. EEG results of Experiment 3. Mismatch Negativity component.....	77
Figure 4.4.2. EEG results of Experiment 3. Neural response to deviants and standards.....	80
Figure 4.4.3. EEG results of Experiment 3. Middle Latency Responses.....	82

## **Experiment 4**

Figure 5.1 Stimulus Material of Experiment 4.....	97
Figure 5.2.1. Behavioral results of Experiment 4. Reaction Times and Correct Detections.....	102
Figure 5.2.2. Behavioral results of Experiment 4. Inverse Efficiency Scores.....	103
Figure 5.2.3. Behavioral results of Experiment 4. Target Anticipations.....	106
Figure 5.3.1 EEG results of Experiment 4. Contingent Negative Variation component.....	108
Figure 5.3.2. EEG results of Experiment 4. Mismatch Negativity and P300 component.....	111
Figure 5.3.3. EEG results of Experiment 4. Steady-State Evoked Potentials .....	113



## Abstract

For decades, researchers have been trying to understand how the human mind/brain processes rhythm and time in general. Within this framework, many studies have explored the influence of long-term musical training on the neural and behavioral correlates of rhythm processing. Some pieces of evidence point to enhanced rhythm processing in musicians as one of the consequences of the structural and functional changes in many brain areas involved in auditory processing, motor synchronization and cognitive control. Yet there is still more controversy than consensus on this field. Indeed, several behavioral and neural studies report opposite results and describe contrasting effects associated to rhythm perception in musicians and non-musicians.

The aim of the project described in this thesis was to shed new light on the effects of long-term musical training on the behavioral and neural correlates of rhythm processing. First, I addressed whether musical expertise influences rhythm processing when this is not task-relevant. Next, I expanded the investigation to the ability of musicians to orient efficiently attention in time. I explored these questions by looking at behavioral and electroencephalographic (EEG) correlates associated to the detection of auditory deviant stimuli.

Chapter 1 is a general introduction to the current knowledge on rhythm processing. After a description of the most influential theories of temporal elaboration, it introduces some of the electrophysiological correlates associated to regularity violation. Then, it provides a detailed description of the neural and behavioral changes triggered by a long-term musical training, focusing on rhythm processing. At the end of this introduction, the aims and hypotheses of each experiment are presented in detail.

In Chapter 2, I describe two behavioral experiments that explored how the processing of different temporal structures (rhythmical, non-rhythmical) influences the detection of deviant stimuli (Experiment 1), and

addressed the role of a refined metrical representation in musicians and non-musicians (Experiment 2). The results revealed an overall superior performance of musicians in all experimental conditions, thus pointing to an enhanced auditory perception as consequence of their musical training. The results also highlight a large-scale processing of rhythm, independent of musical expertise. Furthermore, the results on response speed indicate a refined metrical processing only in musicians. Thus, the first part of the thesis demonstrates that long-term musical training boosts meter processing, whereas (some form of) rhythm processing appears to be overall present in all individuals.

In Chapter 3, I report an EEG experiment which was aimed at exploring the effects of long-term musical training on the neural correlates of auditory deviance violation (Mismatch Negativity – MMN and Middle Latency Responses - MLRs), by inserting deviant stimuli in rhythmical (at strong and weak metrical positions) and non-rhythmical structures. Deviant stimuli within rhythmical structures elicited larger MMN compared to non-rhythmical ones in all participants. Moreover, the MMN was also modulated by meter as showed by the smaller amplitude for deviants at strong than weak positions. Interestingly, a deeper investigation of the neural modulations associated to the strong positions revealed a stronger response in musicians than non-musicians. This demonstrates the prevailing effect of stimulus salience (strong metrical positions and frequency deviant at these positions) over the effect of prediction for musicians. Finally, effect of long-term musical training modulated the MLRs for deviants within rhythmical structures. Taken together, these results indicate an effect of musical expertise at early and late stages of deviance perception, as evidenced by modulations of the MLR and MMN responses.

Chapter 4 reports the results of a cross-modal cueing experiment aimed at investigating the influence of long-term musical training on the ability to orient attention in time using external cues. Besides behavioral and ERP responses, here I examined a particular neural response associated to entrainment: the steady state evoked potential (SS-EP). The results showed that auditory cues greatly facilitate attention orienting in time. Furthermore, targets preceded by short intervals were highly expected and this was visible both at the behavioral (high efficiency and more anticipations) and neural (larger CNV and reduced P300 amplitude) levels. Effects of musical expertise were present only in behavioral data and only when considering the mostly

trained sensory modality (audition). Finally, musicians were less synchronized to the rhythm than non-musicians (reduced SS-EPs). In sum, these results indicate that the auditory modality better guide temporal orienting than the visual one, and that this effect is magnified for musicians. Finally, weaker synchronization to rhythm in musicians may mirror the ease with which they process rhythm.

Chapter 5 is a summary of the main results and of their interpretation.





# Chapter 1

## Introduction

### 1.1 Rhythm and Meter as Structures in Time

Rhythm dictates most of our daily activities. Consider for instance the way in which we walk: our motor acts are rhythmically organized with each movement following the other cyclically. Even at a more basic level, our heartbeats follow a rhythmic pattern. Thus, rhythm can be conceived as a regular distribution of events into coherent patterns in time (Berlyne, 1971). There are also more specific definitions of rhythm, depending on the domain considered. In music, rhythm is defined as a set of temporally organized events encompassing two fundamental levels: (1) the beat and (2) the meter. Previous theories on the processing of rhythm have drawn a distinction between the mechanisms involved in the processing of beat and meter (Large, 2008; Large & Snyder, 2009). According to these theories, beat represents a sequence of regularly recurring events processed sequentially (Cooper & Meyer, 1960; Large, 2008): here, the auditory system elaborates each sound individually and sequentially one after the other. For example, when listening to a piece of music, the beats are those points to which listeners would synchronize their feet. The processing of meter instead involves the association of different strengths (i.e., accents) to the various beats. The presence of accents gives rise to the perception of stronger and weaker beats, and promotes the creation of groups, for example of two (e.g., as in a march) or three (e.g., as in a waltz) elements, which depicts different hierarchical levels of the metrical structure (Martin, 1972; Lerdahl & Jackendoff, 1983). For instance, in the waltz, the sequence of each single sound is part of the first level (beat level or sequential level) of rhythm; at the same time, each first sound of three is perceived as more prominent than the others. This promotes the creation of auditory groups of three sounds that build the second level of rhythm. The prominence of the first sound of three

leads to the perception of a tension between strong and weak beats (i.e., in waltz: strong, weak, weak) which favors the emergence of the metrical structure (Lerdahl and Jackendoff, 1983; London, 2012; Fitch, 2013). For decades, researchers have been trying to understand how the human mind/brain processes rhythmical structures by means of several methodological approaches. In the following sections, I provide some examples of theories exploring the behavioral and neural underpinnings supporting the perception and organization of time and, more specifically, of rhythm. Next, I describe the electrophysiological correlates of deviance detection by focusing on two neural responses associated to regularity violation: the mismatch negativity component (MMN) and the middle latency response (MLR). Most of the descriptions of the two sections will revolve around the comparison between musicians and non-musicians, as musicians represent the ideal model to study rhythm and meter processing. Finally, I examine the effect of endogenous cueing in time orienting of musicians and non-musicians. Here, I will try to bridge two lines of research. The first is related to cueing paradigms. Studies on this topic (Coull & Nobre, 1998; see Niemi & Naatanen, 1981, for a review) have shown that, just as in the case of space, attention can be efficiently oriented in time by means of symbolic cues. The second line of research focuses on the exceptional ability of musicians to encode rhythm, which may act as an additional cue to orient attention to points in time.

## **1.2 Theories of Temporal Structures**

Several theoretical approaches have tried to explain how we perceive and organize temporal structures. The first models describing time processing were the so-called “pacemaker-accumulator” models, introduced in the early 1960s by Creelman (Creelman, 1962) and Treisman (Treisman, 1963). In Creelman’s model (Creelman, 1962), pulses with a fixed frequency (the pacemaker) were considered as an accumulator used to estimate time. Yet, this fixed and not flexible structure failed in discriminating intervals differing slightly in their duration. One year later, Treisman (Treisman, 1963) proposed a more detailed model of time perception. Using the same basic mechanism of Creelman’s model, this model postulated the use of an oscillator as pacemaker, and suggested different processing levels to store and compare time intervals with varying durations. This offered a richest approach to explain how intervals of time of different lengths are

processed. In fact, differing from the fixed structure of Creiman's model, Treisman's model paved the way to more advanced models of time processing and in particular of rhythm, which is often built of different intervals lengths.

Years after the pacemaker-accumulator models, Large and Jones (Large and Jones, 1999) proposed an innovative model of temporal processing outlined in the "Dynamic Attending Theory" (DAT). DAT proposes a dynamic estimation of time; in fact, rather than considering a fixed counting mechanism, as it was previously done, temporal information is encoded by means of a set of oscillators whose frequency can be adapted to multiple intervals' lengths.

DAT is one of the most influential theories used for the analysis of rhythm and meter processing in humans (Jones and Boltz, 1989). It posits that external rhythms capture listeners' attention, resulting in a dynamic fluctuation of attention synchronized to the perceived regularities (entrainment hypothesis). These fluctuations of attentional energy adapt their phase and period based on the metrical structure extracted from a sequence of sounds. This generates specific expectancies and predictions (anticipations) toward each sound of the sequence. In this direction, DAT proposes that attention is maximally oriented toward sounds at strong positions of the metrical structure relative to weak ones. Accordingly, stimuli coinciding at strong positions are highly expected and more effectively processed (Jones, 2009). Further empirical support to DAT comes from behavioral studies observing a facilitation effect prompted by the regularity of an isochronous sequence of sounds on comparative pitches judgments (standard, comparison). Indeed, Jones et al. (2002) demonstrated that listeners were more accurate in judging pitch differences of expected tones within regular sequences compared to pitches embedded in irregular rhythms. These results were interpreted as evidence of a dynamic allocation of attention in time by which tones coinciding with points of maximal attentional energy were best detected.

The Predictive Coding Theory (PC) is another theoretical proposal that explains the effect of predictability (proper of rhythmical sequences) on stimulus perception. In particular, the regularity of the rhythmic sequences is considered as an anticipatory model that, based on different levels of salience attributed to

each sound/beat, is used to create specific expectations towards future events (Vuust & Witek, 2014). As such, sequence regularity should promote better performance for the highly expected/predicted events compared to those less expected/predicted (e.g., Elliott et al., 2014, but see Bouwer & Honing, 2015 for opposite results). PC theory interprets the expectancy driven by rhythm processing in terms of predictive modeling. It postulates that perception follows a Bayesian process by which the brain predicts the causal relationship between sensory percepts in a continuous comparison with previous knowledge (Friston, 2005). In this process, the brain constantly tries to minimize the error between the perceived sensory input and the predicted stimulus based on short- and long- term knowledge accumulated through experience (Vuust & Witek, 2014). Short-term knowledge is built on a moment-to-moment basis, like for example the one that a brief exposure to a rhythmical sequence may provide to a listener. Beside the short-term experience with the experimental manipulations, long-term experiences (such as long-term music training, as in the present thesis) can also influence the ability to make predictions. To maximize the match between prediction and sensory input, a dual mechanism is hypothesized to occur: a backward projection, where the perceived sensory input is compared to previous knowledge, and a forward projection where we use previous knowledge, experience or contextual cues to anticipate and predict sensory inputs (Vuust & Witek, 2014). If there is a mismatch between predicted and perceived sensory input, a residual error signal (i.e., the prediction error) is sent to higher layers via feedforward pathways (Nazimek et al., 2013). Here, considering the dynamics behind rhythm processing, PC suggests two types of mechanisms: a top-down (expectancy-driven) and a bottom-up (stimulus-driven) mechanism (Rao & Ballard, 1999).

Prediction errors have been investigated mainly with oddball paradigms, where rare deviant stimuli are randomly inserted within series of frequent standard stimuli. In terms of electrophysiological correlates, when the predicted stimulus (related to the more frequent, standard stimulus) does not match the actual sensory information (deviant stimuli), the related prediction error gives rise to specific neural responses (i.e., the Mismatch Negativity component (MMN)) (Friston, 2005). Specifically, PC theory postulates that the MMN amplitude is influenced by the magnitude of the expectancy violation (Näätänen et al., 2007). Here, increased

predictability of a stimulus corresponds to a smaller prediction error and to a smaller MMN amplitude (Lecaignard et al., 2015). In the next section, I describe this response more in detail.

### **1.3 Electrophysiological Correlates of Regularity Violation**

One of the effects of temporal expectation is the optimization of behavior for events presented at the expected (regular) moment. When the expected stimulus does not correspond to the actual perception, this regularity violation is reflected at the EEG level in a series of event-related potentials (ERPs). Here, I focus on two ERPs traditionally associated with the perception of deviants in oddball paradigms.

#### **1.3.1 Mismatch Negativity (MMN)**

An example of ERP response associated with regularity violation is the Mismatch Negativity (MMN) component, an index of deviance extraction from a sequence of regular and predictable stimuli (Naatanen et al., 2007). Naatanen and colleagues (Naatanen et al., 1978) described the MMN for the first time as a negative component of the auditory ERPs - originating in the auditory and frontal cortices (van Zuijen et al., 2005), which usually peaks at 200-250 ms after the occurrence of a deviant stimulus. This response has a maximal amplitude over frontal and central scalp locations (Naatanen, 1992). The MMN results from an automatic discriminative process that detects small changes in the acoustic environment and compares them with traces created by the repetitive presentation of sensory stimulation (Naatanen et al., 2010). Thus, the MMN is elicited when an incoming sound is different from the memory traces of a previous input (van Zuijen et al., 2004). This requires that the auditory system has formed a strong representation of the regular and predictable aspect of the (standard) stimuli.

Crucially, the MMN is not simply the result of different states of refractoriness of neurons responding to specific acoustic features of the standard or deviant stimuli (Naatanen, Jacobsen, & Winkler, 2005). This was demonstrated in controlled oddball paradigms, where the auditory stimuli used as deviants in the oddball sequence were inserted, with the same probability, within series of different randomly presented equiprobable auditory stimuli (Schroger & Wolff, 1996; Schroger, 1996). For example, Schroger, 1996

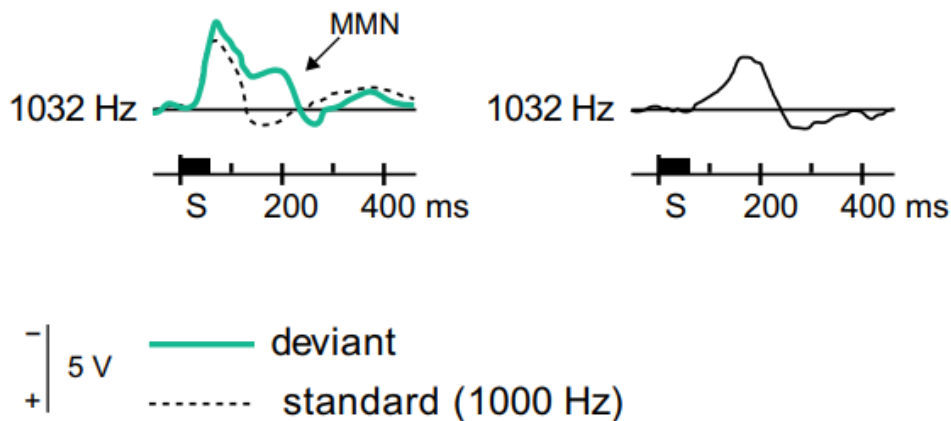
controlled the refractoriness effects using different conditions: MMNs were obtained in oddball blocks with standard and deviant ( $p = .10$ ) sounds of respectively 700 and 750 Hz; and in control blocks, sounds' frequency was 700, 750, 800, 900, 950, 1000, 1050, 1100, or 1200 Hz ( $p = .10$  each). The deviant - control comparison displayed a clear MMN in both oddball and control blocks which could not be explained in terms of differential states of refractoriness (Schroger, 1996).

Earlier studies demonstrated that the MMN is influenced by at least three factors: (1) the deviance magnitude, (2) the variability of the acoustic context (made of standard stimuli), and (3) the type of sequence predictability (Lecaignard et al., 2015). The first factor is related to the fact that a larger MMN amplitude is observed for larger differences between standard and deviant stimuli (Winkler et al., 2009).

Previous studies investigating the effect of the variability of the acoustic context on the MMN (factor 2), manipulated the acoustic properties of standard stimuli (Daikhin & Ahissar, 2012). For example, Daikhin and Ahissar, 2012 used an oddball paradigm in which standard stimuli had variable sound's frequency. Comparing the MMN amplitude elicited by deviants inserted in sequences of standards with different frequencies to the MMN of a standard oddball (all standards with the same frequency), the authors noted that in the first case the MMN amplitude was diminished. This suggested that a varying acoustic context does not allow for a clear standard – deviant differentiation as reflected in a reduction of the MMN amplitude.

The type of sequence predictability (factor 3) could be derived from a set of rules that delineate the statistical dependencies of each sound within a sequence (Lecaignard et al., 2015). It is possible to define local and global rules, the first belonging to short time-scale dependencies (relationship between subsequent sounds) whereas the latter generating large time-scale regularities. For example, consider a sequence of five tones, in which the first four tones are the same and the last can be either the same or a different tone. The short time-scale dependency rule is based on the repetition of the four tones. Hence, the repetition of the four tones predicts the acoustic properties of the fifth tone, which will be equal to the previous ones if the rule is met. Contrariwise, the global rule describes the probability of standard and deviant sequences (xxxxx or xxxxy) within blocks. When the standard sequence (xxxxx) is more probable, the global rule predicts the next sequences to be equal to the most probable ones; instead, when the deviant sequence (xxxxy) is more

probable, the global rule predicts the next sequences to end also in  $\gamma$  (and in contrast with the local rules predictions). Stimulus and sequence predictability affects the amplitude of the MMN response in such a way that there is a decrease of the amplitude with an increase in the stimulus/sequence predictability (Lecaignard et al., 2015).



**Figure 1.1.** The left side of the Figure shows an example of long-latency component of the ERPs elicited by standard (dotted line) and deviant (green line) stimuli. The right side displays an example of Mismatch Negativity component obtained by subtracting the standard stimulus ERP from that of the deviant stimulus. Adapted from Naatanen et al., 2007.

### 1.3.2 Middle Latency Responses (MLR)

Challenging the view that the middle-latency portion of the ERP components is responsible for the perception of stimuli physical properties only, a series of studies showed effects of auditory deviants perception also at these early stages of stimulus elaboration (roughly from 12 to 50 ms post-stimulus onset). Therefore, beside the MMN component, modulations of the Middle Latency Responses (MLRs) reflect also a change-detection process (Grimm et al., 2011) that result from a mismatch between the memory trace of a preceding standard sound and an incoming deviant ones (Alho et al., 2012). The MLRs are a set of positive ("P" waves) and negative ("N" waves) waves occurring between 12-80 ms after sound onset (Grimm et al., 2011; Alho et al., 2012); these responses originate from the inferior colliculus, medial geniculate body, and the primary-secondary auditory cortices (Escera and Malmierca, 2014; Malmierca et al., 2014). The first MLR wave is the

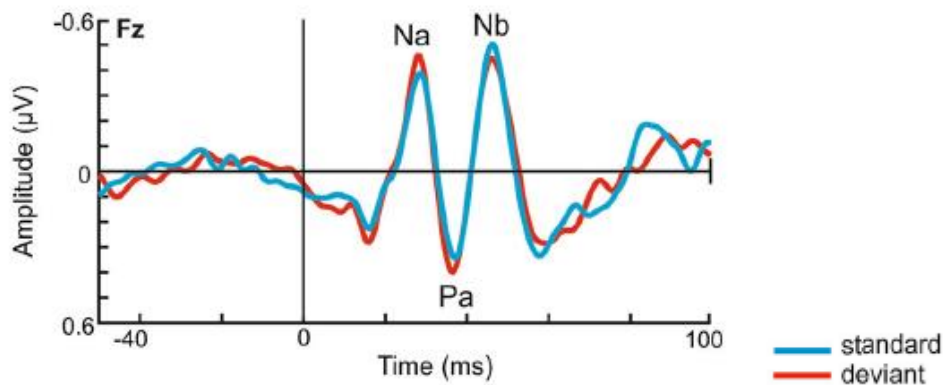


P0 (subcortical origins), followed by Na, Pa and Nb (primary and secondary auditory cortices) (Grimm et al., 2011).

There is not yet a clear understanding of the effects that different stimulus characteristics have on the single waves of the MLR. Some work suggested that each positive-negative wave has its own “specificity”. For instance, modulation of the Na-Pa components are mainly observed following intensity and location changes in the auditory stimulus, while differences in the Nb (and sometimes Pa, Escera et al., 2014) waves are ascribed to frequency changes (Escera and Malmierca, 2014). However, while some studies observed that changes in stimulus frequency, intensity, duration and location influence the latency and amplitude of the MLRs (Borgmann et al., 2001; Escera et al., 2014) others did not confirm these results (Grimm et al., 2011). For example, Grimm and colleagues (Grimm et al., 2011) employed a traditional oddball paradigm where rare and deviant sounds (low and high frequency sounds) were embedded within a train of standard sounds. The authors observed that the Nb wave of the MLRs had a larger amplitude following the detection of rare and unexpected sounds. Yet, the frequency of the deviant sound (low or high) did not have any influence on these components. Importantly, their results were considered a clear hallmark of deviance processing per se since, irrespective of the stimulus frequency (low or high), modulation of the Nb amplitude triggered by deviants’ perception reflected a mismatch between the memory trace formed by the frequent standard stimuli and the actual deviant perception (Grimm et al., 2011).

Moreover, as already noted for the mismatch negativity, deviant predictability affects these earlier components too. Indeed, Lecaigard and colleagues found that predictable deviants resulted in smaller MLRs amplitude compared to unpredictable ones (Lecaigard et al., 2015). This challenged the long-held view by which the MLRs were considered only as the neural correlates of sound physical properties perception (i.e., stimulus frequency) and rather corroborates this response as new electrophysiological markers of auditory deviance detection (Grimm et al., 2011).

Finally, in contrast to what I will describe for the MMN (for which there are numerous studies showing the effects of musical expertise), it is important to note that the effect of rhythm and long-term music training on the MLR has remained unknown.



**Figure 1.2.** Example of Middle Latency Responses elicited by standard (blue line) and deviant (red line) stimuli. Adapted from Escera et al., 2014.

#### 1.4 Long-term Musical Training: Musicians as a Model to Study Rhythm and Meter Perception

The research described in this thesis attempts to investigate the effects of long-term musical training on rhythm/meter processing by looking at behavioral and neural differences between musicians and non-musicians in auditory deviant detection. To this aim, I tested musicians as the ideal model to highlight the mechanisms behind rhythm and meter organization.

In the next sections, I illustrate some of the extant studies describing behavioral and neural differences in rhythm and meter processing between musicians and non-musicians.

##### 1.4.1 Rhythm/Meter Processing and Musical Expertise

In this section I describe some studies investigating differences between musicians and non-musicians in rhythm/meter processing. The method used was to observe changes in specific neural components typically associated with the perception of deviant stimuli inserted within rhythmic sequences of standard sounds. For example, Geiser and colleagues (Geiser et al., 2010) conducted an EEG study on musicians and non-musicians to investigate the effects of long-term musical training on the perception and representation of meter. Participants listened to meter-congruent sequences where deviant events (i.e. intensity accents) were inserted at metrically strong positions and to meter-incongruent sequences where deviants produced a large temporal change. The authors found that the MMN amplitude for deviant stimuli inserted in meter-

congruent (strong and expected position) and meter-incongruent sequences differed by a greater amount in musicians than in non-musicians. Musicians showed larger MMN for accents inserted in meter-incongruent compared to meter-congruent sequences with accents on strong and highly expected positions. This result led to the conclusion that musical training prompts a refined perception of meter as visible from different modulations of the MMN to meter congruent and incongruent deviants. However, the present conclusions have to be considered with caution since the larger MMN amplitude registered for the meter-incongruent sequence might simply echo the processing of a larger acoustic and temporal change present in this condition. In fact, here, a sound was removed from the sequence thus changing the amount of acoustic information delivered; and in addition, this produced a temporal change (shift of accent) within the sequence. Thus, rather than being related to a refined meter processing, the larger difference observed in musicians between meter congruent-incongruent conditions may be simply ascribed to their enhanced perceptual processing per se. In addition, it should be noted that a previous study by the same authors failed to observe neural differences between musicians and non-musicians on rhythm and meter processing (Geiser et al., 2009).

As mention before, it is important to consider that many studies exploring rhythm/meter processing in musicians and non-musicians used sequences of sounds that were either too complex or too simple (Potter et al., 2009; Tierney & Kraus, 2013). Indeed, because musicians are more familiar with complex rhythmical series compared to non-musicians, differences between groups could reflect familiarity with the material used rather than a genuine difference in rhythm and metric perception determined by extensive musical training.

Furthermore, behavioral and neural results reported by previous experiments using isochronous series of sounds (Brochard et al., 2003; Potter et al., 2009, Bouwer and Honing, 2015) are doubtful because listeners might segment the same sequence of equidistant sounds differently (i.e., segments with a different number of sounds). For example, Bouwer and Honing (2015) used an isochronous rhythm to investigate the influence of temporal attending and temporal prediction involved during metrical rhythm processing. Although in their design, a click track sound was superposed on each first sound of the isochronous rhythm in order to induce

a specific sequence segmentation, listeners might have still perceived and group subsequent sounds within the sequence differently.

Another line of research tested differences between musicians and non-musicians by means of direct judgments on auditory stimuli. For example, in the study of Jongsma et al. (2004; see also Jongsma et al., 2003) musicians and non-musicians listened to two bars marking a duple or triple meter. After a silent bar, a test bar with a probe beat was delivered. Participants were asked to pay attention to the metrical structure of the auditory stimulation, and to judge to what extent the timing of the probe beat fitted the experienced metrical structure (duple or triple). The researchers hypothesized that musicians would provide higher ratings for those probe beats that matched the previously heard metrical context, compared to non-musicians. In addition, musicians should show enhanced evoked neural responses (as measured by EEG) to those probe beats whose timing matched the metrical context (Jongsma et al., 2003, 2004). The results were in agreement with the predictions, and suggested that musicians hold in memory more precise metrical representations compared to non-musicians (Jongsma et al., 2003, 2004). However, the observed effects have to be considered with caution since differences between musicians and non-musicians might simply highlight a superior ability of musicians when tested on judgements of subtle differences in meter perception which pinpoints to musical training/ knowledge effects.

Therefore, overall, the literature on the behavioral and neural effects of long-term musical training on rhythm and meter processing do not describe a uniform and clear picture. In fact, it is not clear if differences between musicians and non-musicians may be generalized to tasks that involve auditory sequences (with a rhythmic/metrical structure) beyond musical competences and/or musical material. Moreover, a potentially problematic aspect of previous studies (e.g., Geiser et al., 2009; Geiser et al., 2010; Vuust et al., 2009) is related to the fact that the deviant stimulus often coincided with a rhythm/meter change. Consequently, this might have conferred relevance to the rhythm/metric aspect itself, and made its processing easier for musically experts.

Thereby, a possibility to explore rhythm/meter processing in musicians and non-musicians is to use simple auditory sequences with a clear rhythmical structure and to dissociate the temporal aspect (rhythm/meter

and non-rhythm) from the deviant stimulus, and to look at behavioral and neural changes associated to deviants' detection as consequence of a different temporal processing.

### **1.5 Cueing Effects on Attention Orienting in Time**

An aspect that seems to play a crucial role in the orienting of attention in time is the expectation of future events. This may derive from the ability to efficiently perceive the temporal relationship between events of a rhythmic sequence (rhythm perception), or from the use of symbolic cues indicating the occurrence of a specific events (like for example a green traffic light that cues a pedestrian about when crossing the street). As for rhythm perception, we could consider the intervals of time between sounds of a rhythmical sequence as a special set of cues that provide information on the occurrence of each subsequent sound. These cues boost expectancy and perceptual levels favoring the processing of target sounds coinciding with the cued moment in time. In support to this aspect, it has been demonstrated that perceptual judgements (i.e., pitches, interval lengths) are enhanced for stimuli occurring within regular rhythm and at the expected points in time (Jones et al., 2002; Large & Jones, 1999), compared to stimuli presented before or after these points. As suggested already by the Dynamic Attending Theory, temporal expectations derives from oscillations of attention in synchrony with an external rhythm (Large and Jones, 1999). Thus, stimuli coinciding with maximal picks of attention (larger oscillations amplitude) will be better processed.

From a slightly different perspective, the Predictive Coding Theory underlies the importance of a correct estimation of the intervals of time before each event, which allows an effective prediction of future events. Therefore, the expectancy toward a future stimulus resulting either from a dynamic allocation of attentional resources or from predictive abilities, improves stimulus detection.

On the other hand, an example of controlled temporal expectation is given by experimental paradigms in which symbolic cues are used to manipulate participants' expectancy. For example, knowing the exact location where a stimulus will occur enhances its encoding (Coull and Nobre, 1998). Likewise, it has been demonstrated that knowing the moment in which an event will most likely occur generates a voluntary orienting of attention in time, resulting in an optimization of that stimulus elaboration presented at the

attended moment in time (Coull & Nobre, 1998). This suggests that temporal information is used to guide selectively attention to specific points in time. This results not only in behavioral advantages as visible from faster and/or more accurate responses to target stimuli presented at the expected time, but also in changes of specific neural responses associated with temporal orienting and stimulus processing (see Nobre & Rohenkohl, 2014 for a review).

A method that has been used to investigate the human ability to orient attention in time consists in manipulating the time interval between a warning signal and a target stimulus (i.e., foreperiod, FP) (Weinbach and Henik, 2012). In 1914, Woodrow observed that the length of the FP had a strong influence on the participants' reaction times (Woodrow, 1914). In particular, a short FP (e.g., 800 ms) triggered faster RTs to a target stimulus compared to long FPs (e.g., 2000 ms, Klemmer, 1956; Naatanen et al., 1974), due to the fact that the temporal resolution over short interval of time is higher compared to long ones (Allan & Gibbon, 1991; Wearden & Lejeune, 2008). Additionally, it has been showed that holding constant and predictable the FPs from trial to trial prime a more precise temporal expectation (faster RTs) about the occurrence of the target stimulus compared to conditions in which FPs are variable and unpredictable (Woodrow, 1914).

Besides the behavioral advantages, orienting attention selectively to points in time modulates also a series of electrophysiological responses. For example, in their pioneering work, Walter and colleagues (1964) reported the presence of a physiological response known as the Contingent Negative Variation (CNV), a slow negative component originating in the motor areas of the brain and visible whenever subjects were presented with two stimuli: a warning signal followed by a target stimulus. They observed that a large negative response appeared just after the warning signal and reached maximal amplitude at the anticipated time of the target stimulus (Walter et al., 1964). It has been suggested that the CNV mirrors the ability of the nervous system to anticipate and prepare motor responses to temporally predictable stimuli. As such, it reflects temporal expectations and predictions of a future event (Tecce, 1972; Mento, 2013).

However, the effect of symbolic cues is not only captured in those neural correlates reflecting temporal prediction and anticipation of future events (i.e., pre-target activity). Some EEG studies have focused on the effects that orienting attention in time may have on target processing itself (i.e., post-target onset activity;

Friedman et al., 2001; Lange, 2009; Correa and Nobre, 2008; Griffin et al., 2002; Miniussi et al., 1999). For example, these studies described modulations of the fronto-central P3 component, which usually reaches its maximal positive peaks at about 300 ms from the onset of the target stimulus (Friedman et al., 2001; Lange, 2009; Correa and Nobre, 2008; Griffin et al., 2002; Miniussi et al., 1999). I provide a detailed description of all the aspects concerned with these neural responses in Chapter 4.

### **1.6 The present research: aims and hypotheses.**

Music processing involves many cognitive and sensory functions, which are shared with other domains. For example, attention, working-memory skills and the integration of single elements within a coherent percept based on specific rules are similarly involved in both music and language.

The effects that a long-term musical training has on brain structures and functions have been widely documented (Munte et al., 2002; see Herholz and Zatorre, 2012 for a review); nonetheless, we are still far from a clear understanding of whether these effects are visible also when using auditory stimuli and tasks that do not pinpoint to music expertise. For example, while there are studies showing enhanced evoked potentials and auditory brain stem responses in musicians to pitch changes involved in both language and music (Musacchia et al., 2007), and pointing to an overall superior ability of musicians in timing processing (i.e., rhythm), other studies do not report differences due to musical expertise (Lee and Noppeney, 2011). For instance, Lee and Noppeney (2011) found differences between musicians and non-musicians in audio-visual asynchrony detection task only when using musical stimuli and not for speech (Lee and Noppeney, 2011). In addition, other studies reported advantages in attentional abilities when auditory but not visual stimuli were used (Strait et al., 2010).

The first aim of the research presented in this thesis was to examine whether a long-term training in music could influence auditory processing when the task does not imply music knowledge. Specifically, I tested whether music expertise could improve auditory processing in contexts where rhythm/meter is not task-relevant. To test this aspect, I used a detection task of auditory deviant stimuli inserted within simple auditory temporal structures (rhythmical, non-rhythmical) that were irrelevant for the task at hand. Moreover, the

use of an integrated approach consisting of behavioral and EEG measures, allowed for a more detailed picture of the temporal dynamics of the effect of music expertise in rhythm processing. On the one hand, the use of Signal Detection Theory (SDT) measures entailed the possibility to dissociate whether music expertise influences sensitivity (as measured by  $d'$  values) or response bias (as measured by the criterion). On the other hand, the analysis of two ERP responses associated with deviance detection (MMN and MLR) provided a precise evaluation of the specific (if any) level of stimulus analysis that is modified by long-term musical training.

A secondary aim of the experiments described in the first part of the present thesis (Chapter 2 and 3) was to try to ponder two main theoretical approaches often used to account for rhythm perception and music expertise (DAT: Jones and Boltz, 1989; PC: Friston, 2005): the Dynamic Attending Theory (DAT) and the Predictive Coding (PC) theory. As mentioned in Section 1.2, DAT explains rhythm processing and its facilitation effects as a consequence of a dynamic fluctuation of attention synchronized to the perceived regularities.

In contrast, PC focuses more on the precise prediction of upcoming events for an efficient rhythm processing. While the experiments presented in Chapter 2 and 3 of thesis were not specifically designed to pit the two theories against each other, throughout the experiments presented I tried to evaluate whether the results could be accounted for by a single theory, or if they could be better accounted for by considering the interplay between attentional and predictive mechanisms.

To achieve this aim and to highlight the role of DAT and PC in rhythm processing, two experiments were conducted and are reported in Chapter 2 and Chapter 3, respectively. In Chapter 2, I describe two experiments that explored the behavioral effects of long-term musical training on rhythm and meter processing. I compared musicians and non-musicians looking at differences in the detection of deviants embedded within rhythmical and non-rhythmical auditory sequences (Experiment 1), and at different metrical positions (strong vs. weak metrical positions, Experiment 2) of the rhythm. Therefore, in Experiment 1, I examined the influence of bottom-up structuring of the auditory sequences by assessing the effect of musical expertise on the detection of deviant sounds embedded within rhythmical vs. non-rhythmical



sequences. In Experiment 2, I focused on metric perception by evaluating the role of expectancy to deviant sounds at either the strong (i.e., more salient) or weak (i.e., less salient) metrical positions of rhythm. If long-term musical training enhances rhythm and meter elaboration, I expect to observe faster response times and enhanced accuracy for deviants within rhythmic sequences in musicians compared to non-musicians in both experiments. However, if the role of musical expertise is particularly related to metric processing, differences between groups should become more visible in Experiment 2 (in which expectancy for deviants rely on the processing of both rhythm and meter), compared to Experiment 1 (in which expectancy is primarily driven by rhythm).

Chapter 3 focuses on the neural correlates of the effects described in Chapter 2. By measuring two specific EEG measures (MMN and MLR) associated to deviant detection, I aimed to investigate at which stage of deviant analysis an influence of long-term musical training, as function of a different rhythm/meter elaboration would have been visible. In line with previous findings showing enhanced auditory processing in musicians at late (MMN) and earliest (brainstem) stages of stimulus analysis (Koelsch, 1999; Musacchia et al., 2007), pointing also to musicians' augmented rhythm encoding and production (Rammsayer & Altenmüller, 2006; Chen et al., 2008; see Zatorre, Chen and Penhune, 2007 for a review), I expected that an effect of long-term musical training could be captured even before the MMN component (at the level of the MLR) when considering deviants inserted within rhythmical structures.

The second aim of the thesis project was to investigate the role of endogenous cueing in orienting attention in time, as well as its interaction with music expertise. As mentioned in Section 1.5, research in the last decades has shown that humans can orient attention in time just as well as they can do for space (Coull & Nobre, 1998; see Nobre & Rohenkohl, 2014 for a review). However, a crucial aspect that has remained largely unexplored is whether musicians orient more efficiently attention to specific points in time compared to non-musicians. In fact, although several studies (Koelsch et al., 1999; Grahn & Rowe, 2009; Zuk et al., 2014; Habibi et al., 2014 see Hannon & Trainor, 2007 for a review) underlined the advantages of musicians in many perceptual tasks, less emphasis has been put in understanding if musicians extract information from informative temporal cues more efficiently relative to non-musicians. In addition, it has remained unclear

whether the superior perceptual abilities shown by musicians relative to non-musicians are strictly related to the auditory modality (which is strongly trained in musicians), or if they could generalize to vision as well. To investigate both aspects, I conducted an experiment (described in Chapter 4), in which I used a cross-modal temporal cueing paradigm and I measured the behavioral and neural modulations derived from the perception of auditory deviant targets. Participants listened to a continuous auditory rhythmic sequence where auditory or visual cues informed them about the occurrence of a target stimulus (auditory omission) presented after short, long or mixed (either short or long) intervals of time. On the basis of the findings from studies on the Foreperiod effect (see Section 1.5), I expected both musicians and non-musicians to show changes at the behavioral and neural level for target stimuli preceded by short compared to long/mixed intervals of time, because of the higher temporal resolution over short intervals. Moreover, I hypothesized that compared to non-musicians, musicians would show a superior ability to estimate the cue-target time interval specifically when auditory cues were presented, since this is the mostly trained sensory modality for a musician.



## Chapter 2

# **Behavioral Dynamics of Rhythm and Meter Perception: The Effect of Musical Expertise in Deviance Detection**

### **2.1 Introduction**

The impulse to move feet in time while listening to a rhythmical musical piece is a common experience. This act reveals the cognitive ability to identify and synchronize to ‘pulses’ or ‘beats’ within auditory series. More generally, it speaks to our capacity to structure auditory series into perceptual units broader than the single sound (Fitch, 2013). Previous studies addressing the cognitive mechanisms of beat perception supported the notion that this cognitive ability has a phylogenetic basis shared among some animal species (e.g., Cook et al., 2013; Patel et al., 2009; Parncutt 1994), and world cultures (Wallin & Merker, 2001). Moreover, recent data suggest that the ability to detect and synchronize with regular pulses is present in newborn infants, suggesting innate aspects of rhythm perception (Honing, 2012; Winkler et al., 2009). It is more controversial whether rhythm and meter perception are modulated by ontogenetic influences, such as musical expertise. For instance, it has been proposed that long-term musical training modulates meter perception more than rhythmic perception (Geiser et al., 2009; Geiser et al., 2010).

Although rhythm and meter are inter-related concepts, they contribute separately to the perception of auditory series. Rhythm can be described as the regular alternation of different durations in an auditory series (Bouwer & Honing, 2015; Jongasma et al., 2004), whereas meter involves the attribution of different levels of perceptual prominence to each of the individual auditory event (Fitch, 2013). In addition to being induced by physical auditory cues (accents), the perceptual prominence of each auditory event can be subjectively induced according to the specific metrical structure of the sequence. Consider for instance the waltz, which

is the repeated presentation of three successive isochronous sounds. Listeners tend to impose a subjective accent on the first sound compared to the last two. This introduces a characteristic auditory grouping that assigns strong and weak beats to the series (i.e., strong, weak, weak) and favors the emergence of a metrical structure within an otherwise isochronous series (Fitch, 2013; Lerdaahl & Jackendoff, 1983; London, 2012). Hence, the sound on the first accented beat is on a metrically strong position, while sounds on the second and third beats are on metrically weak positions. Accordingly, previous theories on the processing of regular auditory series have drawn a distinction between sequential versus hierarchical perception of regular series (Martin, 1972). The first and commonly perceived level corresponds to beats processed as recurring events in time. The more complex level encompasses the hierarchical organization of beats in equal subdivisions of two (e.g., as in a march) or three (e.g., as in a waltz) elements. Therefore, in regular auditory series both the beat level (sequential subdivision) and the metrical structure (higher subdivision) can be perceived. From a cognitive point of view, meter perception provides listeners with different expectations as a function of the metrical position of the tone (Large & Snyder, 2009), and influences the ability to anticipate the underlying structure of an auditory series (Vuust et al., 2009). Accordingly, meter is considered as a high-level cognitive phenomenon that is influenced by top-down processes, such as long-term musical training, more than rhythm processing.

Many studies in the literature investigated the processing of rhythm and meter in musicians and non-musicians (Besson & Requin, 1994; Grahn & Rowe, 2009; Geiser et al., 2009; Geiser et al., 2010; Jongsma et al., 2003; Vuust et al., 2005; 2009; Thaut et al., 2014; see Rohrmeier & Koelsch, 2012 for a review).

For example, in the study of Jongsma et al. (2004; see also Jongsma et al., 2003) musicians and non-musicians listened to two bars marking a duple or triple meter. After a silent bar, a test bar with a probe beat was delivered. Participants were asked to pay attention to the metrical structure of the auditory stimulation, and to judge to what extent the timing of the probe beat fitted the experienced metrical structure (duple or triple). The researchers hypothesized that musicians would provide higher ratings for those probe beats that matched the previously heard metrical context, compared to non-musicians. In addition, musicians should

show enhanced evoked neural responses (as measured by EEG) to those probe beats whose timing matched the metrical context (Jongsma et al., 2003, 2004). The results confirmed the predictions, and suggested that musicians hold in memory more precise metrical representations compared to non-musicians (Jongsma et al., 2003, 2004).

A crucial question considered by previous studies is related to whether the influence of rhythm and meter would be present also when there is no explicit requirement to process rhythm/meter. For instance, Geiser et al. (2009) tried to separate an explicit from an implicit processing of different types of deviants inserted within auditory series. In this study, musicians and non-musicians were presented with metrically regular auditory series in which a metric or rhythmic infrequent deviant was introduced. Participants were tested in two conditions. In the attended condition, participants were asked to detect and categorize the type of deviants by reporting whether the deviance perceived was rhythmic or metric. In the unattended condition, they were asked to detect and categorize among different pitches balanced over rhythmic/metric deviants. Behavioral findings showed that musicians performed better in the detection task than non-musicians, particularly for metric deviants.

An MEG study by Vuust et al., (2009) tried to explore even more directly the implicit processing of rhythm/meter. The study investigated the effect of musical expertise on the perception of deviants (rhythmic incongruities). Here deviants were inserted within auditory series of increasing rhythmic incongruence, and their occurrence disrupted the metric expectancy of the auditory series. To direct attention away from the rhythmic incongruities, participants were required to detect a variation of the intensity of a sound (one sound tuned either up or down) inserted at the end of the auditory series. The authors observed quantitative and qualitative differences in the neural responses of expert musicians compared to non-musicians, following the perception of deviant stimuli. Specifically, musicians showed larger amplitude and earlier latency of the magnetic counterpart of the mismatch negativity (MMNm), a response typically associated to the occurrence

of deviant stimuli. The difference was interpreted as an effect of long-term musical training, which allowed musicians to build a more precise rhythmic prediction and thus resulted in an enhanced processing of incongruities within rhythmic sequences (Vuust et al., 2009).

Geiser et al. (2010) also explored the sensitivity of musicians and non-musicians to perceptual accents that were meter-congruent or meter-incongruent, while participants were not explicitly asked to perform such a discrimination (i.e., they were only asked to watch a silent movie, while the EEG signal was recorded). Nevertheless, a manipulation of the temporal structure was present in the auditory stream, resulting from the insertion of metrical deviants. The results indicated a larger difference in the MMN evoked by the meter-incongruent versus meter-congruent condition in musicians compared to non-musicians. This result was interpreted as enhanced, implicit processing of meter in musicians.

A potentially problematic aspect of previous studies (e.g., Geiser et al., 2009; Geiser et al., 2010; Vuust et al., 2009) investigating rhythm/meter effects and music expertise in implicit paradigms is that a change in the rhythm/meter was present in the auditory stream, even when such a change was not explicitly relevant (i.e., it was not the target; see Vuust et al., 2009), or when no task was required (e.g., participants were watching a movie; see Geiser et al., 2010). Thus, the presence of a rhythm/meter change in the auditory stream could have drawn the participant's attention to this feature. Additionally, in addressing potential differences between musicians and non-musicians, it cannot be excluded that musicians are more sensitive to such changes in the auditory stimulation.

For instance, Vuust et al. (2009) (see also Geiser et al., 2009, 2010) tested musicians and non-musicians by means of rhythmic and metric changes inserted within the auditory series. Thus, although the authors tried to test participants in an unattended condition by directing their attention away from the rhythm/meter changes, it cannot be excluded that rhythm/meter was not processed explicitly. This could occur more likely in those participants with previous expertise with rhythmic/meter changes, namely musicians.

In the present experiments, I tested rhythm/meter processing in musicians and non-musicians with a different approach, namely by dissociating the temporal aspects (rhythm/meter and non-rhythm) of the auditory series from the changing event itself (for the case of frequency deviance). In two experiments, I explored potential effects of long-term musical training using a task in which auditory temporal structures were not relevant to the task.

## **2.2 Experiment 1**

Experiment 1 tested whether the difference between musicians and non-musicians found previously would arise also when using simple auditory series, and in tasks that do not require an explicit processing of rhythmicality (rhythm/ meter vs. non-rhythmic). Deviants were inserted in both rhythmic (with a clear metric structure) and non-rhythmic auditory series. If long-term musical training enhances the way in which auditory rhythm is processed and organized, I expect to find differences between musicians and non-musicians in deviance detection, in particular when the rhythmic versus non-rhythmic conditions were considered.

Finally, to investigate if the effect of long-term musical training could be generalized to various types of deviant stimuli, I used frequency and temporal deviant stimuli. Frequency deviant stimuli were 1030 Hz tones that incremented the frequency of the standard sounds of the auditory series. Temporal deviant stimuli were triplets of tones equally spaced by silent intervals, which constituted a clearly distinct group of three tones both in rhythmic and non-rhythmic blocks. As opposed to frequency deviance, temporal deviance detection entailed an explicit processing of rhythm.

## **2.3 Material and Methods**

### **2.3.1 Participants**

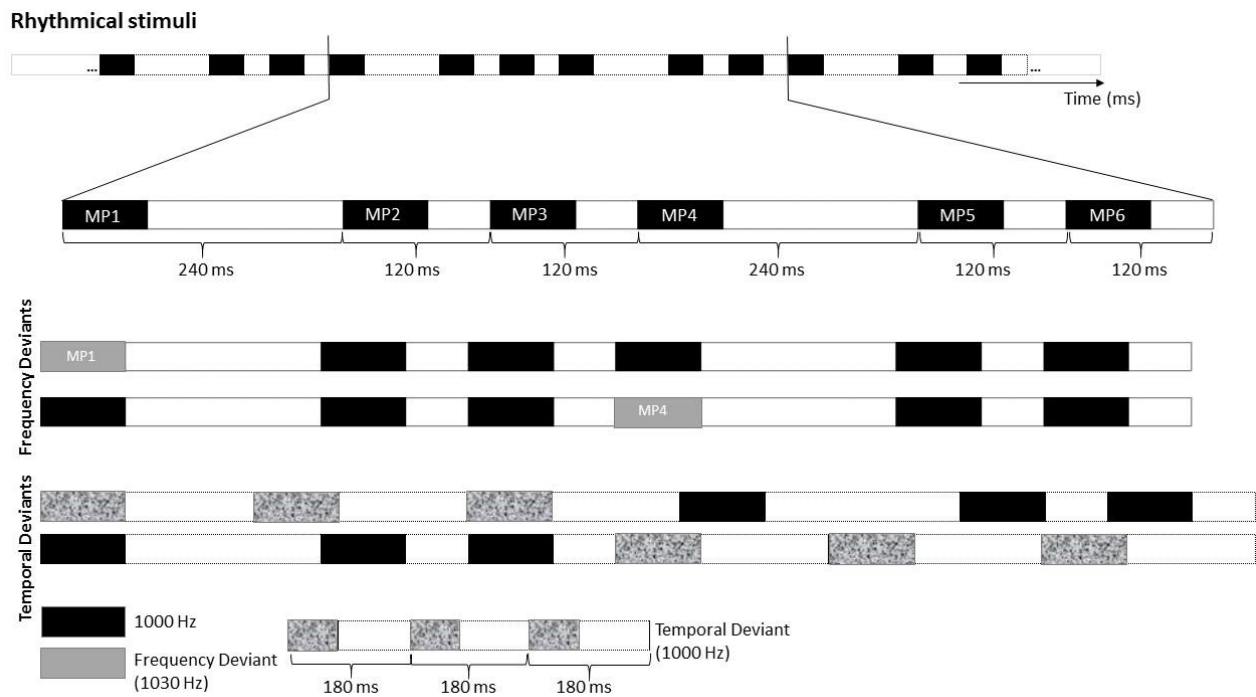
Twenty-eight healthy volunteers took part in the experiment. I tested 14 non-musicians (mean age = 21.6 years; std = 2.1) and 14 musicians (mean age = 26.3 years; std = 3.4; 6 pianists, 4 guitarists and 4 violinists, with 7.2 years of musical education, std = 1.3). In both experiments, I selected musicians who had a minimum



of six years of formal musical training. All musicians reported playing their instrument on average 5 hours per day at the time of the experiments. Non-musicians did not have any formal musical or dance education. All participants provided their written informed consent before starting the experiment. None of the participants reported hearing, neurological or psychiatric disorder, and none was taking any drug at the time of the experiment. The Human Research Ethics Committee of the University of Trento approved the experiments (protocol number 2015-011).

### **2.3.2 Stimuli**

The stimuli comprised rhythmic and non-rhythmic experimental blocks (see Procedure below), built as continuous 240 repetitions of a module comprising 6 sounds. Each sound was a 1000 Hz pure tone, lasting 70 milliseconds (ramped with 5 ms of rise time at the onset and 5 ms fall time at the offset to avoid clicks), and followed by a variable silent interval. In non-rhythmic blocks, each module was created by assembling in random order sounds with different inter-onset intervals (IOIs): 120, 140, 160, 190, 200 and 230 milliseconds (note that these IOIs had non-integer ratio lengths, thus preventing any perception of regularity). In rhythmic blocks, each module was created by assembling in fixed order sounds with the following IOIs: 240, 120, 120, 240, 120, 120 milliseconds. The basic modules of rhythmic and non-rhythmic blocks are shown in the top row of Figure 2.1 and 2.2 respectively.



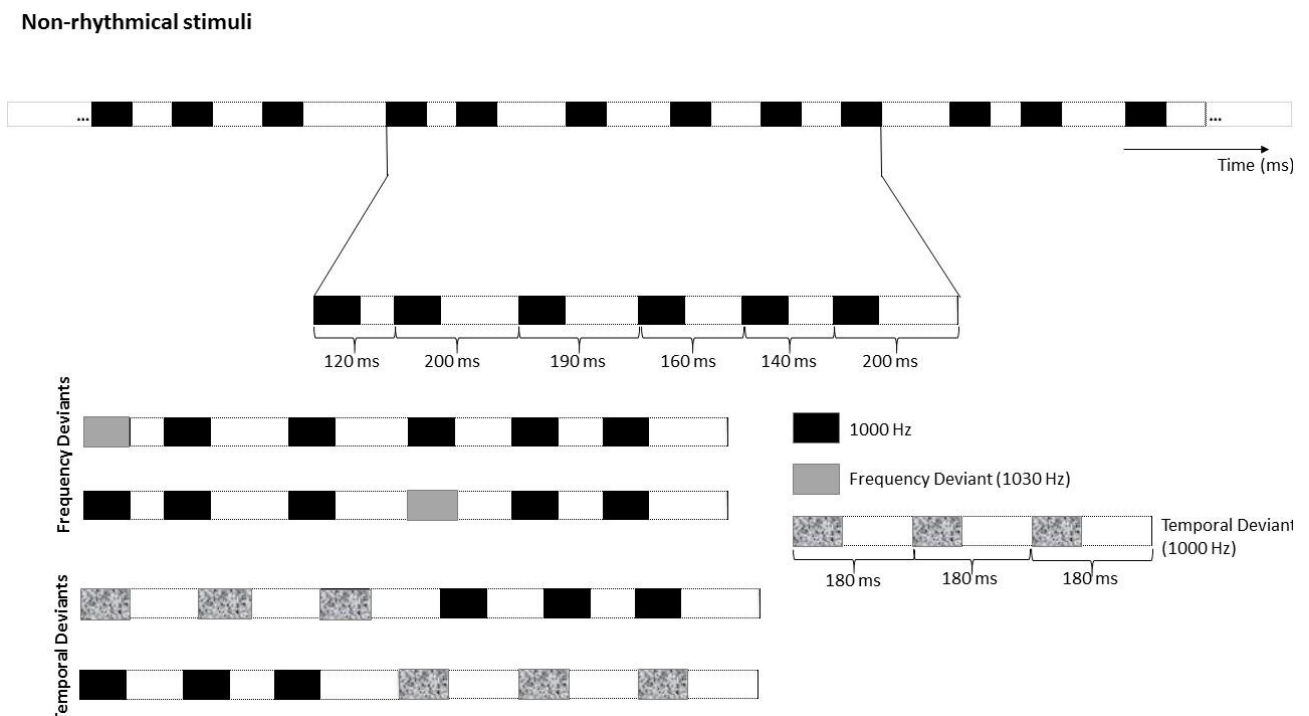
**Figure 2.1. Stimulus material.** Example of the rhythmic module. The top row illustrates an extract of the standard rhythmical module. The middle rows represent an example of rhythmical module with a frequency deviant on the first or fourth strong metrical position (MP1 and MP4). The bottom rows (last 2 rows) illustrate an example of rhythmical module with a temporal deviant replacing the first three or last three tones.

The rhythmic module (i.e., 240-120-120-240-120-120) evoked a clear binary meter based on the perception of tones' onsets. The first and the fourth sound in the module (labelled in Figure 2.1 as MP1 and MP4 - metrical positions 1 and 4, respectively) gained beat saliency due to their long IOIs. From now on, I will refer to sounds with long IOI as metrically 'strong'. By contrast, I will refer to the remaining sounds in the module (i.e., MP2, MP3, MP5 and MP6 in Figure 2.1) as metrically 'weak'.

In each rhythmic and non-rhythmic block, 20% of the modules (i.e., 48 modules out of 240) included a deviant stimulus. Across blocks, the deviant was either a frequency or a temporal change. Frequency deviants were 1030 Hz tones (i.e., 30 Hz higher than the standard tones). Temporal deviants were triplets of tones equally spaced by silent intervals of 110 ms, and constituted a clearly perceivable group both in rhythmic and non-rhythmic blocks. Modules with deviants were inserted in the block in pseudo-random order, with the constraint of a minimum of three and a maximum of seven modules without deviants (i.e., 'standard' modules) before and after each deviant occurrence. Frequency deviants occurred with equal probability on

either the first or the fourth strong metrical position in the module (see Figure 2.1). Temporal deviants replaced with equal probability the first or last three tones in each module (see Figure 2.1).

Auditory stimuli were generated using the PsychToolbox extensions (Version 3.0.12.; October 2014) running under Matlab R2014b (The MathWork, USA). All stimuli had the same acoustic energy and I did not use any musical accent to avoid any potential influence of acoustic sound features per se. Stimuli were presented via loudspeakers placed on a table, at a distance of 60 cm from the participant, at approximately 60 dB SPL. I used a computer keyboard to record participants' responses.



**Figure 2.2. Stimulus material.** Example of the non-rhythmical module. The top row illustrates an extract of the standard non-rhythmical module. The middle rows represent an example of non-rhythmical module with a frequency deviant replacing the first or fourth tone. The bottom rows (last 2 rows) illustrate an example of non-rhythmical module with a temporal deviant replacing the first three or last three tones.

### 2.3.3 Procedure

Participants were instructed to fixate a central cross presented on a computer screen in front of them, and to listen carefully to the continuous auditory stream. They pressed a key as quickly as possible, when they detected the deviant stimulus (go/no go response). Participants completed 12 blocks overall: 6 rhythmic and

6 non-rhythmic blocks. Half of the blocks contained frequency deviants, and the remaining half contained temporal deviants. Block presentation order was counter-balanced across participants.

## **2.4. Results**

Data were analyzed considering the following measures: Reaction Times (RTs) for correct responses, and sensitivity/criterion measures (as defined by Signal Detection Theory, Stanislaw & Todorov, 1999). RTs were analyzed with respect to the sound onset of the frequency deviant and to the onset of the first sound of the temporal deviant. I considered “hits” the responses given in the time window from 200 to 1500 ms post-deviant stimulus onset (the value of 200 ms is approximately the time needed for motor response execution, e.g., see Luce, 1986. The use of 1500 ms as high cut-off was justified by the relatively long duration of the temporal deviants). When no response was recorded in this time interval, the trial was classified as a “miss”. “False alarms” were analyzed considering the sounds on MP1 and MP4 in the rhythmic blocks, or the first and fourth sounds in the non-rhythmic blocks. False alarms were classified as those responses given to (standard) sounds at these positions when no deviant stimulus occurred in the 1500-ms interval before the button press (as explained above, the bars containing a deviant stimulus were preceded and followed by a minimum of 3 and a maximum of 7 standard series).

Each measure was entered separately into an Analysis of Variance (ANOVA) with Deviant Type (frequency or temporal) and Temporal Structure (rhythmic or non-rhythmic) as within-subjects factors, and Group (musicians or non-musicians) as between-subjects factor. Post-hoc analyses were conducted via pairwise comparisons (t-tests). I used Bonferroni correction for all comparisons. Figure 2.3 shows all behavioral results. Statistical analyses have been generated using the IBM SPSS Statistics 21 software.

### **2.4.1 Reaction Times**

The ANOVA indicated a significant main effect of Deviant Type [ $F(1,26) = 268.29, p < .001, \eta p^2 = .912$ ], with faster RTs for frequency compared to temporal deviants (see left section in Figure 3). The main effect of Temporal Structure was also significant [ $F(1,26) = 16.35, p < .001, \eta p^2 = .386$ ], due to faster RTs in rhythmic

than non-rhythmic blocks. Most importantly, I observed a significant main effect of Group [ $F(1,26) = 6.70$ ,  $p = .016$ ,  $\eta^2 = .205$ ], caused by faster RTs in musicians compared to non-musicians in all experimental conditions. There were no significant interactions (Deviant Type x Group:  $p = .507$ ; Temporal Structure x Group:  $p = .963$ ; Deviant Type x Temporal Structure:  $p = .098$ ; Deviant Type x Temporal Structure x Group:  $p = .113$ ).

#### **2.4.2 Sensitivity ( $d'$ )**

To assess perceptual sensitivity to auditory deviants within rhythmic and non-rhythmic blocks for the two groups of participants, I calculated  $d'$  values following the Signal Detection Theory analysis (SDT; Stanislaw & Todorov, 1999). Higher  $d'$  values represent higher sensitivity. Data were analyzed using ANOVA, with the same factors as in the other analyses.

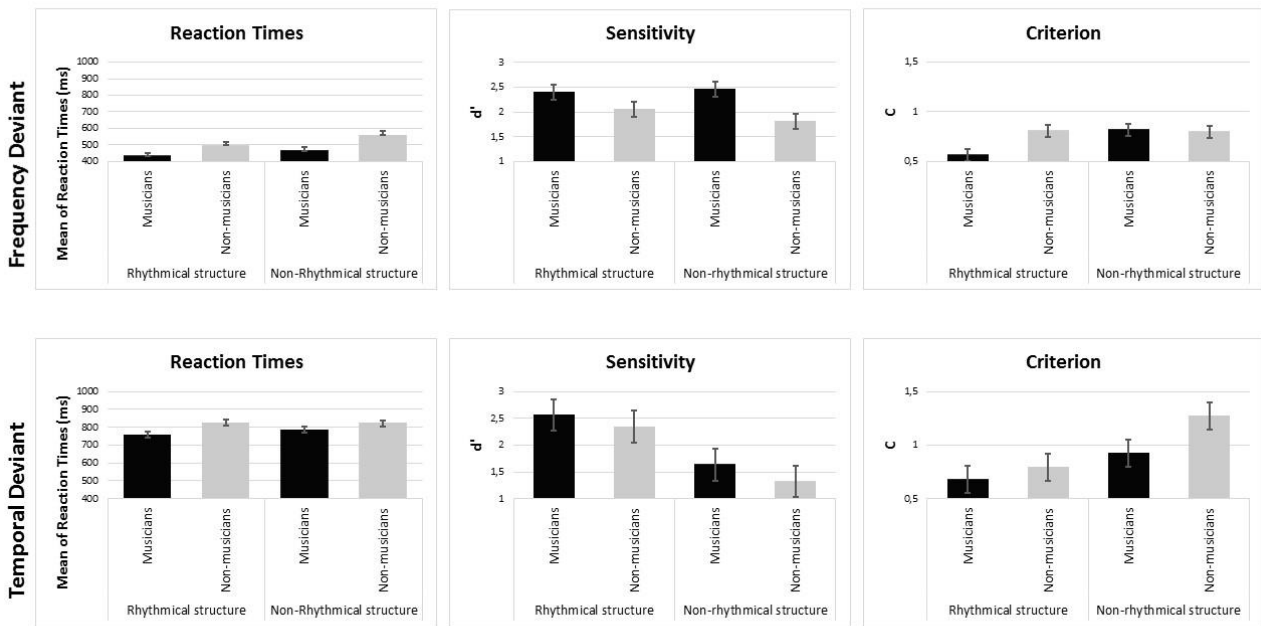
The results showed a marginally significant main effect of Deviant Type [ $F(1,26) = 3.88$ ,  $p = .059$ ,  $\eta^2 = .130$ ], with slightly better discrimination for frequency than temporal deviants, and a significant effect of Temporal Structure [ $F(1,26) = 55.99$ ,  $p < .001$ ,  $\eta^2 = .683$ ] with higher  $d'$  values in the rhythmic compared to the non-rhythmic block (see middle section in Figure 2.3). There was a significant effect of Group [ $F(1,26) = 6.02$ ,  $p = .021$ ], indicating higher sensitivity for musicians than non-musicians in all experimental conditions. The Deviant Type x Temporal Structure interaction was also significant [ $F(1,26) = 33.58$ ,  $p < .001$ ,  $\eta^2 = .564$ ]. Follow-up pairwise-comparisons (t-tests) on the Deviant Type x Temporal Structure interaction showed higher  $d'$  for temporal deviants within rhythmic blocks ( $p = .033$ ) compared to non-rhythmic ones ( $p < .001$ ). No differences were observed for the  $d'$  associated with frequency deviants in rhythmic vs. non-rhythmic blocks ( $p = .458$ ). The Deviant Type x Group ( $p = .294$ ), Temporal Structure x Group ( $p = .174$ ) and Deviant Type x Temporal Structure x Group interactions were not significant ( $p = .504$ ).

### 2.4.3 Response Criterion (c)

To examine whether the criterion adopted by the subjects was modulated by the temporal structure and musical expertise, for each subject and condition I calculated the response criterion (c) following the SDT.

Higher c values indicate the use of a more conservative criterion.

The ANOVA indicated significant effects of Deviant Type [ $F(1, 26) = 4.33, p = .047, \eta^2 = .143$ ], with higher c values for temporal than frequency deviants (see right section in Figure 2.3). The significant effect of Temporal Structure [ $F(1, 26) = 7.95, p < .001, \eta^2 = .234$ ] indicated that subjects were more conservative during non-rhythmic blocks. The ANOVA did not yield a significant effect of Group ( $p = .123$ ). The ANOVA showed statistically significant effects of Deviant Type x Temporal Structure [ $F(1,26) = 4.68, p = .040, \eta^2 = .153$ ] and Deviant Type x Temporal Structure x Group [ $F(1, 26) = 5.05, p = .033, \eta^2 = .163$ ] interactions. The other interactions were not significant (both  $ps > .464$ ). To investigate further the effects of long-term musical training on the criterion adopted for the detection of deviants within rhythmic and non-rhythmic temporal structures, I compared the response criterion of musicians and non-musicians by means of independent sample t-tests. The two groups did not show statistically significant differences in any of the conditions tested (all  $ps > .111$ ; see Note 1).



**Figure 2.3. Results of Experiment 1.** The top row shows reaction times,  $d'$  and criterion measures for frequency deviants inserted within rhythmical and non-rhythmical temporal structures, for musicians and non-musicians. The bottom row shows reaction times,  $d'$  and criterion measures for temporal deviants inserted within rhythmical and non-rhythmical temporal structures, for musicians and non-musicians. Error bars depict the standard error of the means.

## 2.5 Discussion

In Experiment 1, I explored the impact of long-term musical training on the detection of rare deviant stimuli inserted within rhythmic and non-rhythmic auditory blocks. The results indicate that musicians were overall faster and more sensitive than non-musicians, independent of the temporal structure and of the type of deviant. Additionally, rhythm facilitated deviance detection regardless of musical expertise. That is, faster RTs, enhanced sensitivity and a more liberal criterion were measured for deviants inserted within rhythmic blocks as compared to non-rhythmic ones, for both musicians and non-musicians. Finally, data showed an effect of deviant type (Note 4) for both groups, resulting in faster RTs, a marginally enhanced sensitivity and a more liberal criterion adopted for the detection of frequency deviants.

Overall, the results suggest that the large difference between the temporal structures – a rhythmic and highly regular structure compared to a non-rhythmic and irregular one- might have levelled the occurrence of a

behavioral difference between musicians and non-musicians related to the rhythmic versus non-rhythmic conditions. For this reason, I conducted an additional experiment.

## **2.6 Experiment 2**

In Experiment 2, I explored whether the effects of long-term musical training may arise during the processing of a refined metrical structure. I focused on the difference between musicians and non-musicians in detecting deviant stimuli (frequency deviant) occurring at ‘strong’ (MP1, MP4 – more salient) vs. ‘weak’ (MP2, MP3, MP5, MP6 – less salient) metrical positions. I included only the rhythmic series, since strong and weak metrical positions are a peculiar aspect of rhythmic contexts. Finally, I selected only the frequency deviant, as it covers a single sound (while temporal deviant always replace three sounds within each module, thus spanning over strong and weak metrical positions) and allows a more direct comparison between metrical positions. If the role of long-term musical training is strongly associated with meter processing, I expected to observe a different response to deviants on strong versus weak positions for musicians compared to non-musicians.

## **2.7 Material and Methods**

### **2.7.1 Participants**

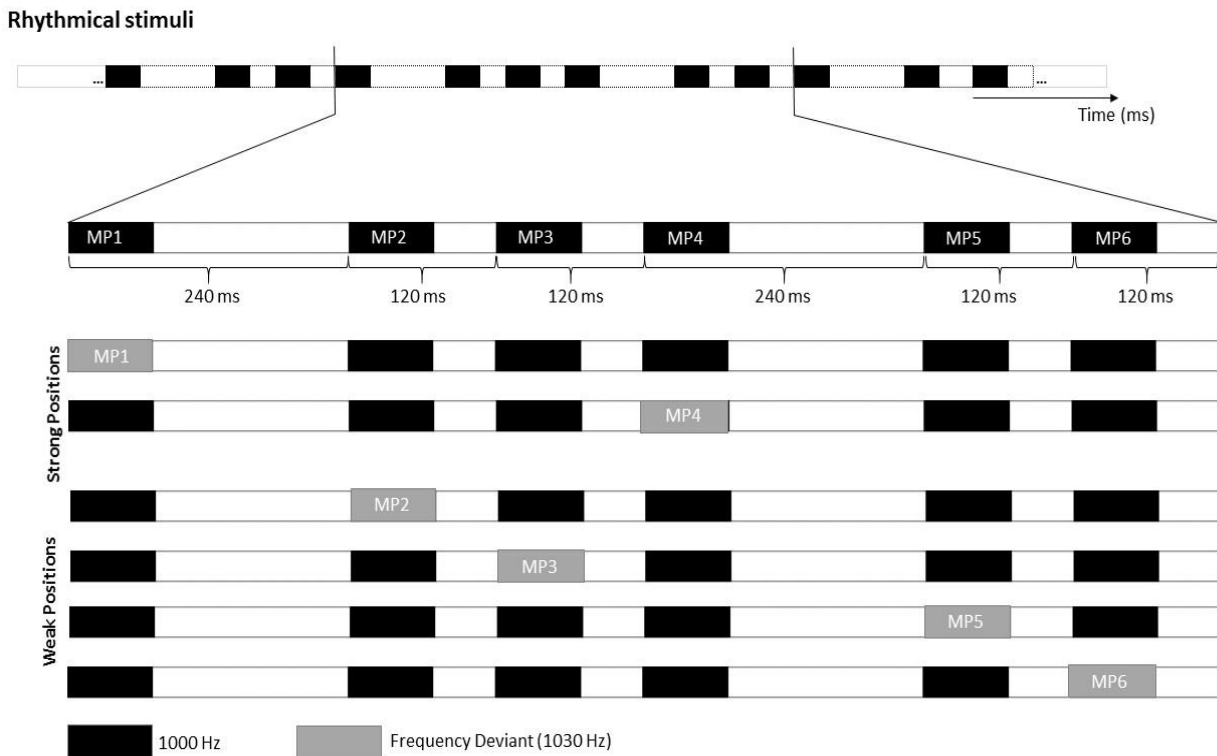
I tested thirty healthy volunteers. None of them had participated in Experiment 1. I tested 15 non-musicians (mean age = 24.3, std = 1.3) and 15 musicians (mean age = 28.1, std = 2.4; 8 pianists, 2 guitarists and 5 violinists with 8.7 years of musical education, std = 1.7). Other details are reported in the dedicated section of Experiment 1.

### **2.7.2 Stimuli**

I selected only the rhythmic block described in the stimuli section of the Experiment 1. Frequency deviants were inserted pseudo-randomly on ‘strong’ or ‘weak metrical’ positions (see Figure 3.1). Because the ratio between strong and weak positions was uneven in each module (i.e., 2 strong positions, 4 weak position), I



tested half of the participants with deviants at weak positions delivered only at MP2 and MP5, and the remaining half with deviants at weak positions delivered only at MP3 and MP6. This resulted in an equal number of deviants at strong and weak positions in each block. In total, out of 240 modules, 24 included deviants at strong position, 24 included deviants at weak positions, and 192 were ‘standard’ modules.



**Figure 3.1. Stimulus Material.** Example of the rhythmic module. The top row illustrates an extract of the standard rhythmical module. The two middle rows represent an example of rhythmical module with a frequency deviant on the first or fourth strong metrical position (MP1 and MP4). The bottom rows (last 4 rows) illustrate an example of a rhythmical module with a frequency deviant replacing the second, third, fifth or sixth weak metrical position (MP2; MP3; MP5; MP6).

Auditory stimuli were generated using the PsychToolbox extensions (Version 3.0.12.; October 2014) running under Matlab R2014b (The MathWork, USA). To guarantee that the effects found were not due to acoustic sound features per se all stimuli had the same acoustic energy and I did not use any musical accent. Stimuli were presented via loudspeakers placed on a table, at a distance of 60 cm from the participant, at approximately 60 dB SPL. I used a computer keyboard to record participants’ responses.

### **2.7.3 Procedure**

The procedure was the same as the one described for the Experiment 1. Participants completed 6 experimental blocks. Each block lasted 4.2 minutes and was followed by approximately 20 seconds of break. Block presentation was self-paced.

## **2.8 Results**

RTs for correct detections and sensitivity/criterion measures data were analyzed following the same procedure as described in Experiment 1, with the exception that here false alarms were calculated considering the sounds on MP1 and MP4 for the strong condition, or MP2/MP3 and MP5/MP6 for the weak condition. Each measure was entered separately in an ANOVA with Position (strong versus weak) as within-subjects factor, and Group (musicians versus non-musicians) as between-subjects factor. Post-hoc analyses were conducted via pairwise comparisons (t-tests). I used Bonferroni correction for all comparisons. Statistical analyses have been generated using the IBM SPSS Statistics 21 software.

### **2.8.1 Reaction Times**

The ANOVA revealed a significant effect of Position [ $F(1, 28) = 14.85, p < .001, \eta^2 = .347$ ], suggesting that participants were overall faster on strong positions (Fig. 4a). Moreover, there was a significant effect of Group [ $F(1, 28) = 10.88, p < .001, \eta^2 = .280$ ], with faster RTs in musicians compared to non-musicians in all conditions. Crucially, the Position x Group interaction was also significant [ $F(1, 28) = 4.21, p = .049, \eta^2 = .131$ ]. Paired sample tests indicated that musicians were faster in detecting deviants on strong compared to weak positions ( $t = -4.94, p < .001$ ) whereas no significant difference was found for non-musicians ( $t = -1.123, p = .280$ ). See Figure 3.2.

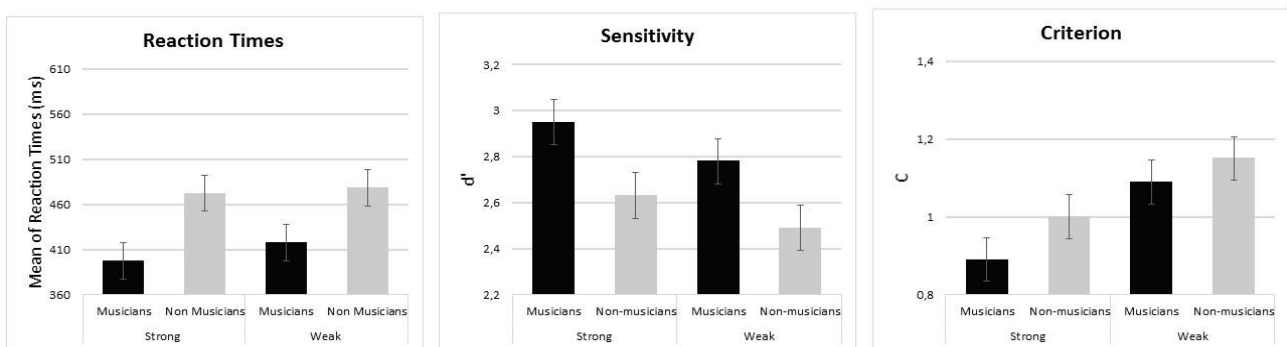
### **2.8.2 Sensitivity ( $d'$ )**

The results showed a significant effect of Position [ $F(1, 28) = 10.03, p = .004, \eta^2 = .264$ ], with higher  $d'$  values at strong versus weak positions for both musicians and non-musicians (Fig. 4b). Furthermore, we

found a significant effect of Group [ $F(1, 28) = 6.08, p = .020, \eta^2 = .179$ ], resulting in higher  $d'$  for musicians compared to non-musicians. The ANOVA did not yield a significant interaction of Position  $\times$  Group ( $p = .714$ ). See Figure 3.2.

### 2.8.3 Response Criterion (c)

The ANOVA indicated a significant effect of Position [ $F(1, 28) = 17.40, p < .001, \eta^2 = .383$ ], underlining a more conservative response criterion for the detection of deviants on weak versus strong positions. No significant effects of Group ( $p = .294$ ) and Position  $\times$  Group ( $p = .572$ ) emerged. See Figure 3.2.



**Figure 3.2 Results of Experiment 2.** Reaction times,  $d'$  measures and Criterion for musicians and non-musicians for deviants at strong and weak metrical positions. Error bars depict the standard error of the means.

## 2.9 Discussion

Musicians were overall faster than non-musicians. Crucially, the effects of long-term musical training were evident in RTs associated to the detection of deviants on strong versus weak positions, with the former identified faster than the latter by musicians. In contrast, response speed in non-musicians was not influenced by the position of the deviant stimuli. Finally, there was an overall increased sensitivity for

musicians compared to non-musicians and a more conservative response criterion adopted by both groups for the detection of deviants on weak versus strong positions.

### **3. General discussion**

In this experiment, I hypothesized that long-term musical training would influence the processing of rhythm and meter, leading to enhanced rhythm elaboration (Experiment 1) and a refined metrical representation (Experiment 2) in musicians as compared non-musicians. In contrast with this prediction, the main results of Experiment 1 did not indicate a training-related effect on rhythm processing, as both musicians and non-musicians were more efficient in the rhythmic versus non-rhythmic conditions. This result could be explained by the large difference between the rhythmic and non-rhythmic auditory series, which could have levelled out any potential effects of long-term musical training in this experiment. Indeed, when I explored the effects of long-term musical training during the processing of a refined metrical structure (Experiment 2), I found that the position of the deviant stimulus influenced response speed only in musicians, with faster responses for strong than weak positions. Thus, expectancy towards specific positions in the auditory sequence (i.e. the strong ones) improved speed in deviance detection.

Below we discuss all the results of the experiment in detail.

The results of Experiment 1 additionally pointed to three aspects.

First, in all conditions responses were faster and more sensitive for musicians than non-musicians. In line with previous research (Herholz & Zatorre, 2012; Klein et al., 2016; Munte et al., 2002) this result indicates that prolonged experience with musical stimuli leads to changes in the way auditory stimuli are processed, even when these are non-musical items. Importantly, results on the criterion and sensitivity measures indicated that the better performance in musicians compared to non-musicians was not the consequence of a different response bias, but likely reflected enhanced perceptual processing in deviance detection. Thus, as pointed out in the introduction of this thesis, the present result highlights a musical training-related effect transferred also to a non-musical context, as long as auditory and timing characteristics are involved.

Second, in line with previous studies (Bouwer et al., 2016; Schwartz et al., 2011), all participants were better at detecting deviants inserted in rhythmic blocks compared to the non-rhythmic ones. Participants were overall faster, had higher sensitivity and adopted a more liberal criterion when detecting deviants in rhythmic blocks compared to non-rhythmic ones. This effect can be framed in two theories that describe the mechanisms of rhythm perception: temporal prediction, as theorized by the Predictive Coding theory (PC); and temporal attention, described by the Dynamic Attending Theory (DAT).

According to PC theory, the perception and organization of sounds in regular units is triggered by the existence of an internal predictive model built on previous knowledge. The model is grounded on the probability of observing a certain event based on prior statistical information (Vuust & Witek, 2014), and it has been associated with the successful temporal estimation of incoming target sounds (Friston, 2005; Ono et al., 2015). According to PC theory, perception follows a Bayesian process by which the brain predicts the causal relationship between sensory percepts in a continuous comparison with previous knowledge (Friston, 2005). In this process, the brain constantly tries to minimize the prediction error between the input and the internal expectation (Vuust & Witek, 2014). To maximize the match between prediction and sensory input, a dual mechanism is hypothesized to occur: a backward mechanism, where the perceived sensory input is compared to previous knowledge, and a forward mechanism where we use previous knowledge, experience or contextual cues to anticipate and predict sensory inputs (Vuust & Witek, 2014). Beside the short-term experience with the experimental manipulations, long-term experiences (such as long-term music training, as in the present experiment) can also influence the ability to make predictions. In line with PC theory, the regularity extracted from the rhythmic series is used to predict future events (Bendixen, 2014). In particular, the regularity of the rhythmic series is considered as an anticipatory model, namely a prior that, based on different levels of salience attributed to each sound/beat, is used to create specific expectations towards future events (Vuust & Witek, 2014). As such, sound regularity should promote better performance for the highly expected and predicted events compared to those less expected and predicted (e.g., Elliott et al., 2014, but see Bouwer & Honing, 2015 for opposite results).

The Dynamic Attending Theory (DAT) also provides a description of the processes involved in the perception of metrical structures and rhythm processing. The theory argues that the listener's attention is synchronized to the regularity of the series and this promotes the creation of a temporal grid used to generate specific expectancies toward future events (Large, 2000; Jones & Boltz, 1989; Large & Jones, 1999). The mechanism of synchronization reflects the distinction between strong and weak positions of the metrical structure, with the former being more relevant and more attended than the latter. One consequence of this process is that high attentional resources on metrical strong positions result in enhanced sensitivity and thus better detection of events occurring at those positions. Previous studies on pitch and duration judgements have provided support to DAT (Jones et al., 2002; Large & Jones, 1999). For example, subjects judged pitch differences better if the sounds were embedded in auditory series with a regular inter-onset interval between tones as compared to irregular series (Jones et al., 2002). According to these studies, the regularity of the sound onset captured attention to specific time points, allowing for a precise anticipation of the incoming stimulus and hence an easier comparison of pitches occurring at these points. Therefore, regular series induced an increase in attentional levels and an anticipation of future events aligned to expected time points (Jones et al., 2002; Large & Jones, 1999). Contrariwise, the series irregularity widened attentional focus and decreased expectancy level, thus making the comparison between pitches more difficult to achieve (Jones et al., 2002).

The two theories have several aspects in common, such as the importance of item salience for both expectancy and selection processes. Indeed, both theoretical approaches converge on the explanation of the rhythmic facilitation (i.e., shorter RTs and higher sensitivity to deviant stimuli in the rhythmic compared to non-rhythmic blocks) found here. Since all deviant stimuli occurred on strong metrical positions, we suggest that the superior performance observed when deviants were in rhythmic blocks stem from the synchronization to relevant time points (i.e. the strong positions), as well as from a refined mechanism of prediction and preparation to highly expected positions (but see Bower and Honing, 2015, for opposite predictions). Hence, although these specific results did not reveal differences between musicians and non-

musicians as hypothesized in the general introduction (Chapter 1), it seems that both a dynamic allocation of attentional resources to specific time points (as proposed by DAT) and a high precision in predicting time intervals preceding future events (as proposed by PC) led to an augmented perception of stimuli coinciding with expected points in time (Ono et al., 2015; Friston, 2005), which would in turn facilitate deviance detection in this experiment.

Third, the results comparing frequency versus temporally deviant stimuli indicated that all participants detected frequency deviants more easily. I speculate that the identification of frequency deviants required an online comparison between the sound frequencies of the standard and deviant stimuli. Thus, a single odd sound was sufficient to initiate deviance detection in the case of frequency deviance. In contrast, the identification of temporal deviance could be accomplished in different ways, ranging from the detection of the first unexpected interval length within the triplet of deviant sounds to the unified recognition of the three sounds with different intervals as a group of deviant elements. Thus, the difference in time sluggishness between frequency and temporally deviant stimuli could explain the difference in detecting frequency versus temporal deviance. More specifically, in the rhythmical structure, the temporal deviant disrupted the temporal predictions built based on the previous inter-onset-intervals (IOIs) characterizing the rhythm. In contrast, the frequency deviant replaced a single tone of the rhythm, and did not modify the perceived rhythm and the temporal predictions made. In line with the PC theory, I could speculate that the detection of temporal deviants is more difficult than for frequency ones specifically because the change in the rhythm created by the different intervals of time separating each sound disrupts any efficient temporal prediction to detect the deviant.

The results additionally indicated that temporal deviants occurring in non-rhythmical blocks resulted in lower sensitivity compared to those presented in rhythmic ones. Non-rhythmic structure are highly irregular, and as such largely unexpected with respect to rhythmic ones. Thus, the difference found could be related to different degrees of unpredictability (one intrinsic to the nature of temporal deviance and one specifically related to rhythmic versus non-rhythmic contexts).

Three crucial aspects emerged from the results of Experiment 2.

First, as mentioned at the outset of the General Discussion, I found that a refined metrical structure prompted different response times as a function of musical expertise. Indeed, the results indicated that only musicians had faster responses in detecting deviants inserted at strong versus weak positions. Previous research (Jogsma et al., 2003; 2004; Geiser et al., 2010) demonstrated that musicians identify the metrical structure of rhythmic auditory series better than non-musicians. Accordingly, musicians may form different levels of expectancy as a function of beat salience, with higher temporal expectancy toward more salient beats (i.e. strong positions) as compared to less salient beats (i.e. weak positions).

Second, metrical structure influenced the sensitivity in deviance detection independent of musical expertise, as deviant stimuli led to higher sensitivity and more liberal response bias on strong than weak positions for all participants (Note 3). This indicates that some effects of metric processing are present in all participants, regardless of musical expertise. This is consonance with the hypothesis stated in the general introduction (Chapter 1) of a possible interplay between the DAT and PC supporting rhythm/meter processing. Specifically, according to DAT, this result occurs because attention synchronizes with the stimulus rhythmicity and it is maximally focused on events coinciding with most salient metrical positions (the strong beats compared to the weak ones in the present experiment) (Jones & Boltz, 1989; Large & Jones, 1999; Large, 2000). This in turn results in an enhanced processing of deviant events occurring on the most salient positions. PC could account for these results in a similar way, for the following reasons (but see Bouwer & Honing, 2015, for an opposite view). Since strong positions are salient, an increment in the sound frequency, which confers to the sound a higher salience level, is hypothetically expected more on strong positions than on weak ones (as it is in the 'subjective rhythm' phenomenon showed for isochronous sequences, see Bolton, 1894; Woodrow, 1909). Accordingly, participants in this experiment benefit from a more refined prediction (and enhanced detection) of the frequency deviant (1030 Hz, which represented an increment of the standard sound frequency, 1000 Hz), when this was presented on strong compared to weak metrical positions.



Third, in line with the results of Experiment 1, effects of long-term musical training in a non-musical context were captured in an overall faster response time and more sensitive deviant detection in musicians than non-musicians, regardless of the metrical position. As in Experiment 1, the difference could not be attributed to a different response bias in the two groups, given that no significant effect of group emerged from the analysis of the response criterion.

In conclusion, the results of the present experiment provided new evidence of a large-scale processing of rhythm, regardless of musical training (Experiment 1), and of specific effects of long-term musical training in implicit processing of meter (Experiment 2). The observation that this effect was present when using simple auditory series and in a task that did not require an explicit processing of temporal structures suggests that long-term learning experience can influence cognition beyond the specific domain of training (Boot et al., 2008; 2011; Brochard et al., 2004; Klein et al., 2016; Lee & Noppeney, 2014; Munte et al., 2002).



## Chapter 3

# The time course of auditory deviance processing: the effect of long-term musical training

### 3.1 Introduction

Rhythm refers to a subdivision of time in regular intervals characterized by periodicity. In this respect, rhythm organizes time and defines a regular distribution of events into coherent and comprehensible patterns (Berlyne, 1971). The grouping mechanism triggered by rhythm often has a hierarchical structure, with different levels of complexity, and each level of the hierarchy presents a sequential organization of events. This hierarchical structure is the metrical component of rhythm. Consider for instance the waltz, where each sound subdivides the time in equal intervals. This defines the rhythmical aspect of the sequence. At the same time, some sounds are perceived as more prominent than others (i.e., in waltz, the first sound of the triplet). The prominence of the first sound guides the creation of groups of three sounds representing the first metrical level of the sequence. By perceiving the first sound as prominent, listeners perceive a tension between strong (more prominent) and weak (less prominent) sounds (i.e., in waltz: strong, weak, weak) (Lerdahl and Jackendoff, 1983; London, 2012; Fitch, 2013). Hence, the first more prominent sound will be at a metrical strong position (the one on which most listeners will spontaneously tap their feet on while listening to a musical piece), while the other two sounds will be at metrical weak positions. Accordingly, the perception of rhythmic regularity stems from a series of expectations created in different levels of the metrical structure

(Meyer, 1956): sounds at prominent positions will tend to be highly expected and attended, relative to sounds at weak positions (see, e.g., Cooper and Meyer, 1963).

In the last decades, many studies have explored the neural and behavioral correlates of rhythm and meter processing, focusing primarily on musicians, because of their extensive training and expertise with this type of material (i.e. rhythmic/metric stimuli). Some of these studies demonstrated that long-term musical training improves the encoding and production of rhythm (Rammsayer & Altenmüller, 2006; Chen et al., 2008), and triggers a refined meter perception (Jongsma et al., 2004, 2005; Vuust et al., 2005, 2009; Geiser et al., 2010; Kung et al., 2011).

To investigate rhythm processing and musical expertise effects researchers have often exploited the modulation of a specific electroencephalographic (EEG) response associated with auditory expectancy violation: the mismatch negativity (MMN) (Näätänen et al., 1978). The MMN is typically observed in oddball paradigms, in which deviant and rare stimuli inserted within a sequence of standard and frequent stimuli violate listeners' expectancy. Deviance detection elicits a negative deflection visible between 100 and 250 ms post stimulus-onset. This deflection, known as MMN, is interpreted as the mismatch between the perception of a novel stimulus (the deviant) and the memory trace created by the preceding standard ones (Näätänen et al., 1978, 2007). The amplitude and latency of the MMN is influenced by the magnitude of the regularity violation, such that larger deviations from standard stimuli elicits larger MMN amplitude and shorter MMN latency, compared to a smaller deviations. The MMN is also strongly influenced by musical expertise. Indeed, some previous work (Koelsch et al., 1999; Tervaniemi et al., 2006; Yu et al., 2015 for a review) observed a larger MMN amplitude associated with auditory expectancy violation in musicians compared to non-musicians. Based on this result, it has been suggested that musicians show enhanced auditory stimulus processing because of their intensive training with sounds (Koelsch et al., 1999; Russeler et al., 2001).

A theory that has been used to explain the MMN is the Predictive Coding theory (PC), which associates the amplitude modulation of the MMN response to the prediction error between the expected stimulus and the

actual percept. According to PC, the brain seeks to minimize the prediction error to obtain the best prediction of the external events and it suggests that a small prediction error of a forthcoming stimulus results in smaller MMN amplitude compared to unpredictable stimuli. Recently, Vuust and colleagues (2009) have suggested that PC is an effective approach to understand music perception and, more specifically, rhythm and meter perception. A central aspect of this approach is that the MMN is modulated by different violations of metrical expectations.

According to PC, short-term experiences can promote precise predictions for future events, like when external cues are used before a specific stimulus. However, the ability to make predictions is influenced also by long-term experiences. Musicians show refined metrical representations (Jongsma et al., 2004), a more precise temporal integration window (Russeler et al., 2001), and enhanced sensitivity to detect small time differences inserted within simple sequences (Jones and Yee, 1997), which in turn could trigger enhanced predictive abilities and influence neural responses associated to rhythm/meter violations. Moreover, if musicians show a refined meter elaboration, this should be visible in a more efficient prediction of meter congruent deviants as compared to meter incongruent ones, as reflected by a reduced MMN amplitude for the first (meter congruent) as compared to the second (meter incongruent) event (Vuust et al., 2009). In this direction, previous studies exploring meter/rhythm processing in musicians and non-musicians observed a stronger neural response to highly incongruent metrical violations in musicians compared to non-musicians (Vuust et al., 2009), and a larger difference in the MMN amplitude between meter congruent and meter incongruent deviants in musicians than non-musicians (Geiser et al., 2010). For example, in the study of Vuust et al., (2009) musicians and non-musicians were compared on the perception of two metric violations: a strong metric violation and a metrically acceptable one. Although metric violations elicited an MMNm (the magnetic equivalent of the MMN, registered with MEG) response in both participants, MMNm to highly incongruent metrical deviations had a larger amplitude and earlier latency in musically trained subjects compared to non-musicians. These results are in line with the PC theory and point to a direct modulation of neural responses to meter incongruences due to musical experience, since the prediction error generated by meter deviations correlated positively with musical expertise.

However, the existing literature on the influence of long-term musical training on rhythm and meter processing at a behavioral and neural level (MMN) still presents inconsistent results (for a related review, see Grahn, 2009), and often reports contradicting conclusions (Honing et al., 2009; Ladinig et al., 2009; Geiser et al., 2009; Geiser et al., 2010; Bouwer et al., 2014).

For example, Geiser et al. 2009 measured behavioral and neural (MMN) responses to meter and rhythm changes inserted with metrically regular sequences in musicians and non-musicians. Participants were tested in attended and unattended conditions. In the attended condition, participants were explicitly asked to detect and categorize the type of changes by reporting whether they perceived a rhythmic or a metric change. In the unattended condition, they were explicitly asked to detect and categorize among different pitches balanced over rhythmic/metric changes. Musicians performed better on the detection task than non-musicians and this was particularly evident in the detection of meter changes. Nonetheless, the electrophysiological finding did not show any difference between groups in any of the conditions tested. This has been interpreted as reflecting similar neural processing of meter and rhythm changes in musicians and non-musicians.

An opposite result comes from the experiment of Geiser et al. 2010, where behavioral and EEG data were recorded from musicians and non-musicians listening to sound sequences with intensity accents at metrically congruent or incongruent positions. Behaviorally, musicians and non-musicians did not differ in the meter deviant detection task. This was interpreted as reflecting the ease of the meter manipulation, even for non-musicians. The EEG data showed a stronger neural response (MMN) to accents at metrically incongruent positions, compared to metrically congruent positions regardless of musical expertise. However, the difference in the MMN amplitude between meter congruent and incongruent accents was larger in musicians compared to non-musicians. Specifically, musicians showed a smaller MMN amplitude for congruent deviants (metrically expected intensity accents) and a larger neural response to metrically incongruent accents relative to non-musicians. This difference was explained as reflecting a refined mechanism of meter

extraction and, consequently, of accent prediction at congruent metric positions due to long-term musical training.

It is important to note that some of the pioneering work investigating the dynamics of meter/rhythm processing in musicians and non-musicians left several questions unanswered. For example, a critical aspect of some previous studies (e.g., Geiser et al., 2009; Geiser et al., 2010; Vuust et al., 2009) is that the deviant stimulus often coincided with a rhythm/meter change. One consequence of this is that rhythm/meter becomes relevant and, thus might be processed explicitly. For instance, Geiser et al., 2009 explored differences between musicians and non-musicians using rhythmic and metric changes inserted within sequences of sounds with a rhythmic structure. Although the authors tested subjects also in a condition in which meter was unattended (pitch detection task), it cannot be excluded that some participants still processed rhythm/meter explicitly, since this was directly manipulated in the auditory series. Moreover, it cannot be excluded that musicians were even more susceptible to such changes in the auditory stimulation, even when not instructed to pay attention/respond to them.

Another critical aspect to take in consideration when interpreting previous results is that neural and behavioral responses have been often explored for stimuli that differed in terms of the amount of acoustic information delivered (e.g., Bolger et al., 2013; 2014). For instance, Bolger et al., 2013 (see also Bolger et al., 2014) used series of sounds of which the first had a higher frequency compared to the following sounds. This difference might affect the registered neural response and facilitate deviant detection coinciding with the first sound simply because of the higher salience of the stimulus with higher frequency compared to the stimuli with lower frequency. Crucially, this effect might be even more evident in participants with a long-term musical training, for which neural responses to auditory stimuli are larger compared to non-musicians (Koelsch et al., 1999).

Finally, another limitation is that some of the past studies have sought to investigate rhythm/meter processing employing isochronous sequences of sounds (Schwartz et al., 2011; Bolger et al., 2014), which potentially could be organized in various ways according to the metrical subdivision that each listener decide

to superimpose on the sequence. In this case, it is difficult to obtain a rigorous control of the specific metrical organization a listener imposes to the isochronous sequence and, as a consequence, of the specific relevance that each sound in the sequence obtains as function of the specific metric organization.

Here, I aim to investigate the influence of long-term musical training on rhythm and meter processing by looking at behavioral and neural modulations associated to the detection of deviant stimuli (i.e. sounds with higher frequency compared to standard sounds) inserted within different temporal structures: rhythmical and non-rhythmical. Participants were tested in a relatively easy task that required neither explicit processing of auditory series, nor explicit categorization of deviant stimuli.

If long-term musical training enhances the way rhythm is processed and organized, I expect to observe a faster and more sensitive (higher  $d'$ ) detection of deviant stimuli inserted in the rhythmical structure as compared to the non-rhythmical ones. Moreover, given their refined meter elaboration musicians should show an improved detection of deviants occurring at strong metrical positions relative to weak ones. Indeed, since a sound with a higher frequency (frequency deviant in this experiment) is more expected/predicted at strong metrical positions than at weak ones, the prediction error associated with a correctly predicted event is smaller than the one associated with an unpredictable event (increment of sound frequency at weak positions), and this should be even magnified in musicians.

Overall, in line with previous evidence (Koelsch et al., 1999; Vuust et al., 2012; see Herholz & Zatorre, 2012 for a review), I expected larger neural responses to frequency deviant detection in musicians than non-musicians.

Moreover, the Predictive Coding Theory makes a specific hypothesis concerning the MMN amplitude as a function of stimulus predictability. Accordingly, given the design of the present experiment (where I always present an increment of sound frequency on weak and strong metrical positions), the best prediction should be associated with the increment of sound frequency (frequency deviant) at strong positions, where I expect smaller MMN amplitude than the MMN amplitude registered for frequency deviant at weak positions.



However, when considering the additional role of long-term musical training within the PC framework, the weak positions are not informative about the direction of the effect in the present experiment. Indeed, given the refined meter processing of musicians, these metrical positions may be still more expected by a musician than a non-musician, which should lead to the prediction that the MMN amplitude be smaller for musicians. However, because the deviant stimulus is a sound increment, it should be less expected at this position relative to the strong position, and this should lead to a larger MMN amplitude especially for musicians. Since I cannot estimate the relative contribution of each of these factors in modulating the MMN amplitude, it is hard to make specific predictions on the MMN amplitude due to musical expertise for deviants at weak positions. Conversely, because the deviant stimulus is a frequency increment in this experiment, it is highly expected to occur at strong positions and this is even magnified for musicians. Thus, in line with PC, musicians would display a smaller MMN amplitude compared to non-musicians because of their strong expectancy/prediction associated to an increment of sound frequency coinciding with the strong metrical positions.

As explained in the introduction of this thesis, besides the MMN component, the detection of a deviant stimulus within trains of standard ones (i.e., regularity violation) elicits also an enhancement of a set of ERP responses visible roughly from 12 to 50 ms post-deviant onset, known as Middle Latency Responses (MLRs) (Grimm et al., 2011). Thus, challenging the view that early neural responses to sounds are associated only with the processing of the physical properties of an incoming stimulus, it has been demonstrated that MLRs are larger for deviant versus standard sounds (Grimm et al., 2011; Alho et al., 2012), reflecting early evidences of mismatch between the memory traces for standard and deviant stimulus.

Here, I aimed to explore the existence of an effect of long-term musical training at these early stages of auditory stimulus processing.

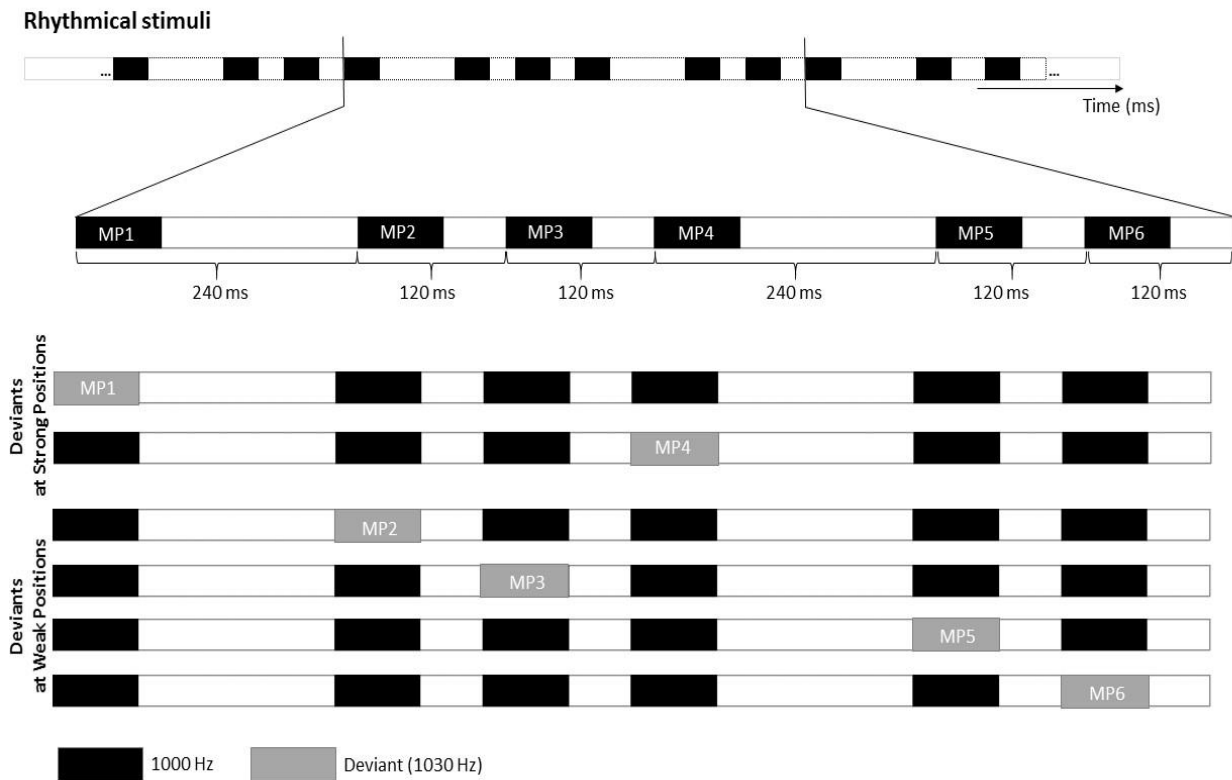
## **3.2 Material and methods**

### **3.2.1 Participants**

Twenty healthy volunteers took part in the experiment, 10 were non-musicians (mean age = 23.9 years; std = 1.5) and 10 were musicians (mean age = 27.8 years; std = 1.2). The criterion to select musicians was a minimum of six years of formal musical training. All musicians reported playing their instrument on average 5 hours per day at the time of the experiment. Non-musicians never had any formal musical or dance education. All participants provided their written informed consent prior to the experiment. None of the participants reported hearing, neurological or psychiatric disorder, and none was not taking any drug at the time of the experiment. The Human Research Ethics Committee of the University of Trento approved the experiments (protocol number 2015-011).

### **3.2.2 Stimuli**

The stimuli comprised rhythmic and non-rhythmic experimental blocks (see Procedure below), built as continuous repetitions of a module comprising 6 sounds. Each sound was a 1000 Hz pure tone, lasting 70 milliseconds (ramped with 5 ms of rise time at the onset and 5 ms fall time at the offset to avoid clicks), and followed by a variable silent interval. In rhythmic blocks, each module was created by assembling in fixed order sounds with the following inter-onset intervals (IOIs): 240, 120, 120, 240, 120, 120 milliseconds. In non-rhythmic blocks, each module was created by assembling in random order sounds with different inter-onset intervals (IOIs): 120, 140, 110, 160, 200 and 230 milliseconds (note that these IOIs had non-integer ratio lengths, thus preventing any perception of regularity). The basic modules of rhythmic and non-rhythmic blocks are shown in Figure 4.1 and 4.2 respectively.



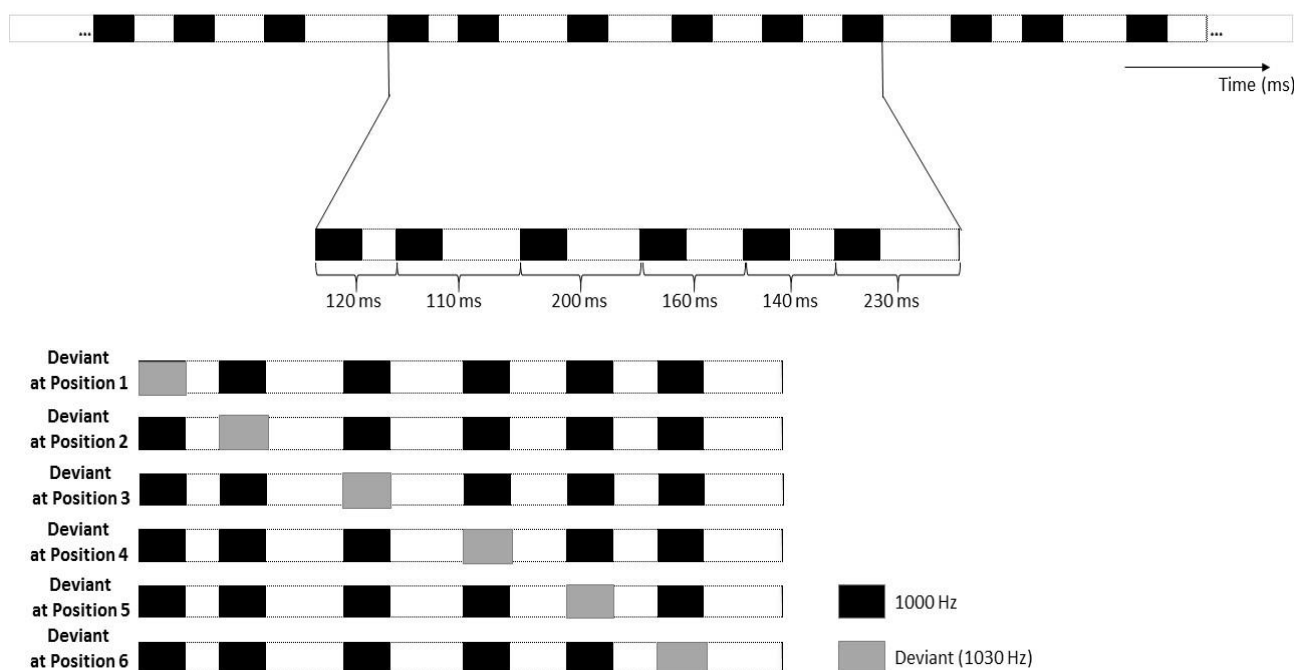
**Figure 4.1. Stimulus material.** Example of the rhythmic module. The top row illustrates an extract of the standard rhythmic module. The middle rows represent an example of rhythmic module with a frequency deviant replacing the first or fourth strong metrical position (MP1 and MP4). The last four rows represent an example of rhythmic module with a frequency deviant on the weak metrical positions (MP2, MP3, MP5 and MP6).

The rhythmic module (i.e., 240-120-120-240-120-120) evoked a clear binary meter based on the perception of tones' onsets. The first and the fourth sound in the module (labelled in Figure 4.1 as metrical positions 1 and 4, MP1 and MP4 respectively) gained beat saliency due to their long IOIs. From now on, I will refer to sounds with long IOI as metrically 'strong'. By contrast, I will refer to the remaining sounds in the module (i.e., MP2, MP3, MP5 and MP6 in Figure 4.1) as metrically 'weak'.

In each rhythmic and non-rhythmic block, 20% of the modules (i.e., 48 modules out of 240) included a deviant stimulus. The deviant stimulus was a frequency change, a 1030 Hz tones that increase of 30 Hz the frequency of the standard tones. Modules with deviants were inserted in the block in pseudo-random order, with the constraint of a minimum of three and a maximum of seven modules without deviants (i.e., 'standard' modules) before and after each deviant occurrence. Frequency deviants occurred with equal probability on either the first or the fourth strong metrical position in the module (see Figure 4.1), or on either the second,

third, fifth or sixth weak metrical position. Within the rhythmical structure, deviant stimuli occurred pseudo-randomly and with equal probability on ‘strong’ or ‘weak’ metrical position (see Figure 4.1). Because the ratio between strong and weak positions was uneven in each module (i.e., 2 strong positions, 4 weak position), I tested half of the participants with deviants at weak positions delivered only at MP2 and MP5, and the remaining half with deviants at weak positions delivered only at MP3 and MP6. This resulted in an equal number of deviants at strong and weak positions in each block.

### Non-rhythmical stimuli



**Figure 4.2. Stimulus material.** Example of the non-rhythmical module. The top rows illustrate an extract of the standard non-rhythmical module. Frequency deviants replacing either the first, second, third, fourth or fifth sound of the module are represented at the bottom of the figure.

The auditory stimuli were generated using the PsychToolbox extensions (Version 3.0.12.; October 2014) running under Matlab R2014b (The MathWork, USA) and presented binaurally through electromagnetically shielded insert earphones (ER 1, Etymotic Research, Elk Grove Village, IL, USA). To guarantee that the effects

found were not due to acoustic sound features per se all stimuli had the same acoustic energy and I did not contain any musical accent. A computer keyboard was used to record participants' responses.

### **3.2.3 Procedure**

Participants were seated comfortably in an electrically shielded cabin. They fixated a central cross presented on a computer screen in front of them and were required to listen carefully to the continuous auditory stream. They pressed a key as quickly as possible, when they detected the deviant stimulus (go/no go response). Each block lasted four minutes and was followed by approximately 20 seconds of break. Participants completed 12 blocks: 6 rhythmical and 6 non-rhythmical blocks. Blocks' presentation order was counter-balanced across participants. At the end of each EEG session, participants' musical competence was tested via the Musical Ear Test (MET) (Wallentin et al., 2010). The test consists of 208 short musical phrases of which 104 were melodic phrases played with sampled piano sounds, and 104 were rhythmical phrases played with wood block sounds. Musical phrases were presented in pair to each subject. Participants' task was to listen carefully to the musical phrases and to judge whether these were identical or not by crossing one out of two boxes on a response sheet. The different melodies in the melodic trials contained one pitch violation. The difference in the rhythmical trials was prompted by a rhythmic change. Participants did not receive any feedback during the test. The MET lasted 18 minutes and was conducted using a laptop computer. Stimuli were presented through the same earplugs used during the EEG session.

### **3.3 EEG data collection and pre-processing**

EEG data were recorded from 29 Ag/AgCl electrode sites (bandpass filter: 0.016–250 Hz) with a BrainAmp amplifier system (Brain Products GmbH, Gilching, Germany). EEG channels were placed according to the 10-20 system. The reference electrode was placed on the right mastoid, and the ground electrode was positioned on AFz. The signal was amplified and digitized using a sampling rate of 1000 Hz, and re-referenced offline to mean activity of the left and right mastoids. Impedance values were kept below 10 k $\Omega$ . The

amplified and digitized EEG was stored on a computer disc for off-line analysis. EEG pre-processing was performed using Brain Vision Analyzer.

Ocular correction ICA (Fast ICA restricted algorithm) was applied on the continuous data to remove artefacts related to eye movements. Continuous EEG signals were bandpass-filtered using a 0.1-40 Hz Butterworth zero-phase filter for the MMN component analysis (Ladinig & Honing, 2009; Tierney & Kraus, 2013), and a 15-250 Hz filter for the MLR components analysis (Baess et al., 2009; Slabu et al., 2010). Epochs starting 50 ms before and ending 450 ms after the deviant onset were taken from the continuous recordings, after baseline correction (-50 ms to 0 relative to deviant onset). Epochs with an amplitude exceeding  $\pm 100$   $\mu\text{V}$  (for MMN) and  $\pm 80$   $\mu\text{V}$  (for MLR) at any electrode in a 0-450 ms time window relative to deviant stimulus onset were excluded from further analysis (e.g., Tervaniemi et al., 2006; Ladinig et al., 2009 for the MMN and Grimm et al., 2011; Slabu et al., 2010 for the MLR). EEG epochs were averaged separately for each participant and condition (rhythmical, non-rhythmical, strong and weak position; stimulus type: standard and deviant), separately for musicians and non-musicians.

### **3.3.1 MMN**

Visual inspection of grand-averages for all conditions indicated deflections with a largest negative peak between 100 and 300 ms after the onset of the deviant stimulus, with a frontocentral localization. This is in line with the latency and scalp distribution of the MMN (Schroger, 1998). On the basis of this time range, for each participant and condition, I extracted MMN mean amplitudes at each electrode within a fronto-central region of interest (ROI) (Cz; Fz; FCz; F4; F3; C3; C4), considering a 60 ms time window around the most negative peak. I used this time window and these electrodes for the analyses of both deviant and standard stimuli responses. Peak latencies were defined as the time point of the most negative peak. I calculated difference waves subtracting the ERP obtained from the standard stimuli (except for those comprised in three bars before and after each deviant stimulus - Peter, V., McArthur, G., & Thompson, W. F., 2010) from the ERP obtained for the deviant ones. The difference waves for each participant are thought to reflect only the

additional activity elicited by the deviant stimulus, namely the MMN. Finally, for further statistical analysis I considered the MMN mean amplitude and latency averaged across the ROI.

### **3.3.2 MLR**

Mean amplitude of P0, Na, Pa, Nb and Pb components of the MLR were extracted for the two stimulus types (standard and deviant), the two temporal structures (rhythmical and non-rhythmical) and for both musicians and non-musicians. MLR mean amplitude were measured at Cz, considering a 10 ms time window (Althen et al., 2016; Grimm et al., 2011; Leung et al., 2013) around the most negative/positive peak within the following intervals: 10 - 22 ms (P0), 22 - 32 (Na), 28 - 38 ms (Pa), 38 - 48 ms (Nb), and 48 - 58 ms (Pb) (Malmierca et al., 2014; Grimm et al., 2011; Slabu et al., 2010). Because I present an explorative investigation of the MLRs, and a specific experimental paradigm designed to highlight these early components should be considered in future work, I limited the analyses at the components' peak amplitude.

## **3.4 Results**

### **3.4.1 Behavioral Results**

I analyzed the data based on the following measures: Reaction Times (RTs) for correct responses, and sensitivity/criterion measures (as defined by Signal Detection Theory, Stanislaw & Todorov, 1999). RTs were analyzed with respect to the deviants' onset. I considered "hits" each response registered in the time window from 200 to 1000 ms post-deviant stimulus onset (the value of 200 ms is approximately the time needed for motor response execution, e.g., see Luce, 1986). When no response was recorded in this time interval after the deviant, the trial was classified as a "miss" in the rhythmic blocks. "False alarms" were computed considering the sounds on MP1 and MP4 for the strong metrical positions, or MP2/MP3 and MP5/MP6 for the weak metrical positions in the rhythmic blocks. In the non-rhythmic blocks, the first and fourth sounds, the second/third and the fifth/sixth sounds were considered in the non-rhythmic blocks. A response given on these positions when no deviant stimulus was presented in the 1000-ms interval before a registered

response (as explained above, the bars containing a deviant stimulus were preceded and followed by a minimum of 3 and a maximum of 7 standard series) was considered a false alarm.

Each measure was entered separately in two Analyses of Variance (ANOVA). The first ANOVA with Temporal Structure (rhythmical or non-rhythmical) as within-subjects factor, and Group (musicians or non-musicians) as between-subjects factor. The second ANOVA with Position (strong or weak) as within-subjects factor, and Group (musicians or non-musicians) as between-subjects factor. Post-hoc analyses were conducted via pairwise comparisons (t-tests). I used Bonferroni correction for all post-hoc comparisons. It is worth noting that I performed this second ANOVA only considering the rhythmical structure since strong and weak metrical positions are present only in this condition and the comparison of metrical positions within the non-rhythmical structure was meaningless. Behavioral results are displayed in Figure 4.3. Statistical analyses have been generated using the IBM SPSS Statistics 21 software.

### **3.4.1.1 Reaction Times**

#### *Rhythmical vs. non-rhythmical blocks*

The first ANOVA indicated a significant main effect of Temporal Structure ( $F(1, 18) = 13.24, p = .002, \eta_p^2 = .424$ ), due to faster RTs in rhythmical than non-rhythmical structures. I observed a significant main effect of Group ( $F(1, 18) = 21.75, p < .001, \eta_p^2 = .547$ ), with faster RTs in musicians compared to non-musicians in all experimental conditions. The Temporal Structure x Group ( $F(1, 18) = 7.62, p = .013, \eta_p^2 = .298$ ) interaction was also significant. Follow-up pairwise comparisons (t-tests) on Temporal Structure x Group interaction, showed a significant difference in the RTs between rhythmical and non-rhythmical structures only for musicians ( $p < .001$ ). Specifically, musicians were faster in detecting deviant stimuli within rhythmical compared to non-rhythmical structures ( $t(9) = -4.036, p = .003$ ). Non-musicians did not show significant differences in the RTs associated with deviance detection between rhythmical and non-rhythmical structures ( $t(9) = -.720, p = .490$ ). Moreover, independent sample t-test displayed faster RTs for musicians than non-musicians for deviants within both rhythmical ( $t(18) = -5.124, p < .001$ ) and non-rhythmical structures ( $t(18) = -3.897, p = .001$ ).



### *Strong vs. weak positions*

The ANOVA indicated a significant main effect of Position ( $F(1, 18) = 25.56, p < .001, \eta_p^2 = .587$ ), suggesting that participants were overall faster on strong positions than weak ones. I also observed a significant effect of Group ( $F(1, 18) = 25.65, p < .001, \eta_p^2 = .588$ ), with faster RTs in musicians compared to non-musicians in all experimental conditions. The Position x Group interaction was also significant ( $F(1, 18) = 12.96, p = .002, \eta_p^2 = .419$ ). Follow-up pairwise comparisons (t-tests) on Position x Group interaction revealed a significant difference in the RTs between strong and weak positions only for the musicians group ( $t(9) = -4.967, p = .001$ ), with musician yielding shorter response speed on the strong relative to the weak metrical positions. Non-musicians showed the same RTs regardless of deviant position ( $t(9) = -1.485, p = .172$ ). Finally, musicians showed faster RTs as compared to non-musicians for both strong metrical positions ( $t(18) = -5.887, p < .001$ ), and weak ones ( $t(18) = -3.849, p = .001$ ).

### **3.4.1.2 Sensitivity ( $d'$ )**

#### *Rhythmical vs. non-rhythmical blocks*

The first ANOVA showed a significant main effect of Temporal Structure ( $F(1, 18) = 19.28, p < .001, \eta_p^2 = .517$ ), with higher  $d'$  values in the rhythmical compared to the non-rhythmical condition. In SDT, higher  $d'$  values represents higher sensitivity. There was a significant effect of Group ( $F(1, 18) = 14.46, p = .001, \eta_p^2 = .445$ ), indicating higher sensitivity for musicians than non-musicians in all experimental conditions. A marginally significant Temporal Structure x Group interaction was observed ( $F(1, 18) = 3.84, p = .066, \eta_p^2 = .176$ ). In particular, post hoc analysis on the Temporal Structure x Group interaction revealed a significant difference in the  $d'$  to deviant inserted within rhythmical as compared to non-rhythmical blocks only for musicians ( $t(9) = 4.922, p = .001$ ), with rhythmical blocks eliciting higher  $d'$  than the non-rhythmical ones. Non-musicians did not show significant differences in the  $d'$  associated with deviance detection in rhythmical vs. non-rhythmical blocks ( $t(9) = 1.591, p = .146$ ). Finally, independent sample t-test revealed higher sensitivity for musicians compared to non-musicians in both rhythmical ( $t(18) = 3.56, p = .002$ ) and non-rhythmical condition ( $t(18) = 3.84, p = .001$ ).

#### *Strong vs. weak positions*

The second ANOVA showed a significant effect of Position ( $F(1, 18) = 22.46, p < .001, \eta_p^2 = .555$ ), with higher  $d'$  level for deviants on strong versus weak positions for both musicians and non-musicians. Furthermore, I found a significant effect of Group ( $F(1, 18) = 6.04, p = .024, \eta_p^2 = .252$ ), resulting in enhanced  $d'$  for musicians compared to non-musicians. No significant interaction was reported ( $p = .230$ ).

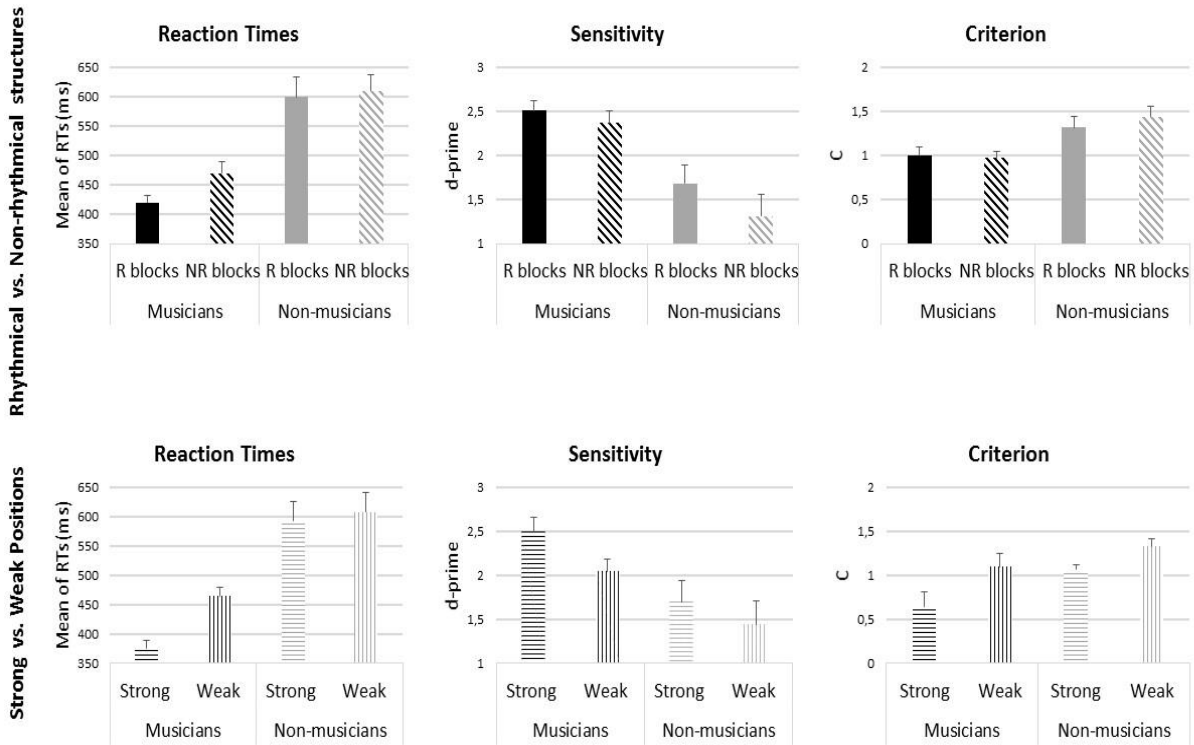
#### **3.4.1.3 Response Criterion (c)**

##### *Rhythmical vs. non-rhythmical blocks*

The first ANOVA indicated only a significant effect of Group ( $F(1, 18) = 11.18, p = .004, \eta_p^2 = .383$ ), with higher  $c$  values for non-musicians than musicians: higher  $C$  values indicate the use of a more conservative criterion. The main effect of Temporal Structure and the Temporal Structure X Group interaction were not significant (all  $p$ s  $> .439$ ).

#### *Strong vs. weak positions*

The second ANOVA indicated a significant effect of Position ( $F(1, 18) = 22.95, p < .001, \eta_p^2 = .560$ ), underlining a more conservative response criterion for the detection of deviants on weak compared to strong positions. Finally, I observed a marginally significant effect of Group ( $F(1, 18) = 4.10, p = .058, \eta_p^2 = .186$ ), due to higher  $c$  values for non-musicians than musicians. The Position X Group interaction was not significant ( $p = .214$ ).



**Figure 4.3 Results of Experiment 3.** The top row shows reaction times,  $d'$  and criterion measures for deviants within rhythmical (R) and non-rhythmical (NR) blocks, and for musicians and non-musicians. The bottom row shows reaction times,  $d'$  and criterion measures for deviants at strong or weak metrical positions of the rhythmical structure, and for musicians and non-musicians. Error bars depict the standard error of the means.

#### 3.4.1.4 Musicality Tests

The musicality test scores were analyzed by means of an ANOVA with Test (melodic or rhythmic) as within-subjects factor, and Group (musicians or non-musicians) as between-subjects factor. The ANOVA indicated only a significant effect of Group ( $F(1, 18) = 9.77, p = .006, \eta_p^2 = .352$ ), showing higher musicality (both rhythmic and melodic) scores for musicians relative to non-musicians.

#### 3.4.2 EEG results

#### 3.4.3 Mismatch Negativity (MMN)

To analyze if deviant stimuli elicited a response significantly different from zero in all conditions and for both groups, one-sample t-tests were performed on the mean amplitude of the difference waves (deviant minus

standard stimuli; 60 ms around the most negative peak within the time window from 100 to 300 ms post stimulus onset) elicited within rhythmical and non-rhythmical blocks, and at strong and weak metrical positions for musicians and non-musicians separately. The analysis yielded statistically significant results for each condition considered (for musicians, rhythmical condition:  $t(9) = -6.59, p < .001$ ; non-rhythmical condition:  $t(9) = -6.75, p < .001$ ; strong position:  $t(9) = -8.62, p < .001$ ; weak position:  $t(9) = -4.68, p = .001$ ; for non-musicians, rhythmical condition:  $t(9) = -8.27, p < .001$ ; non-rhythmical condition:  $t(9) = -7.23, p < .001$ ; strong position:  $t(9) = -7.24, p < .001$ ; weak position:  $t(9) = -5.52, p < .001$ ), showing that all deviants elicited a MMN.

To explore the effects of long-term musical training on the MMN, I analyzed separately mean amplitudes and latencies by means of two ANOVAs with the same factors used for the analysis of the behavioral data (See Figure 4.4.1).

All statistical analyses have been carried out using the IBM SPSS Statistics 21 software.

### **3.4.3.1 MMN amplitude**

#### *Rhythmical vs. non-rhythmical blocks*

The ANOVA revealed a main effect of Temporal Structure ( $F(1, 18) = 5.73, p = .028, \eta_p^2 = .242$ ), with a more negative mean amplitude for rhythmical than non-rhythmical structures. Furthermore, I observed a significant main effect of Group ( $F(1, 18) = 6.23, p = .022, \eta_p^2 = .257$ ), with more negative MMN amplitude for musicians compared to non-musicians in all experimental conditions. I did not observe significant interactions ( $p = .794$ ).

#### *Strong vs. weak positions*

The analysis revealed a trend towards significance for the main effect of Position ( $F(1, 18) = 3.042, p = .098, \eta_p^2 = .145$ ), with larger MMN amplitudes for deviant stimuli coinciding with weak metrical positions relative to strong ones. This result is in line with the Predictive Coding theory. No other effect was significant (all  $p$ s  $> .185$ )

Nonetheless, as mentioned in the Introduction, the Predictive Coding Theory also allows for a more specific hypothesis on the effects of long-term musical training associated with meter processing. In particular, given that meter expertise should lead to smaller prediction errors at highly expected positions of a metrical structure (strong metrical positions); the MMN to expected stimuli at these positions should be smaller in musicians than non-musicians. Thus, by means of independent t-tests I explored the difference between musicians and non-musicians in MMN amplitude at strong positions. In contrast with the initial hypothesis, the result highlighted a more negative MMN amplitude in musicians relative to non-musicians ( $t(18) = -2.13$ ,  $p = .047$ ).

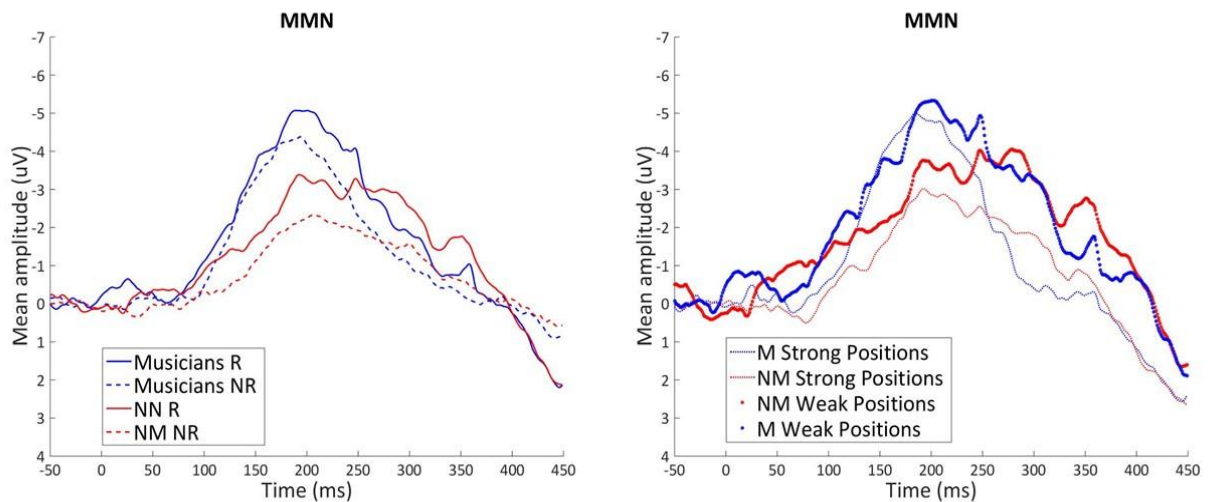
### 3.4.3.2 MMN latency

Rhythmical vs. non-rhythmical blocks

The ANOVA did not show any significant main effect or interaction (Temporal Structure:  $F(1, 18) = .108$ ,  $p = .746$ ,  $\eta_p^2 = .006$ ; Temporal Structure X Group:  $F(1, 18) = .165$ ,  $p = .295$ ,  $\eta_p^2 = .061$ ; Group:  $F(1, 18) = .096$ ,  $p = .760$ ,  $\eta_p^2 = .005$ ).

Strong vs. weak positions

The analysis resulted in a marginally significant interaction Position X Group ( $F(1, 18) = 3.47$ ,  $p = .079$ ,  $\eta_p^2 = .162$ ). No other significant effects were found (all  $ps > .531$ ). Post-hoc analysis on the Position X Group interaction did not show significant effects (all  $ps > .1$ ).



**Figure 4.4.1.** Grand averages of the MMN recorded from 10 musicians (blue line) and 10 non-musicians (red line) are shown for a fronto-central region of interest (ROI) (Cz; Fz; FCz; F4; F3; C3; C4), considering a 60 ms time window around the most negative peak. MMN for rhythmical (solid line) and non-rhythmical (dashed line) conditions are depicted on the left. The figure on the right illustrates MMN reposes elicited by deviants at strong (dotted line) and weak (bold type points) metrical positions of the rhythmical structure.

### 3.4.4 Analyses on deviant and standard stimuli responses (100-300 ms interval)

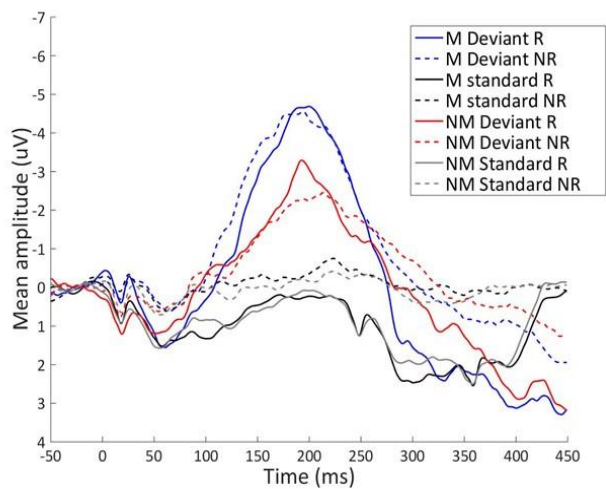
#### 3.4.4.1 Mean amplitude

In addition to the analysis of the MMN amplitude and latency, I analyzed the ERPs amplitude and latency elicited by deviant and standard stimuli separately for all experimental conditions. All standard stimuli except for those comprised in three bars before and after each deviant stimulus were considered for the analysis (Peter, V., McArthur, G., & Thompson, W. F., 2010). This analysis allows for the evaluation of the distinct contribution of standard and deviant stimuli to the generation of the MMN. Furthermore, I wanted to explore if the difference between musicians and non-musicians was also evident from a different elaboration of standard stimuli, or if instead it was only linked to the perception of deviant stimuli. Analyses were performed as explained in the sections above, with the additional factor Stimulus Type (deviant or standard).

### *Rhythmical vs. non-rhythmical blocks*

The ANOVA revealed a main effect of Stimulus Type ( $F(1, 18) = 83.24, p < .001, \eta_p^2 = .822$ ), with more negative mean amplitudes for deviant than standard stimuli in both groups of subjects. The main effect of Group was also significant ( $F(1, 18) = 8.12, p < .011, \eta_p^2 = .311$ ), due to larger ERP amplitudes in musicians compared to non-musicians. The interactions Stimulus Type X Group ( $F(1, 18) = 6.22, p = .023, \eta_p^2 = .257$ ) and Stimulus Type X Temporal Structure ( $F(1, 18) = 11.22, p = .004, \eta_p^2 = .384$ ) were also significant. Follow-up pairwise comparisons (t-tests) on the Stimulus Type X Group interaction showed larger ERP amplitudes for deviant stimuli in musicians than non-musicians ( $t(18) = -2.69, p = .015$ ). No significant differences between groups were found for the standard stimuli ( $t(18) = -.818, p = .424$ ). Finally, post-hoc analysis (t-tests) on the Stimulus Type X Temporal Structure interaction revealed a difference between rhythmical and non-rhythmical structures for the ERP amplitude elicited by standard stimuli. Specifically, standard stimuli within non-rhythmical structures yielded a more negative ERPs amplitude compared to standard stimuli building the rhythmical structures ( $t(9) = 3.59, p = .006$ ). No significant effects were observed for deviants' ERPs amplitude comparing the two temporal structures ( $t(9) = -1.19, p = .262$ ).

These results are shown in Fig. 4.4.2.



**Figure 4.4.2.** Grand-averaged ERPs elicited by deviant and standard stimuli are shown for 10 musicians (blue line for deviant; black line for standard) and 10 non-musicians (red line for deviant; grey line for standard) for a fronto-central region of interest (ROI) (Cz; Fz; FCz; F4; F3; C3; C4), considering a 60 ms time window around the most negative peak. The ERPs for rhythmical (solid line) and non-rhythmical (dashed line) conditions are depicted for both groups.

#### *Strong vs. weak positions*

Finally, ERPs mean amplitude elicited by deviants on strong and weak metrical positions were analyzed in an ANOVA with Stimulus Type (deviant or standard) and Position (strong or weak) as within-subjects factor, and Group (musicians or non-musicians) as between-subjects factor. The ANOVA revealed a main effect of Stimulus Type ( $F(1, 18) = 60.178, p < .001, \eta_p^2 = .770$ ), with more negative ERP mean amplitudes for deviant than standard stimuli in both groups of subjects. The analysis yielded a trend towards significance for the main effect of Position ( $F(1, 18) = 3.123, p = .094, \eta_p^2 = .148$ ), due to larger ERP amplitudes for the weak compared to strong positions. The main effect of Group was also marginally significant ( $F(1, 18) = 3.747, p = .069, \eta_p^2 = .172$ ), due to larger ERP amplitudes in musicians compared to non-musicians. No other effects were observed (all  $p$ s  $> .109$ ).



As in the analysis of the MMN response, by means of independent t-tests I explored the difference between musicians and non-musicians in ERP responses amplitude at strong positions. In line with the MMN t-test results and in contrast with the initial hypotheses, the analysis revealed a significant difference between musicians and non-musicians, with musicians showing a more negative ERP amplitude compared to non-musicians ( $t(18) = -2.740, p = .013$ ).

#### **3.4.4.2 Mean latency**

*Rhythmical vs. non-rhythmical blocks*

No significant effect emerged (all  $p$ s > .104).

*Strong vs. weak positions*

The analysis resulted in a marginally significant interaction Position X Group ( $F(1, 18) = 3.47, p = .079, \eta_p^2 = .162$ ). No other significant effects were found (all  $p$ s > .531). Post-hoc analysis on the Position X Group interaction did not reveal significant effects (all  $p$ s > .1).

#### **3.4.5 Middle Latency Responses (MLR)**

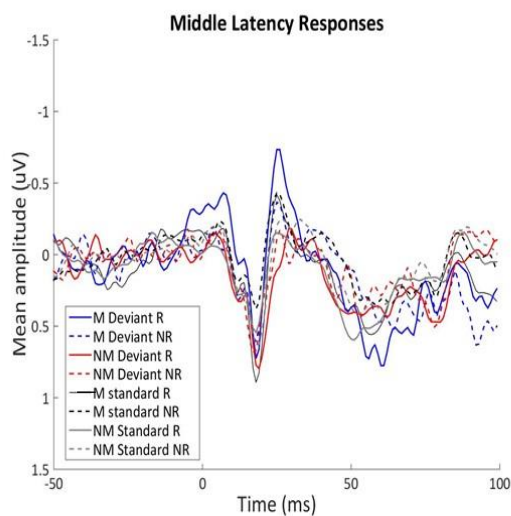
For each MLR component (Na, Nb, P0, Pa, Pb), mean amplitudes were analyzed separately following the same procedure described for the MMN data. The analyses revealed statistically significant effects only for two components of the MLR: the Na and P0. Therefore, the results are reported for these two components only.

##### **3.4.5.1 Na amplitude**

*Rhythmical vs. non-rhythmical blocks*

The ANOVA revealed statistically significant Stimulus Type X Group ( $F(1, 18) = 5.35, p = .03, \eta_p^2 = .229$ ) and Stimulus Type X Temporal Structure X Group ( $F(1, 18) = 5.72, p = .02, \eta_p^2 = .241$ ) interactions. There were no other significant results (all  $p$ s > .1). Post-hoc analyses (ANOVAs) on the Stimulus Type X Temporal Structure X Group interaction were conducted to investigate further the effects of long-term musical training on the

Na amplitude elicited by stimuli within rhythmical and non-rhythmical temporal structures. I performed two ANOVAs with Temporal Structure (rhythmical, non-rhythmical) and Group (musicians, non-musicians), separately for stimulus type (deviant, standard). The ANOVA on deviant stimuli resulted in a significant Group X Temporal Structure ( $F(1, 18) = 5.93, p = .025, \eta_p^2 = .248$ ) interaction (Figure 4.4.3). Main effects of Temporal structure ( $F(1, 18) = 1.85, p = .190, \eta_p^2 = .093$ ) and Group ( $F(1, 18) = 2.33, p = .144, \eta_p^2 = .115$ ) were not significant. Follow-up pairwise comparisons (t-tests) on the Group X Temporal Structure interaction revealed a significant difference between musicians and non-musicians only for the rhythmical structure. Specifically, I observed more negative Na amplitude for deviants within rhythmical structures in musicians ( $t(18) = -2.263, p = .036$ ) compared to non-musicians. No difference between musicians and non-musicians were found in the analysis of the Na amplitude elicited by deviants within non-rhythmical structures ( $t(18) = -.255, p = .802$ ). The ANOVA on standard stimuli did not yield significant results (all  $p$ s > .546).



**Figure 4.4.3.** Grand-average MLRs recorded at Cz (10 ms time window around the most negative/positive peak) from 10 musicians (blue line for deviant; black line for standard) and 10 non-musicians (red line for deviant; grey line for standard) are shown. The MLRs to rhythmical (solid line) and non-rhythmical (dashed line) conditions are depicted for both groups. A statistical difference for the Na component peaking at about 22 ms was observed for deviants within rhythmical structures for musicians vs. non-musicians.

### 3.4.5.2 P0 amplitude

The ANOVA on the P0 amplitude did not yield any significant results (all  $p$ s > .159).

#### *Strong vs. weak positions*

Since I did not have specific hypothesis on the effects of metrical positions on either the Na nor the P0 components, I did not analyse this further.

### 3.4.5.3 Control Analysis

The visual inspection of the Grand-averaged waveforms (see Figure 4.4.3) revealed that the neural responses registered at earlier latencies were particularly noisy, especially when considering the responses before the deviant onset in the rhythmical sequences for musicians. Thus, to test for the presence of statistically significant differences between musicians and non-musicians before deviant onset I analyzed the signal mean amplitude in the time windows from -10 ms to 0 ms and from 0 ms to 10 ms after deviant onset by means of two ANOVAs with Temporal Structure (rhythmic, non-rhythmic) as within subject factor, and Group (musicians, non-musicians) as between subjects factor.

The ANOVAs did not show significant effects (all  $p$ s > .149) except for a marginally significant effect of Group ( $F(1, 18) = 3.617$ ,  $p = .073$ ,  $\eta_p^2 = .167$ ) found in the interval from 0 to 10 ms after deviant onset showing a more negative response amplitude for musicians than non-musicians.

## 3.5 Discussion

The aim of the present experiment was: (1) to investigate the effects of long-term musical training on deviance detection in the context of rhythmic, metric (strong or weak metrical positions), and non-rhythmic auditory sequences; and (2) to explore the effects of long-term musical training at early stages of deviance processing. I tested musicians and non-musicians using non-musical stimuli and a task that did not entail explicit processing of auditory sequences, or explicit categorization of different types of deviants.

### **3.5.1 Behavioral results**

In all conditions, musicians were faster and more sensitive in the detection of deviants than non-musicians. They also adopted a more liberal criterion in the detection of deviants than non-musicians. This is in line with previous studies (Koelsch et al., 1999; Tervaniemi et al., 2001; Vuust et al., 2012b; see Herholz & Zatorre, 2012 for a review) highlighting an enhanced processing of auditory stimuli in musicians as compared to non-musicians. Crucially, the influence of long-term musical training was specifically associated with the rhythmical condition. Indeed, musicians were faster at detecting deviants inserted within rhythmical versus non-rhythmical structures, whereas non-musicians did not show such an advantage. As hypothesized in the general introduction, this further highlighted the influence of a long-term musical training outside the musical domain to the extent that the stimuli used are auditory.

Moreover, in line with previous findings (Jogsma et al., 2003; 2004; Geiser et al., 2010) only musicians' response times differentiate between deviant at strong versus weak positions, with the first eliciting faster response time than the latter. This supports the existence of a different hierarchy associated to the metrical structure of rhythm, in which stimuli coinciding with salient beats (i.e. strong positions) trigger higher temporal expectancy and/or more efficient attentional processing levels than less salient beats (i.e., weak positions).

### **3.5.2 Electrophysiological results**

Three results emerged from the analysis of the MMN.

First, the MMN was larger in both musicians and non-musicians for deviant stimuli within rhythmical structures compared to non-rhythmical ones. The additional analysis separately for deviant and standard stimuli revealed that the difference between rhythmic and non-rhythmic conditions was mainly driven by the response to standard stimuli. Indeed, the standard stimuli building the rhythmical structure resulted in a less negative neural response than the neural response elicited by the standard stimuli within the non-rhythmical structure. Since stimulus rhythmicity created a regular context, it likely prompted a better temporal expectancy toward the occurrence of stimulus onset for standard stimuli in the rhythmical context.

Therefore, the enhancement in MMN for the rhythmical condition was the result of a higher degree of adaptation in the response to the standard stimuli within a rhythmical structure, rather than being the effect of an augmented processing of the deviant stimuli in this condition.

Second, the MMN amplitude was also modulated by the metrical structure: I observed a smaller MMN amplitude for deviants (frequency deviants) at strong than weak positions. As stated in the introduction, this result is in line with the hypothesis of the PC theory. Indeed, it could be interpreted as reflecting a reduced prediction error associated to the fact that an increment of sound frequency (which were the deviant stimuli in the present experiment) is more expected at strong positions as compared to weak ones. In fact, as it is in the 'subjective rhythm' phenomenon showed for isochronous sequences (see Bolton, 1894; Woodrow, 1909), a sound with a higher frequency is expected to mark the strong positions of a metrical structure. Moreover, because I used the same deviant stimulus across all metrical positions, the observed MMN modulations for deviant at strong vs. weak positions did not derive from the perception of different auditory frequency and confirm the extraction of a specific metrical structure in the brain. This further corroborates the fast response speed associated to deviants on strong positions by both groups. Indeed, as stated in the introduction (section 3.1), earlier work often compared neural and behavioral responses associated to acoustically varying stimuli (e.g., Bolger et al., 2013; 2014). For instance, Bolger et al., 2013 (see also Bolger et al., 2014) used series of sounds of which the first had a higher frequency compared to the subsequent sounds. Crucially, this difference might affect the neural response registered and facilitate behavioral detection simply because of the higher salience of the stimulus with higher frequency compared to the stimuli with lower frequency.

Third, despite the larger MMN amplitude observed for musicians than non-musicians in all conditions, which is, as pointed out before, in line with previous work showing that prolonged experience with musical stimuli leads to an increased auditory representation for tones of a musical scale, enhanced sensitivity to the timbre of the instrument on which subjects were trained, and increased sensitivity to melodic aspect of musical phrases (Pantev et al., 1998; 2001; Fujioka et al., 2004), the specific comparison between musicians and non-musicians for the MMN amplitude for deviant at strong position revealed a larger response for the first group

relative to the latter. According to the assumptions of the PC theory, a less surprising event (i.e., an increment of sound frequency) at the most expected position (i.e., the strong one) would result in a reduced prediction error, and thus smaller MMN amplitude for musicians compared to non-musicians. This prediction is based on the assumption that an increment of sound frequency (i.e., frequency deviant in this experiment) is generally more expected to mark strong metrical positions than weak ones. In contrast with this hypothesis, I found that the MMN of musicians had a larger amplitude than that of non-musicians. This result clearly suggests a prevailing effect of stimulus salience (deviant stimuli occurring at strong metrical positions) in boosting neural responses for musicians than non-musicians. This was also in line with the shorter RTs registered to deviants on strong positions for musicians than non-musicians. Therefore, the current results lend support to the core aspect of the Dynamic Attending Theory, which posits that high attentional levels are allocated to events coinciding with metrically strong positions of a rhythm. The data further indicate that long-term musical training influences this ability, suggesting a close relationship between music expertise and the neural dynamics supporting the processing of temporally salient events.

Finally, the explorative investigation of the effects of long-term musical training on the earliest cortical responses (MLR) to deviance detection showed a stronger effect of rhythm for musicians as compared to non-musicians. Indeed, deviants within rhythmical structures elicited a larger Na amplitude in musicians relative to non-musicians. In addition to findings showing a modulation of the MMN amplitude due to musical expertise (Koelsch, 1999), this result suggests that an influence of long-term musical training is already detectable at early stages of stimulus processing (at around 22 ms post-stimulus) specifically when considering deviants within rhythmical structure. This is in line with previous research describing improved rhythm encoding and production in musicians (Rammsayer & Altenmüller, 2006; Chen et al., 2008; see Zatorre, Chen and Penhune, 2007 for a review), and it further highlights that this enhanced rhythm processing even modulates earliest cortical responses to a stimulus inserted within a rhythmical structure. Contrariwise, as I have described above, effects of rhythm appeared only at a later stage (approximately, 200 ms post-

stimulus) for non-musicians. To my knowledge, this is the first experiment to report a modulation of the MLR due to music expertise. Therefore, future work will need to expand and replicate these findings.

### **3.6 Conclusion**

These experiments demonstrated a general effect of rhythm, which modulates behavioral and neural responses associated to deviant detection in both musicians and non-musicians. This points to the ability of the auditory system to (1) organize events in time and to (2) use such regularity to facilitate many aspects of our everyday activities, like for example speech comprehension (Friedrich et al., 2004; Goswami et al., 2003). Contrariwise, the fact that both groups showed meter-related effects only at later stages of stimulus processing (i.e., in MMN range) indicates that, as compared to the basic rhythm elaboration, the processing of a more complex structure comprising different hierarchical levels of stimulus organization required more complex operations.

Finally, this experiment demonstrated finer rhythmic and meter processing in musicians than non-musicians, as it was evident both from behavioral and modulations of early and late latency neural responses. While this data cannot disentangle the question of nature or nurture on temporal perception, I highlighted a difference between rhythm and meter processing showing a clear proficiency of musicians, which was preserved also in non-musical auditory contexts and in tasks that did not require an explicit monitoring of temporal structures.





## Chapter 4

# **Orienting Attention in Time: The Effect of Cue Modality and Musical Expertise**

### **4.1 Introduction**

Musical training is a rich multisensory experience related to the integration over time of auditory, visual, and tactile stimuli with motor responses. The complexity of this experience rests on many skills associated with certain executive functions (e.g., sustained attention, goal-directed behavior, task-switching, working memory, processing speed) (Zuk et al., 2014; see Hannon & Trainor, 2007 for a review). As a consequence, long-term musical training entails refined auditory perception (Koelsch et al., 1999), enhanced rhythm processing (Grahn & Rowe, 2009; Habibi et al., 2014), augmented temporal discrimination and attentional abilities (see Hannon & Trainor, 2007 for a review), as visible from a better ability to judge/reproduce intervals' lengths (Aagten-Murphy et al., 2013), a narrower temporal integration window for judging audio/visual musical and non-musical stimulus asynchronies (Lee and Noppeney, 2011). An example of the enhanced rhythm processing of musicians derives from the study of Habibi and colleagues (Habibi et al., 2014), where behavioral and EEG data were recorded from musicians and non-musicians to unexpected rhythmic deviations inserted within unfamiliar melodies. Musicians showed to better perceive rhythm deviations than non-musicians. Furthermore, the neural responses (N100 and P200) associated to these deviations had a larger amplitude in musicians compared to non-musicians. These results were interpreted as a clear signature of the augmented perceptual ability of musicians in detecting rhythmic changes.

The extent to which the advocated advantages generalizes to non-auditory sensory domains is a matter of debate. Some studies (Helmbold et al., 2005; Patston et al., 2007) report the effects of long-term musical training also with stimuli other than auditory or musical ones. For example, musicians show better visuo-spatial and attention skills than non-musicians when tested in the detection of visual letters inserted among digits (Helmbold et al., 2005), and visual line bisection tasks (Patston et al., 2007). However, other studies did not find a generalized effect of musical training. For instance, Lee and Noppeney (2011) found differences between musicians and non-musicians in audio-visual asynchrony detection task only when using musical stimuli and not for speech (Lee and Noppeney, 2011). In addition, other studies reported advantages in attentional abilities when auditory but not visual stimuli were used (Strait et al., 2010).

Another line of research has addressed the impact of musical training on cross-modal processing. The accumulating evidence suggests a stronger cross-modal processing in musicians compared to non-musicians (Kussner et al., 2014; Vassena et al., 2016), because of their training on the integration of multimodal stimuli. For instance, a recent study by Vassena et al. (2016) tested the ability to predict future stimuli in musicians and non-musicians using a cross-modal attentional cueing paradigm. Target stimuli (auditory or visual) were preceded by compatible (cue and stimulus had the same pitch/location) or incompatible (cue and stimulus had different pitch/location) auditory, visual or cross-modal cues. Compatible conditions occurred 80% of the total and represented the predictable condition. The authors observed an increased sensitivity for the compatible (predictable) compared to the incompatible conditions only in musicians. Interestingly, the advantage was present not only unimodally, but also across modalities (auditory, visual). The result was interpreted as a clear hallmark of an enhanced cross-modal perception in musicians, which highlighted the influence of long-term musical training on predictive abilities beyond the auditory modality.

However, the paradigm adopted in the study of Vassena et al., 2016 does not fully investigate subjects' predictive abilities in time as orienting of attention in time, where subjects extract information from valid cues to efficiently predict and optimize stimulus elaboration. In fact, in their study, indicating if a target stimulus is compatible (target stimulus and cue with the same pitch or correct spatial-pitch association) with

a previously presented cue does not pinpoint to the ability to extract information from a cue to orient attention in the time domain. Hence, it could be argued that this task entails predictions on stimulus identity/location rather than temporal orienting/ predictions.

We continuously extract sensory information from the environment to optimize behavior (Coull & Nobre, 1998). For example, knowing the exact location where a stimulus will occur enhances its encoding (Coull and Nobre, 1998). Likewise, it has been demonstrated that knowing the moment in which an event will most likely occur generates a voluntary orienting of attention in time, resulting in an optimization of stimulus elaboration at the attended moment in time (Coull & Nobre, 1998). This suggests that temporal information is used to guide selectively attention to specific points in time. This results not only in behavioral advantages as visible from faster and/or more accurate responses to target stimuli, but also in changes of specific neural responses associated with temporal orienting and stimulus processing (see Nobre & Rohenkohl, 2014 for a review).

When a target stimulus is presented at the expected time, faster reaction times (RTs) are measured compared to conditions in which the time of occurrence of the target stimulus is unpredictable. This result is in line with earlier studies in the field of alertness demonstrating that when a warning signal (WS) is presented before a target stimulus, subjects are faster compared to conditions in which subjects are not alerted (Woodrow, 1914). The WS is used to inform subjects on the interval length before a forthcoming target stimulus, allowing for a correct estimation of the moment in which the target stimulus will be presented (thus triggering a faster response to it).

It has been observed that variation of the interval length between the WS and the target stimulus, the foreperiod (FP), influences temporal preparation, expectancy and thus reaction times to the target (see Niemi & Naatanen, 1981, for a review). In particular, a short FP (e.g., 800 ms) triggers faster RTs to a target stimulus compared to a long FP (e.g., 2,000 ms, Klemmer, 1956; Naatanen et al., 1974), due to the fact that temporal resolution over short interval of time is higher compared to long ones (Allan & Gibbon, 1991; Wearden & Lejeune, 2008). Finally, it has been showed that holding constant and predictable the FPs from trial to trial

primes a more precise temporal expectation (faster RTs) about the occurrence of the target stimulus compared to variable and unpredictable FPs (Woodrow, 1914).

The WS used in these tasks was initially interpreted as triggering a general mechanism of involuntary and automatic alertness (Posner & Boies, 1971). A more recent approach indicates the WS not only as trigger of a rigid mechanism of alertness but also of a voluntary dynamic allocation of attentional resources in time based on specific temporal expectancies (Coulle and Nobre, 1998; Nobre et al., 2007). Accordingly, knowledge about the time intervals before a target stimulus guide the orienting of attention voluntarily and dynamically toward its occurrence. Thus, faster RTs reflect an efficient extraction of temporal information from the WS, which is used as temporal cue to orient the attention to the moment when the target should appear.

A possible way to induce temporal expectancy/ prediction toward the occurrence of a target stimulus is by using the cueing paradigm. This paradigm uses symbolic cues to direct participants' expectancy/ prediction to future stimuli presented after cued intervals. In this paradigm, predictive and valid temporal cues (usually 80% of the total) are often combined with invalid cues (usually 20% of the total). The first type of cues predicts the forthcoming stimulus correctly while the second – invalid cues – are not followed by a target stimulus or its occurrence is unpredictable. Studies using the cueing paradigm observed that subjects show different behavioral and neural dynamics when a target stimulus is preceded by valid cues (cue that predict the most likely moment at which a target would occur) compared to invalid ones (Rohenkohl & Nobre, 2011). Valid cues typically elicit faster response times, increased accuracy and greater neural response to the target stimuli compared to invalid cues.

At the neural level, orienting attention selectively to points in time modulates a series of electrophysiological responses. For example, in their pioneering work, Walter and colleagues (1964) reported the presence of a physiological response known as the Contingent Negative Variation (CNV), a slow negative component originating in the motor areas of the brain and visible whenever subjects were presented with two stimuli: a warning signal followed by a target stimulus. They observed that a large negative response appeared just after the WS and reached the maximal amplitude peak at the anticipated time of the target stimulus (Walter

et al., 1964). It has been suggested that the CNV mirrors the ability of the nervous system to anticipate and prepare motor responses to temporally predictable stimuli. As such, it reflects temporal expectations and predictions of a future event (Tecce, 1972; Mento, 2013).

In addition to the neural modulations observed as function of the warning signal and expectancy related effects, other EEG studies have focused on the effects that orienting attention in time may have on target processing (Friedman et al., 2001; Lange, 2009; Correa and Nobre, 2008; Griffin et al., 2002; Miniussi et al., 1999). These studies described modulations of the fronto-central P3 component, which usually reaches its maximal positive peaks at about 300 ms from the onset of the target stimulus (Friedman et al., 2001; Lange, 2009; Correa and Nobre, 2008; Griffin et al., 2002; Miniussi et al., 1999). Larger P3 amplitudes were observed in condition where invalid cues preceded target stimuli compared to condition with valid cues (Gomez et al., 2009; Mangun and Hillyard, 1991). This has been interpreted in terms of a violation of the temporal expectations about the predicted target stimulus (Gomez et al., 2008; Mangun and Hillyard, 1991). Thus, target stimuli presented after invalid cues are elaborated as novel stimuli (Escera, 1998; Friedman, 2001), similarly to the expectancy violation effect observed in oddball paradigm, where subjects detect rare deviant stimuli randomly inserted among frequent standard stimuli and the P3 is larger for deviants than for standard ones (Ritter & Vaughan, 1969).

Taken together, these findings point to the human ability to predict/ anticipate stimulus appearance to improve its elaboration. Nevertheless, although several studies (Koelsch et al., 1999; Grahn & Rowe, 2009; Zuk et al., 2014; Habibi et al., 2014 see Hannon & Trainor, 2007 for a review) underlined the advantages of musicians in many perceptual tasks, less emphasis has been put in understanding whether musicians handle these perceptions more efficiently when orienting attention in time.

By means of behavioral and electrophysiological measures, the first aim of the present experiment was to investigate the effects of long-term musical training on the orienting of attention in time by looking at the processing of auditory target stimuli preceded by auditory or visual cues. I used a cross-modal cueing paradigm where auditory or visual cues preceded auditory target stimulus requiring participants' response.

Participants listened to a continuous auditory rhythmic sequence where auditory or visual cues informed them about the occurrence of a target stimulus (auditory omission) presented after short, long or mixed (either short or long) intervals of time.

First, in line with previous evidence (Allan & Gibbon, 1991; Wearden & Lejeune, 2008), I expected both musicians and non-musicians to show faster reaction times to target stimuli preceded by short compared to long intervals of time, because of the higher temporal resolution over short intervals of time. Moreover, I hypothesized that musicians would keep a high temporal resolution also for target stimuli presented after long intervals, due to an efficient tracking of events in time. Second, I explored whether effects of long-term musical training held cross-modally or if instead these strongly depend on the cue sensory modality. In this respect, I hypothesized that compared to non-musicians; musicians would show a superior ability to estimate the interval of time between cues and target stimuli specifically when auditory cues (since this is the mostly trained sensory modality for a musician) are used compared to visual cues.

At the neural level, the aim was to understand if effects of cues and musical expertise could be captured in different neural modulations before targets onset, which as such represented anticipatory activities (pre-stimulus activity, CNV), and/or were more linked to the actual stimulus elaboration, and thus, visible more at the level of the MMN and the P300.

As mentioned in the Introduction, musicians show enhanced rhythm processing abilities than to non-musicians as reflected in a sharp synchronization and a more precise expectation/prediction toward salient points of a rhythmical structure. Based on this evidence, the second aim of this experiment was to investigate differences in rhythm processing between musicians and non-musicians by looking at a particular neural response: the steady state evoked potential (SS-EP). This component has been observed in a series of EEG studies (Nozaradan et al., 2011, 2012, 2013; Celma-Miralles et al., 2016) investigating the neural mechanisms behind rhythm entrainment – that is, the synchronization to an external periodicity – which showed a selective enhancement of neural response at the exact frequency of the perceived rhythm. More specifically, these studies observed that the perception of a stimulus repeated periodically and with a fixed frequency, results in a change of the electrical activity recorded by electroencephalogram (EEG). This change, known as

Steady-State Evoked Potential (SSEP), has been showed to be stable in phase and amplitude over time. Therefore, the frequency of presentation of an external stimulus to which we attend delineates the frequency of the neural response. Hence, in keeping with the idea of an augmented rhythm processing in musicians, I hypothesized the presence of differences between the two groups in the neural entrainment to rhythm, as visible from higher SS-EPs amplitude associated to the processing of rhythmically presented sounds (the standard stimuli). Hence, although previous work demonstrated that regularly presented stimuli triggers a steady-state response at the exact frequency of the perceived regularity, it is not clear yet if this response is modulated by long-term musical experience.

## **4.2 Material and Methods**

### **4.2.1 Participants**

Twenty-four healthy volunteers took part in the experiment. I tested 12 non-musicians (age range = 18 - 36 years) and 12 musicians (age range = 18 - 36 years; 4 pianists, 2 guitarists, 1 violinist, 1 drummer and 4 clarinetists). I selected musicians who had a minimum of 6 years of formal musical training. All musicians reported playing their instrument on average 5 hours per day at the time of the experiments. Non-musicians did not have any formal musical or dance education. All participants provided their written informed consent before starting the experiment. None of the participants reported hearing, neurological or psychiatric disorder, and none was taking drugs at the time of the experiment. The Human Research Ethics Committee of the University of Trento approved the experiments (protocol number 2015-011).

### **4.2.2 Stimuli**

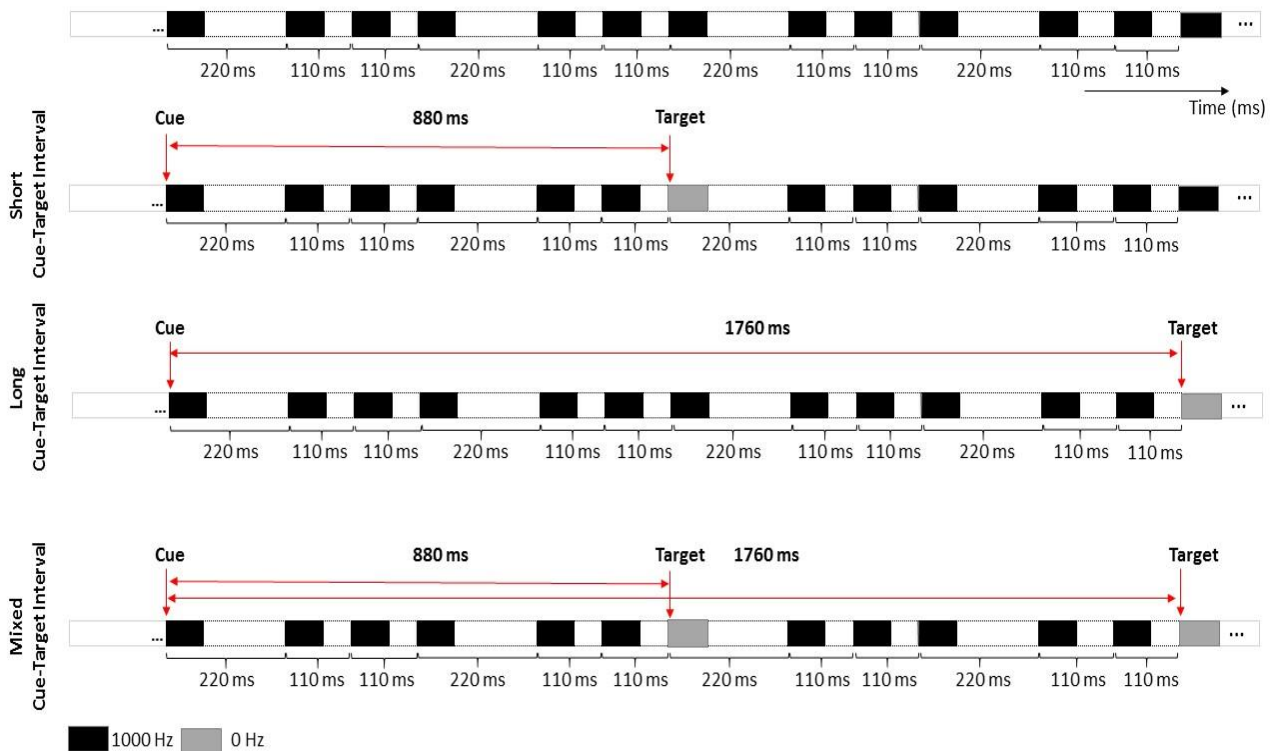
The stimuli comprised rhythmic blocks (see Procedure below) built as continuous repetitions of a module of 6 sounds, continuously looped for seven minutes and presented at a volume of 60 db SPL. Each sound was a 1000 Hz pure tone, lasting 60 milliseconds (ramped with 5 ms of rise time at the onset and 5 ms fall time at the offset to avoid clicks), and followed by a variable silent interval. Each module was created by assembling

in fixed order sounds with the following IOIs: 220, 110, 110, 220, 110, 110 milliseconds. The basic module of the rhythmic blocks is shown in Figure 5.1.

The rhythmic module (i.e., 220-110-110-220-110-110) evoked a clear binary meter based on the perception of tones' onsets. The first and the fourth sound in the gained beat saliency due to their long IOIs. From now on, I will refer to sounds with long IOI as metrically 'strong'. By contrast, I will refer to the remaining sounds in the module as metrically 'weak'. In each rhythmic block, 23.5 % of the modules included a target stimulus. Across blocks, the target stimulus was a zero Hz tone (i.e., an omission). Modules with a target stimulus were inserted in the block in pseudo-random order, with the constraint of a minimum of three and a maximum of seven modules without any target stimulus (i.e., 'standard' modules) before and after each target stimulus occurrence. Target stimuli occurred always on a strong metrical position (see Figure 5.1).

In each module, auditory or visual cues informed participants about the occurrence of a forthcoming target stimulus. I used 6 types of cues, 3 auditory (500 Hz tone, 1500 Hz tone and white noise) and 3 visual (yellow, violet and white cross). Each cue was associated with a specific cue-target interval. Low-frequency tone (500 Hz) and violet cross were associated with a short cue-target interval (880 ms), high-frequency tone (1500 Hz) and yellow cross with a long cue-target interval (1760 ms), white-noise and white cross were associated with mixed cue-target interval (either 880 ms or 1760ms, with unpredictable presentation; see Figure 5.1). Cues were inserted in correspondence of the first sound of the modules. Auditory cues were 60 milliseconds sounds (ramped with 5 ms of rise time at the onset and 5 ms fall time at the offset to avoid clicks). Visual cues lasted on the screen for 100 ms. both cue modality (auditory or visual) and cue-target interval (short, long or mixed) were blocked.





**Figure 5.1. Stimulus material.** Subjects listened to continuous auditory rhythmic sequences. A cue (either an auditory or a visual cue) indicated the length of the interval of time (880, 1760 ms or an unpredictable mixed of the first or second interval length) before the presentation of a future target stimulus (sound omission, 0 Hz sound).

Auditory stimuli were generated using the PsychToolbox extensions (Version 3.0.12.; October 2014) running under Matlab R2014b (The MathWork, USA). To guarantee that the effects found were not due to acoustic sound features per se all stimuli had the same acoustic energy and I did not use any musical accent. Stimuli were presented binaurally through electromagnetically shielded insert earphones (ER 1, Etymotic Research, Elk Grove Village, IL, USA), at approximately 60 dB SPL. I used a computer keyboard to record participants' responses.

#### 4.2.3 Procedure

Participants were seated comfortably in an electrically shielded room. They were instructed to fixate a central cross presented on a computer screen in front of them, and to listen carefully to the continuous auditory stream. I instructed participants on the association between each specific cue and interval length before a

target stimulus. Before starting the recording, a practice block was completed to memorize the cue-target intervals associations. Participants pressed a key as quickly as possible when a target stimulus was detected. Participants were tested over two experimental sessions of 3 hours each. In each experimental session, they completed 2 blocks per condition (2 blocks with the short cue-target interval, 2 with the long cue-target interval, and 2 with the mixed cue-target interval) for each cue modality. Each block included 130 standard modules (without any target event) and 40 modules with a target stimulus. Each block lasted 7 minutes and was followed by approximately 30 seconds of break. Block presentation order was counter-balanced across participants. The participants were asked to suppress saccades, avoid blinking and any other movement when a cue was presented.

#### **4.3 EEG Data Collection and Preprocessing**

EEG data were recorded from 28 Ag/AgCl electrode sites (bandpass filter: 0.016–250 Hz) with a BrainAmp amplifier system (Brain Products GmbH, Gilching, Germany). EEG channels were placed according to a subset of the 10-20 system. The reference electrode was placed on the right mastoid, and the ground electrode was positioned on AFz. The signals were amplified and digitized using a sampling rate of 1000 Hz, and re-referenced offline to mean activity of the left and right mastoids. Impedances were kept below 10 k $\Omega$  during EEG recording. The amplified and digitized EEG was stored on a computer disc for off-line analysis. EEG preprocessing was performed using Brain Vision Analyzer. Ocular correction ICA (Fast ICA restricted algorithm) was applied on the continuous data to remove artefacts related to eye movements. Continuous EEG signals were bandpass-filtered using a 0.1-40 Hz Butterworth zero-phase filter (Ladinig & Honing, 2009; Tierney & Kraus, 2013). On the bandpass-filtered data, two different analyses were carried out. One for the analyses of the neural responses following auditory or visual cues (CNV), and elicited by the target stimuli (MMN and P300). The second, to analyze the neural responses elicited by the processing of the standard stimuli (Steady-State Evoked Potentials, SS-EPs). For the first analysis, EEG epochs starting 50 ms before and ending 420 ms after the deviant onset were taken from the continuous recordings, after baseline correction (-50 ms to 0 relative to target stimulus onset). Epochs with an amplitude exceeding  $\pm 100$   $\mu$ V at any electrode in a 420 ms

time window relative to target stimulus onset were excluded from further analysis. EEG epochs were averaged separately for each participant and condition (Cue-Target interval short; Cue-Target interval long; Cue-Target interval mixed), separately for auditory and visual cues and for musicians and non-musicians. In the following sections, I describe separately the analyses performed on the CNV, MMN and P300.

#### **4.3.1 CNV**

Target stimuli occurred after two time intervals: either 880 or 1760 ms, or a combination of these in the mixed condition. Therefore, in order to compare the CNV amplitude for all conditions across the same interval of time, I extracted the CNV mean amplitude from 500 to 800 ms after cue onset at each electrode within a fronto-central ROI (Fz, Cz, F3, F4, C3, C4).

#### **4.3.2 MMN**

Visual inspection of grand-averages indicated a tendency to negative deflections between 100 and 300 ms after the onset of the target stimulus onset, with a frontocentral localization. This is in line with the latency and scalp distribution of the MMN (Schroger, 1998). For each participant, MMN mean amplitudes were extracted for each condition at Cz electrode (Näätänen, R., Gaillard, A. W., & Mäntysalo, S., 1978) considering a 60 ms time window around the most negative peak. Peak latencies were defined as the time point of the most negative peak. I calculated difference waves subtracting the ERP obtained from the standard stimuli from the ERP obtained for the deviant ones. The difference waves for each participant are thought to reflect only the additional activity elicited by the deviant stimulus, namely the MMN.

#### **4.3.3 P300**

P300 was analyzed considering the mean amplitude in a time window from 200 to 400 ms post target stimulus onset, registered at the electrode Cz (Näätänen, R., Gaillard, A. W., & Mäntysalo, S., 1978).

#### 4.3.4 Steady-State Evoked Potentials (SS-EPs)

EEG epochs from +1 to +23 s relative to target onset were taken from the continuous recordings. The first second of recording was removed from the analysis as justified in Nozaradan et al., (2011), in order to remove transient auditory evoked potentials linked to the percept of stimulus onset (Nozaradan et al., 2011, 2012). The overall interval was chosen to allow for the presence of several cycles of stimulation, which are necessary to prompt a steady state (Regan, 1989), and because a steady perception of rhythm is achieved only after a series of repetition of the same regular structure (Repp, 2005). Epochs containing artefacts exceeding  $\pm 150$   $\mu\text{V}$  at any electrode were excluded from further analysis. In order to enhance the signal-to-noise ratio of beat-related EEG activations and to attenuate any activity not phase-locked with the stimulation train (Nozaradan et al., 2011, 2012, 2013; Mouraux et al. 2011), for each subject, EEG epochs were averaged in the time-domain. I then transformed the average waveforms in the frequency-domain by applying the Fast Fourier Transform (FFT). Frequency spectrum had a resolution ranging from zero to 500 Hz with a frequency resolution of 0.045 Hz. The obtained signal is likely associated with the EEG activity elicited by the auditory beat (beat related SS-EPs). However, it may also contain residual background noise deriving from spontaneous activity, muscle activity, or eye movements. Therefore, as described by Nozaradan et al., (2011), in order to obtain valid beat-related SS-EPs I removed the contribution of this noise by subtracting the averaged amplitude of the two surrounding non-adjacent frequency bins ranging from -0.15 to -0.09 Hz and from +0.09 to + 0.15 Hz, at each frequency bin from 0.5 to 6 Hz. The rationale of this procedure is based on the assumption that if no beat-related SS-EPs are registered, than the signal amplitude in a frequency bin of interest (frequency of stimulation – 2 Hz) should be similar to that registered from neighboring ones, and should tend towards zero (Mouraux et al., 2011). Contrariwise, the subtraction procedure will not affect the presence of specific beat-related SS-EPs. Moreover, this subtraction procedure is important because the magnitude of the background noise is not equally present across electrodes. For instance, low-frequency noise derived from eye movements might be more evident in frontal electrodes, while high-frequency noise related to muscle activity might be stronger at temporal electrodes. Thus, this procedure is expected to result also in more reliable topographical maps of SS-EPs amplitude.

Finally, I estimated the magnitude of beat-related SS-EPs by averaging the signal amplitude from the three frequency bins centered on the corresponding frequencies of interest (1.08 Hz auditory SS-EPs: bins ranging from 1.04 to 1.13 Hz; 2.17 Hz auditory SS-EPs: bins ranging from 2.12 to 2.21 Hz; 4.34 Hz auditory SS-EPs: bins ranging from 4.30 to 4.39 Hz). With this procedure, I corrected for possible spectral leakage from our frequency of interest because the FFT do not estimate the SS-EPs at exactly the frequency of each SS-EP (Nozaradan et al., 2011, 2012a; 2012b).

## **4.4 Results**

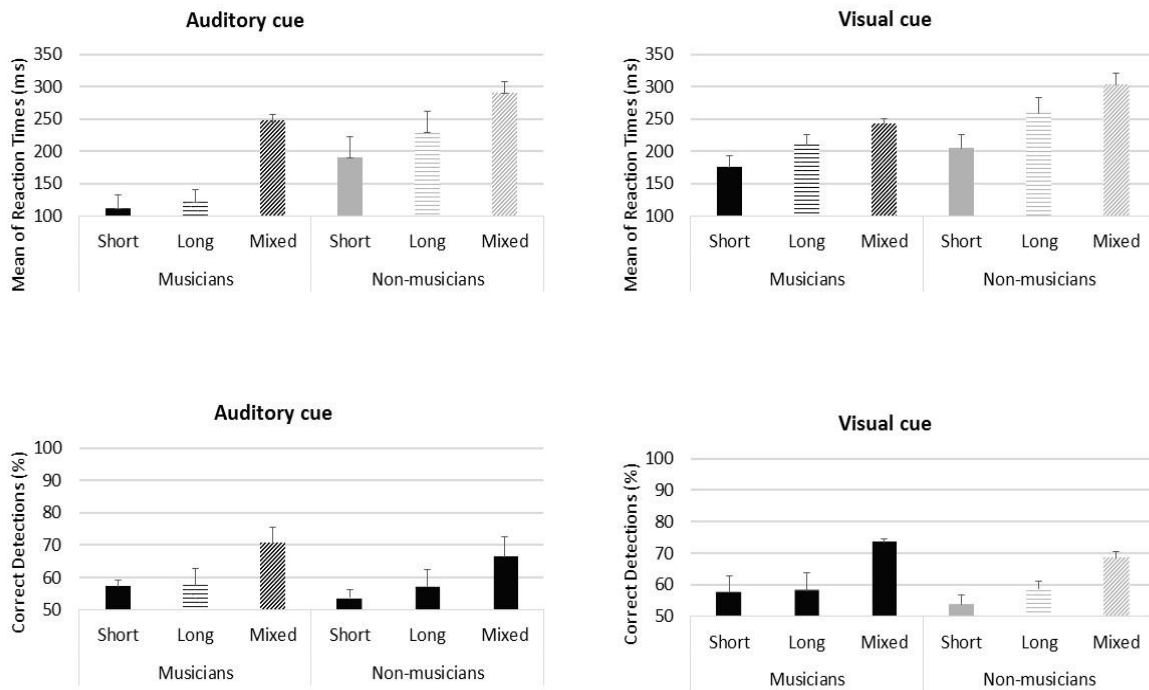
### **4.4.1 Behavioral results**

Behavioral data were analyzed with the following measures: Reaction Times (RTs) for correct responses (in the 0-100 ms post-target interval), Correct Detections, and Responses around target onset. Two post-hoc observations guided the choice of analysis.

First, as can be seen in Figure 5.2.1, there was a speed-accuracy trade-off, with participants being faster but more error prone in many conditions. Thus, to circumvent this problem, I calculated the Inverse Efficiency Scores (IES) by dividing the RTs by the proportion of correct responses. IES is suggested to be an “observable measure that gauges the average energy consumed by the system over time”. High IES values represent low efficiency. Second, I observed a low accuracy rate especially for the predictable conditions. This last observation guided the hypothesis that participants used the predictable cues to efficiently predict the moment in which a target would appear, thus leading to responses already before its actual onset (anticipations). For this reason, I decided to perform a frequency count on the time interval around target onset considering bins of 50 ms around it (see analyses steps below).

Each measure was entered separately in an Analysis of Variance (ANOVA) with Cue Target Interval (short, long and mixed) and Cue Modality (auditory, visual) as within-subjects factors and Group (musicians or non-musicians) as between-subjects factor. Post-hoc analyses were conducted via pairwise comparisons (t-tests). I used Bonferroni correction for all post-hoc comparisons. All statistical analyses have been carried out using the IBM SPSS Statistics 21 software.

RTs and Accuracy analyses are reported in the appendix at the end of the thesis. In the following section, I describe the analyses on IES and on the responses around the deviant onset.



**Figure 5.2.1. Behavioral Results of Experiment 4.** Reaction times (top row) and correct detections (bottom row) for deviants at the short, long and mixed cue-target intervals, per musicians and non-musicians and separated according to the factor cue-modality (results from auditory cues condition are shown on the left; results from visual cues condition are shown on right side of the figure). The figure displays a clear speed accuracy trade off, showing overall faster reaction times and lower accuracy levels.

#### 4.4.1.2 Inverse Efficiency score (IES)

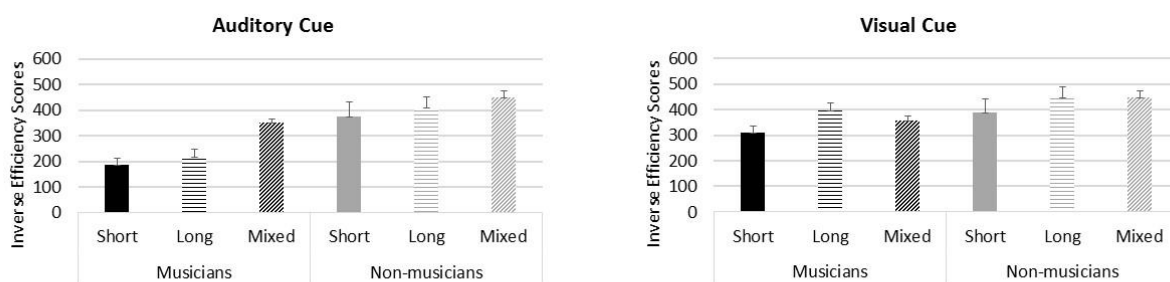
As mentioned above, an increase in IES indicates a decrease in efficiency. The analysis showed a significant main effect of Cue Target Interval ( $F(2, 44) = 12.839, p < .001, \eta_p^2 = .369$ ), a significant main effect of Cue Modality ( $F(1, 22) = 8.873, p = .007, \eta_p^2 = .287$ ), and a significant effect of Group ( $F(1, 22) = 11.303, p = .003, \eta_p^2 = .339$ ). The Cue Target Interval x Cue Modality ( $F(2, 44) = 5.317, p = .009, \eta_p^2 = .195$ ) and Cue Modality X Group ( $F(1, 22) = 5.021, p = .035, \eta_p^2 = .186$ ) interactions were also significant.

Follow-up pairwise comparisons (t-tests) were used to explore the Cue Target Interval x Cue Modality interaction. Considering the auditory cues, the analysis did not yield a significant difference between short

and long intervals ( $t(23) = -1.133, p = .270$ ), while a difference was found between the short and mixed intervals ( $t(23) = 4.100, p < .001$ ), highlighting smaller IES for the short intervals compared to the mixed one. Long intervals were also associated to lower IES than mixed ones, ( $t(23) = 3.346, p = .003$ ).

Considering the visual cues, the IES between short and long intervals revealed a significant difference ( $t(23) = 4.100, p < .001$ ), due to smaller IES for the short interval relative to the long one. Moreover, the data revealed a difference also between the short and mixed intervals ( $t(23) = 3.787, p = .001$ ), because of reduced IES for the short intervals compared to the mixed ones. No difference emerged between long and mixed intervals ( $t(23) = -.883, p = .386$ ).

Post-hoc analyses (t-tests) on the Cue Modality X Group interaction yielded significant differences between musicians and non-musicians mainly when target stimuli were preceded by auditory cues ( $t(22) = -4,164, p < .001$ ), showing smaller IES for musicians than non-musicians. When the visual cues were considered, a trend towards significance was found between the two groups ( $t(22) = -1,775, p = .090$ ), with lower IES for musicians than non-musicians. Figure 5.2.2 displays these results.



**Figure 5.2.2. Behavioral Results of Experiment 4.** Inverse Efficiency Scores for deviants at the short, long and mixed cue-target intervals, per musicians and non-musicians and separated according to the factor cue-modality (results from auditory cues condition are shown on the left; results from visual cues condition are shown on right side of the figure). The figure displays a clear speed accuracy trade off, showing overall faster reaction times and lower accuracy levels.

#### 4.4.1.3 Response frequency around target stimulus onset (-150-150 ms interval relative to stimulus onset)

The overall low accuracy observed guided the rationale supporting the analyses of the responses registered before target stimulus onset. To evaluate the time intervals where there were significant effects, without choosing a priori a fixed time interval, I followed a two-step procedure.

First, I conducted an ANOVA on the frequency counts for the responses around target stimulus onset with Bin (13 bins, from 500 ms before deviant onset to 150 ms after deviant onset in steps of 50 ms), Cue Target Interval (short, long, mixed) and Cue Modality (auditory vs. visual) as within-subjects factors, and Group (musicians vs. non-musicians) as between-subjects factor. The ANOVA indicated a significant main effect of Cue Target Interval ( $F(2, 44) = 53.893, p < .001, \eta_p^2 = .710$ ), with the highest number of responses before those stimuli that would have been presented after a short cue-target interval. A significant main effect of Cue Modality ( $F(1, 22) = 25.898, p < .001, \eta_p^2 = .541$ ), due to the largest number of responses recorded for auditory cues compared to visual ones. A significant main effect of Bin ( $F(12, 264) = 20.205, p < .001, \eta_p^2 = .479$ ), displaying a larger amount of target anticipations from 100 ms before its onset to 150 ms after it. The Cue Target Interval X Cue Modality ( $F(2, 44) = 7.093, p = .002, \eta_p^2 = .244$ ), Cue Modality X Bin ( $F(12, 264) = 16.424, p < .001, \eta_p^2 = .427$ ), Cue Target Interval X Bin ( $F(24, 528) = 12.459, p < .001, \eta_p^2 = .362$ ), Cue Target Interval X Cue Modality X Bin ( $F(24, 528) = 6.199, p < .001, \eta_p^2 = .220$ ) interactions were significant. Finally, the Cue Target Interval X Cue Modality X Bin X Group ( $F(24, 528) = 1.696, p = .021, \eta_p^2 = .072$ ) interaction was also significant. No other significant effect was observed (all  $p$ s  $> .114$ ).

Second, I conducted separate ANOVAs for each bin, and whenever the Cue Target Interval X Cue Modality X Bin X Group interaction was constant and showed the same direction in at least three adjacent bins, these were collapsed and re-considered in a single analysis. The results revealed the following constant effects:

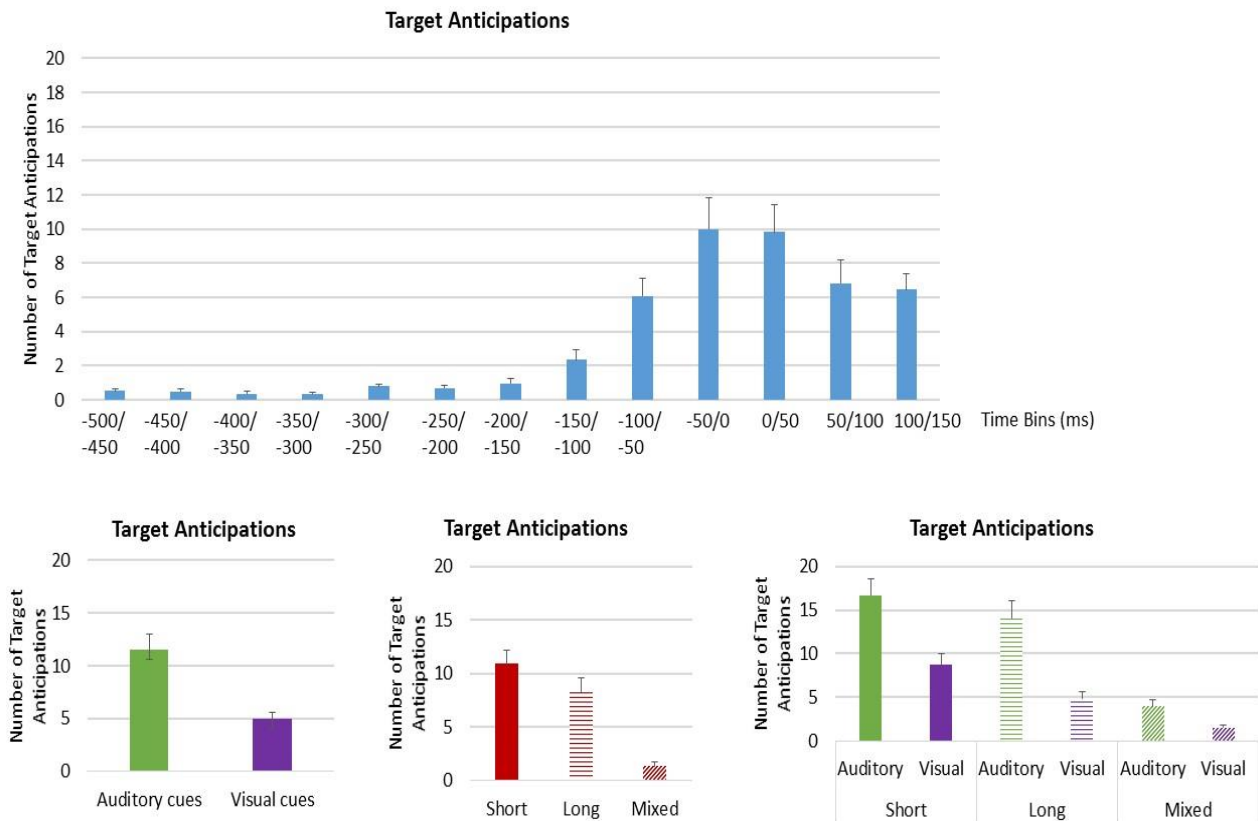
(1) A main effect of Cue Target Interval from 150 ms before to 50 ms after target onset. As expected, the highest number of responses were associated with the predictable short intervals as compared to the long intervals ( $t(11) = 3.997, p = .002$ ) and mixed ones ( $t(11) = 9.169, p < .001$ ). Additionally, more responses were found also for the long relative to the mixed intervals ( $t(11) = 6.730, p < .001$ ). Overall, as expected,



this underlies that the predictability of the target stimuli triggered an anticipatory response more than that observed for target stimuli presented after mixed (unpredictable) time intervals.

(2) A main effect of Cue Modality from -50 ms to 150 ms relative to deviant onset, due to a higher number of responses associated to auditory cues relative to visual cues. Differences between cue modalities were addressed by means of post-hoc analyses (t-tests) considering the entire time window collapsing the aforementioned bins. The results confirmed the general trend reported above ( $t(11) = 7.150, p < .001$ ).

(3) An interaction of Cue Target Interval X Cue Modality from 50 ms before deviant onset to 150 ms after deviant onset. Considering the auditory cues, the t-tests on the collapsed bins did not reveal any difference in the target anticipation between short and long intervals ( $t(23) = 1.544, p = .136$ ). Contrariwise, a significant effect was found between both predictable (short, long) and mixed intervals (short vs. mixed:  $t(23) = 7.946, p < .001$ ; long vs. mixed:  $t(23) = 6.490, p < .001$ ). The visual cues prompted the highest number of responses only for the short interval compared to all other intervals (short vs. long:  $t(23) = 4.371, p < .001$ ; short vs. mixed:  $t(23) = 6.794, p < .001$ , long vs. mixed:  $t(23) = 4.900, p < .001$ ). This highlights the influence that different interval lengths have on the ability to predict target stimuli cued by the visual modality. The reported results are showed in Figure 5.2.3.



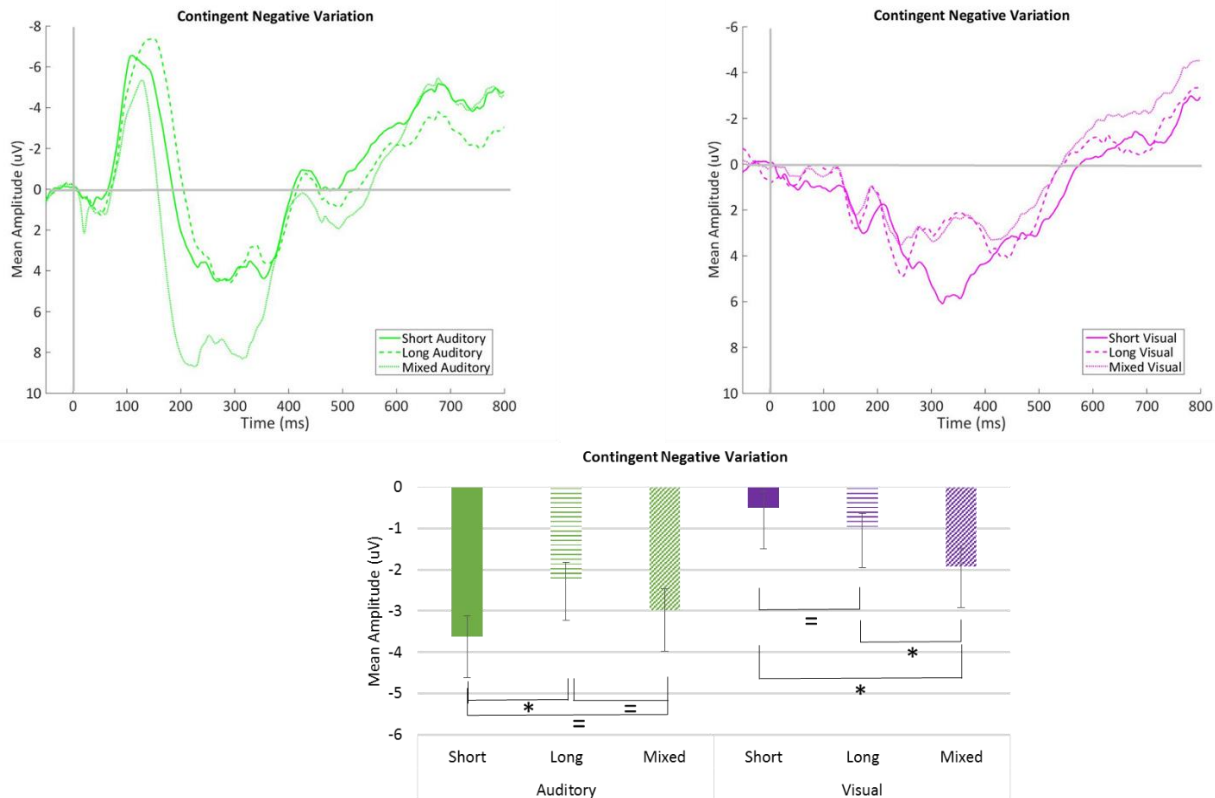
**Figure 5.2.3. Behavioral Results of Experiment 4.** The top row shows the mean of target anticipations in each time bin from -500 ms before target onset to 150 ms after target onset (in time steps of 50 ms). Results are collapsed over groups, cue-target intervals and cue modalities. In the bottom row is represented the mean of target anticipations separated per auditory and visual modality (left), considering each cue-target interval (center), and comparing auditory vs. visual modality for each cue-target interval (short, long, mixed). The factor group is collapsed in all graphs since no differences between musicians and non-musicians were observed.

#### 4.4.2 EEG results

##### 4.4.2.1 Contingent Negative Variation (CNV)

CNV amplitude values were analyzed by means of an ANOVA with Cue Target Interval (short, long, mixed) and Cue Modality (auditory and visual) as within-subjects factors, and Group (musicians or non-musicians) as between-subjects factor. All statistical analyses that will be reported have been carried out using the IBM SPSS Statistics 21 software. The ANOVA revealed main effects of Cue Target Interval ( $F(2, 44) = 3.270, p = .047, \eta_p^2 = .129$ ), Cue Modality ( $F(1, 22) = 33.046, p < .001, \eta_p^2 = .600$ ), and of the interaction Cue Target Interval X Cue Modality ( $F(2, 44) = 13.296, p < .001, \eta_p^2 = .377$ ). No other significant effects were found (all  $p > .113$ ). The observed interaction is shown in Fig. 4.3.1. Post-hoc analyses (t-tests) on the Cue Target

Interval x Cue Modality interaction were conducted to explore the influence of different interval lengths on the CNV amplitude as function of cue modality. As concerns the auditory cues, t-tests highlighted a significant difference between short and long Cue Target Interval ( $t(23) = -4,492, p < .001$ ), with the short interval boosting the CNV amplitude compared to the long one. However, no significant difference emerged between short and mixed intervals ( $t(23) = 1,295, p = .208$ ), and between long and mixed intervals ( $t(23) = -1,477; p = .153$ ). As for the visual cues, the t-tests depicted an opposite scenario. In fact, no differences were observed between the short and long Cue Target Interval ( $t(23) = 1,443, p = .163$ ). Instead, the CNV amplitude was significantly different for the short compared to the mixed interval ( $t(23) = -3,879, p = .001$ ), and between the long and mixed one ( $t(23) = -2,558; p = .018$ ), due to the larger amplitude elicited by the mixed interval. Overall, the results showed a reduction of the CNV amplitude following visual cues. This reflects the lower efficiency of the visual modality to guide attention in time and, in particular, to trigger expectancy toward target stimuli presented in a different modality (auditory).



**Fig. 5.3.1. CNV.** Grand-averaged ERPs recorded over a fronto-central ROI (Fz, Cz, F3, F4, C3, C4) representing the anticipations for forthcoming target stimuli triggered by auditory (left, green) and visual (right, violet) separately for short (880 ms, solid thick line), long (1760 ms, dashed line) or mixed (either 880 ms or 1760 ms, solid thin line) intervals of time. The results are collapsed for musicians and non-musicians since the analysis did not reveal any significant effects of group nor interaction. The bottom row show a bar graph of the CNV mean amplitude depicted above.

#### 4.4.2.2 CNV over centro-frontal and posterior-occipital Regions of Interest (ROI)

Since the cue could be visual or auditory, I conducted an analysis aimed to disentangle potential differences in the CNV amplitude elicited by auditory and visual cues by comparing responses in areas generally activated by visual or auditory stimuli. To this aim, two regions of interest (ROIs) were chosen according to electrodes position over centro-frontal (Cz, Fz, F4; F3; C3; C4) and posterior-occipital (Oz, O1, O2, P3, P4, Pz) sites (Fig...). Statistical analyses were conducted on CNV mean amplitudes by means of an ANOVA with Cue Target Interval (short, long, mixed), Cue Modality (auditory or visual) and ROI (visual or auditory) as within-subjects factors, and Group (musicians or non-musicians) as between-subjects factor.

The ANOVA revealed a trend towards a significant effect of Cue Target Interval ( $F(2, 44) = 2.613, p = .085, \eta_p^2 = .106$ ), a significant effect of Cue Modality, ( $F(1, 22) = 21.960, p < .001, \eta_p^2 = .500$ ), and a significant main

effect of ROI ( $F(1, 22) = 102.868, p < .001, \eta_p^2 = .824$ ), due to a larger CNV amplitude registered over Centro-frontal ROI as compared to the parietal occipital one. The Cue Target Interval X Cue Modality ( $F(2, 44) = 16.336, p < .001, \eta_p^2 = .294$ ), Cue Modality X ROI ( $F(1, 22) = 8.583, p = .008, \eta_p^2 = .281$ ), Cue Target Interval X Cue Modality X ROI ( $F(2, 44) = 5.816, p = .006, \eta_p^2 = .209$ ) interactions were significant. No other significant effects were found (all  $ps > .127$ ).

Post-hoc analyses (ANOVAs) were conducted separately for the auditory and visual cues with Cue Target Interval and ROI as factors. The ANOVA for the auditory cues revealed a significant main effect of ROI ( $F(1, 22) = 132.975, p < .001, \eta_p^2 = .858$ ), and of Cue Target Interval X ROI ( $F(2, 44) = 5.245, p = .009, \eta_p^2 = .193$ ). No other significant effects were found (all  $ps > .1$ ). Pairwise comparisons (t-tests) to evaluate the Cue Target Interval X ROI interaction showed larger CNV amplitude values over frontal central ROI as compared to the parietal occipital ROI for each cue target interval considered (all  $ps < .001$ ).

The ANOVA for the visual cues revealed a significant effect of Cue Target Interval ( $F(2, 44) = 8.206, p < .001, \eta_p^2 = .272$ ), with the mixed interval producing the largest CNV amplitude than all other intervals. A significant main effect of ROI ( $F(1, 22) = 56.723, p < .001, \eta_p^2 = .721$ ) also emerged, again with larger CNV amplitude over fronto-central ROI as compared to the parietal occipital ROI. No other significant effects were found (all  $ps > .233$ ).

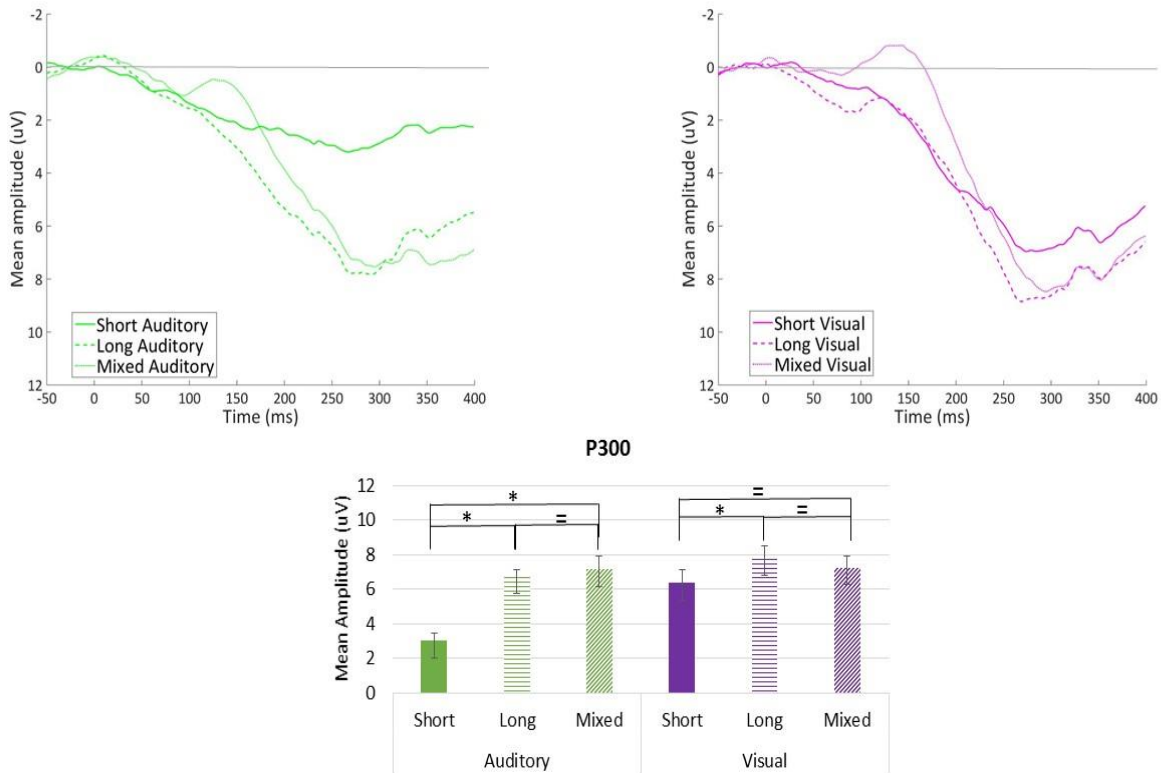
Overall, these analyses confirmed a larger CNV amplitude over fronto-central sites than parietal occipital ones irrespective of the cues' sensory modality.

#### **4.4.2.3 MMN**

To test for the presence of a reliable MMN, I compared the MMN amplitude difference values against the null mean in all conditions by means of one-sample t-tests. The analyses indicated only a trend for more negative values for the visual mixed condition ( $t(23) = -1.830, p = .080$ ) (see Figure 5.3.2). Therefore, I decided not to analyze this component further.

#### 4.4.2.4 P300

Amplitude of the P300 component were analyzed by means of ANOVA with the same factors used for the behavioral analyses. The results revealed significant effects of Cue Target Interval ( $F(2, 44) = 15.472, p < .001, \eta_p^2 = .413$ ), Cue Modality ( $F(1, 22) = 13.184, p = .001, \eta_p^2 = .375$ ), and of the Cue Target Interval X Cue Modality ( $F(2, 44) = 16.094, p < .001, \eta_p^2 = .422$ ) interaction. There were no other significant effects (all  $p$ s  $> .619$ ). Follow-up comparisons (t-tests) on the Cue Target Interval X Cue Modality interaction were conducted to evaluate the impact of intervals lengths on the P300 amplitude elicited by target stimuli cued by either the auditory or the visual modality. Considering the auditory cues, the analyses indicated a significant difference between the short and long ( $t(23) = -5.656, p < .001$ ) as well as mixed ( $t(23) = 5.812, p < .001$ ) intervals, with the target stimuli after the short intervals resulting in a smaller P300 compared to both the other intervals (long and mixed). No significant difference emerged from the comparison between long and mixed intervals, ( $t(23) = .547, p = .590$ ). Thus, higher expectancy level might be associated to stimuli presented after a short interval; consequently, these are not processed as novel stimuli as those presented after long or mixed intervals. As for the visual cues, the t-tests yielded a significant difference between the short and long intervals ( $t(23) = -2.489, p = .020$ ), resulting in smaller P300 amplitude for the short than long interval. However, the effect was not strong enough to be visible also when comparing the short and mixed intervals ( $t(23) = 1.702, p = .102$ ), nor for the long and mixed ones ( $t(23) = -1.267, p = .218$ ). Figure 5.3.2 shows the reported results.



**Fig. 5.3.2. MMN and P300.** Grand-averaged ERPs recorded at Cz elicited by target stimuli after auditory (left, green) or visual (right, violet) cues separately for short (880 ms, solid thick line), long (1760 ms, dashed line) or mixed (either 880 ms or 1760 ms, solid thin line) intervals of time. The results are collapsed for musicians and non-musicians since the analysis did not reveal any significant effects of group nor interaction. The bottom row show a bar graph of the P300 mean amplitude depicted above.

#### 4.4.2.5 SS-EPs

Figure 5.3.3 shows the average frequency spectra (Hz) of the EEG amplitude signals after the noise-subtraction procedure, for musicians (blue line) and non-musicians (red line) for each frequency of interest (1.08, 2.17 and 4.34 Hz), averaged across all scalp electrodes. As shown in the figure, the auditory rhythm elicited clear SS-EPs at 1.08, 2.17 and 4.34 Hz corresponding to the frequency and related subharmonic and harmonic of the beat, and referred as beat-related SS-EPs. A clear increase in the signal amplitude appears for both musicians and non-musicians at the frequency of 1.08 Hz, which corresponds to the repetition of the first tone of a module of 6 tones. Hence, the neural activity captured in this SS-EP suggests that participants associated specific relevance to the first tone (metrical position strong, first hierarchical subdivision) out of the six tones of a rhythmic module. However, non-musicians displayed a larger SS-EP

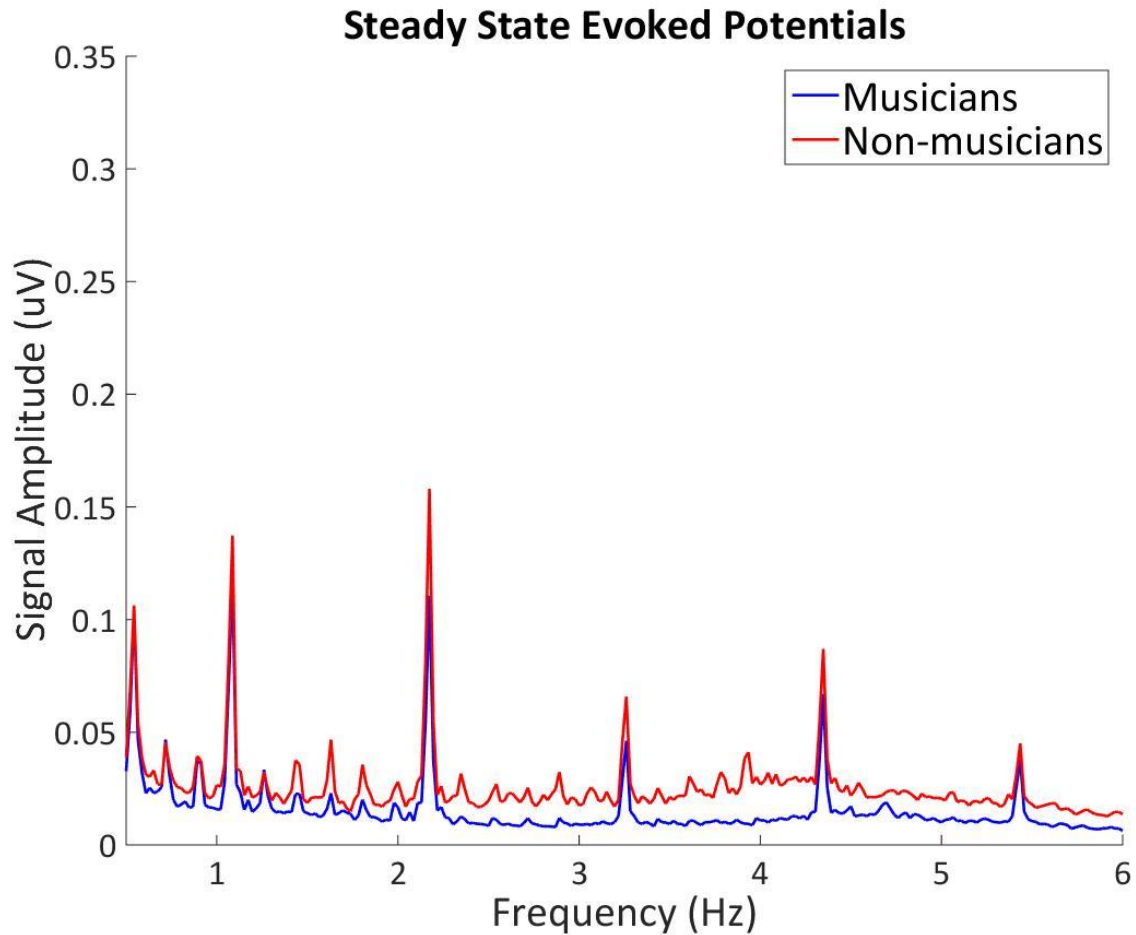
amplitude at 2.17 Hz compared to musicians. Furthermore, there is a large peak for the first harmonic (4.34 Hz) for all participants.

I did not have a priori hypotheses on the scalp topography related to beat-induced responses. Thus, to exclude any electrode selection bias, for each subject and frequency of interest, I averaged the magnitude of the SS-EPs across all electrodes. I first performed one-sample t-tests in order to determine if the noise-subtracted SS-EPs amplitudes were significantly different from zero. In fact, if SS-EPs were not present, the average of the noise-subtracted signal amplitude would tend towards zero. The analyses demonstrated that for all frequency of interest the signal amplitude were significantly different from zero (for 1.08 Hz:  $t = 12,425$ ,  $p < .001$ ; for 2.17 Hz:  $t = 12,403$ ,  $p < .001$ ; for 4.34 Hz:  $t = 11,944$ ,  $p < .001$ ).

To explore whether musicians differ from non-musicians in the strength of rhythm synchronization, I compared the amplitude of the beat-related SS-EPs averaged across auditory and visual cues, by means of ANOVA with Frequency of Interest (1.08 Hz, 2.17 Hz and 4.34 Hz) as within subject factor, and Group (musicians and non-musicians) as between subjects' factor. Bonferroni correction was applied for all post hoc comparisons.

The ANOVA revealed a significant main effect of Frequency of Interest ( $F(2, 22) = 24.781$ ,  $p < .001$ ,  $\eta_p^2 = .530$ ), showing smaller SS-EPs amplitude elicited at 1.08 Hz frequency beat than both 2.17 Hz ( $t = -5.951$ ,  $p < .001$ ) and 4.34 Hz ( $t = -4.28$ ,  $p = .001$ ) frequencies. It also indicated a marginally significant greater SS-EPs amplitude elicited by the beat frequency of 2.17 Hz than 4.34 Hz ( $t = 1.877$ ,  $p = .087$ ). There was a significant effect of Group ( $F(1, 22) = 4.200$ ,  $p = .053$ ,  $\eta_p^2 = .160$ ), due to higher SS-EPs amplitude observed for non-musicians compared to musicians. The Frequency of Interest X Group ( $F(2, 44) = 5.656$ ,  $p = .007$ ,  $\eta_p^2 = .204$ ) interaction was also significant. Pairwise comparisons (t-tests) indicated greater SS-EPs amplitude in non-musicians as compared to musicians for the beat frequency of 2.17 Hz ( $t = -3.514$ ,  $p = .002$ ), whereas 1.08 Hz and 4.34 Hz beat frequencies did not indicate any significant difference between groups (1.08 Hz:  $t = -.607$ ,  $p = .550$ ; 4.34 Hz:  $t = -.233$ ,  $p = .818$ ).





**Fig. 5.3.3. SS-EPs.** Grand-average frequency spectra (Hz) of the noise-subtracted EEG amplitude signals elicited by the rhythmical sequences for musicians (blue line) and non-musicians (red line), averaged across all scalp channels. For all subjects, the rhythmical stimulus elicited a clear steady state EP at 1.08 Hz, 2.17 Hz and 4.34 Hz. Difference between groups was reported only in the SS-EP amplitude at 2.17 Hz.

#### 4.5 Discussion

The present experiment sought to investigate the effects of long-term musical training on the ability to orient attention in time using a cross-modal cueing paradigm with auditory and visual cues. With this manipulation, I examined whether the effects of long-term musical training transfer also to a different sensory modality than the mostly trained one for a musician (i.e., the auditory one). The behavioral results revealed a speed-accuracy trade off showing overall fast response time and low percentage of correct detections. Thus, the analyses were centered on inverse efficiency scores. Moreover, the low accuracy level for predictable targets was considered as reflecting a mechanism of anticipation associated to target occurrence. Thus, to

understand if predictable targets indeed prompted a response before or immediately after target onsets, I report the response frequencies obtained in time bins around target onset.

The paradigm of this experiment allowed for the investigation of three EEG components: the CNV component, which reflects the amount of anticipation before the actual occurrence of a target stimulus; the MMN and P300 components, which are linked to the actual target analysis. Finally, I considered SS-EPs to test for effects on musical expertise on rhythm synchronization.

#### **4.5.1 Behavioral results**

Overall, the results points to three aspects.

First, I observed lower inverse efficiency scores (i.e., higher efficiency) in musicians compared to non-musicians in all experimental conditions, indicating an enhanced ability to detect target stimuli due to long-term musical training. Given that cues and target stimuli were inserted within rhythmical sequences, I speculate that the high efficiency in musicians may depend on a refined rhythm processing. Indeed, as demonstrated by previous work (Rammsayer & Altenmüller, 2006; Chen et al., 2008), long-term training with musical stimuli strengthen the encoding and production of rhythm. This may act as an additional cue, together with the instantaneous cues used in the experiment, to boost musicians' efficiency in detecting the target elements. Moreover, the fact that musicians' efficiency was magnified with auditory relative to visual cues highlights that musical training is (at least partially) modality-dependent, as the effects are most evident for the mostly trained sensory modality (i.e. audition).

Second, regardless of musical expertise, the results indicated that target detection was more efficient after short cue-target intervals. This finding is in line with the existing literature using a constant foreperiod paradigm (where the interval between a cue and a target is kept constant in a block; e.g., Niemi & Naatanen, 1981; Klemmer, 1956), which demonstrated a decrease in reaction times as the duration of the interval of time between cue and target stimuli decreases (Coull & Nobre, 1998; Griffin et al., 2001; Miniussi et al., 1999). It could be suggested that the temporal certainty peculiar of short time intervals enhances target

stimuli encoding (Niemi & Naatanen, 1981) and results in an overall higher efficiency level relative to long and mixed intervals, in which temporal certainty is reduced. Alternatively, temporal certainty may lower the decision threshold (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010) and, as such, lead to an anticipation of the responses to the target stimulus. The idea that short intervals may induce lower response thresholds is supported by the results on the frequency of responses registered around the target onset. Indeed, I found that a large amount of responses were centered in the time window around predictable target onset (from -150 ms to 50 after it) relative to unpredictable ones.

Third, in line with the idea that the auditory modality is more tuned for temporal information than the visual modality (Posner et al., 1980), target stimuli preceded by auditory cues resulted in higher efficiency levels and a higher number of anticipations (in the time window from 50 milliseconds before to 150 milliseconds after target onset). In addition, while the auditory cues were equally efficient, regardless of the cue-target interval length, visual cues facilitated deviance detection only for short target-cue intervals. The effect was visible in efficiency scores as well as target anticipations. Indeed, when looking at responses around target onset (from 50 ms before deviant onset to 150 ms after it), the auditory cues triggered an equal anticipatory activity over both intervals (short and long), while the opposite scenario was depicted for the visual cues, which prompted the highest number of responses specifically for the short interval compared to all other intervals. A possible explanation of the reason why visual cues efficiently orient subjects' attention only over short intervals only is linked to the paradigm of this experiment, which was a modified version of the traditional paradigm used in the literature of temporal orienting. Specifically, here Foreperiods were not silent intervals but subjects listened to a continuous auditory rhythm throughout each block, and as such also between each cue and target event, leading to a higher relevance of the auditory modality compared to the visual one in orienting attention to events presented after short as well as long intervals.

## **4.5.2 EEG results**

In contrast with our hypothesis, EEG results did not reveal any effects of long-term musical training. As I will discuss in the General Conclusion (Chapter 5), one possibility to account for this results is that the experimental paradigm I used was not designed to investigate differences at the level of movements' production and organization on which usually musicians and non-musicians show the largest differences.

### **4.5.2.1 Pre-target stimulus response**

The results showed a larger CNV amplitude when auditory cues were used as compared to visual ones. This result is in line with previous evidence and emphasizes the ease with which the auditory modality guides attention in time (Gaillard and Naatanen, 1976; Correa et al., 2006; Rockstroh et al., 1982) compared to the visual modality that is more apt to space processing (Posner, 1980). Earlier work described the larger CNV amplitude elicited by auditory cues in terms of the higher alerting property of an auditory stimulus compared to a visual one (Posner et al., 1976) when orienting attention in the time domain. Moreover, this effect may also reflect the ease of a unimodal processing (anticipation of a target in the same sensory modality of the cue) compared to the load associated to the cross-modal processing of a visual cue and a subsequent auditory target stimulus.

Furthermore, according with previous findings reporting larger CNV amplitude for short than long cue-target intervals (McAdam, 1969; Miniussi et al., 1999), these data further indicated that auditory cues predicting a target stimulus after a short interval triggered a larger expectancy level as reflected in a larger CNV amplitude relative to long intervals, and an equal amount of expectancy as the mixed (unpredictable) condition. This is in line with the assumption of the CNV as neural signature of expectancy to forthcoming events (Mento, 2013). Thus, the data suggest the occurrence of a high level of preparation also in the unpredictable condition, in order to reduce the overall uncertainty. Nonetheless, the relatively low robustness of the statistical outcomes (see Results section) recommends caution in the interpretation of this pattern.

In the visual cue condition, cues predicting the short and long intervals resulted in a smaller CNV amplitude with respect to the mixed condition. Thus, in contrast with the auditory modality, predictive visual cues seem

to cause interference in the amount of expectancy towards an auditory events. This may be due to the need to interpret the visual cues and their predictability level, and to the fact that the visual modality does not match the modality of the target event (which was always an auditory event).

#### **4.5.2.2 Post-target stimulus response**

The paradigm used in this experiment was a modified oddball paradigm wherein deviant stimuli were always preceded by a cue predicting their occurrence after various cue target intervals; consequently, detection of deviants did not result in a MMN response, which is usually elicited by unpredictable and rare deviant stimuli.

In contrast, temporal orienting influenced the P300 amplitude, showing a different modulation as a function of cue-target interval for the visual and auditory cue modalities. In particular, stimuli presented after the short interval resulted in the smallest P300 amplitude compared to all other intervals. Compatible with previous studies, temporal orienting had larger effect for target occurring after short rather than long foreperiods (Miniussi et al., 1999; Correa et al., 2006; see Nobre, 2001 for a review). Here, I speculate that the short and predictable intervals elicited the highest level of certainty about the target occurrence, reducing the need to update the stored information about the auditory context as well as reducing the perception of “novelty” of the deviant stimulus. Therefore, consistent with the idea that the P300 amplitude reflects the subjective expectancy to the forthcoming presentation of a target stimulus (Donchin, 1981; Muller et al., 2003), smaller P300 amplitude were observed for the short cue-target intervals in which, as reported above, a higher level of certainty was experienced as compared to the long cue-target intervals.

Interestingly, differences in the P300 amplitude might also depend on a different modulation of the preceding CNV response. Indeed, as reported in the section 4.5.2.1, the higher expectancy level associated to target after a short cue-target interval was evidenced by a larger CNV amplitude and as these data showed, by a reduction of the P300 amplitude highlighting a reduction of novelty associated to the target presentation. This further suggest that target expectancy, as captured by behavioral (anticipation) and neural (CNV) modulations facilitate stimulus detection as evidenced by long latency neural components.

Stimuli predicted by visual cues did not show clear effects in the P300 component. Overall, the P300 amplitude elicited by stimuli preceded by visual cues was larger than that cued via the auditory modality, suggesting that visual cues are overall associated with high levels of uncertainty. This corroborates previous findings pointing to a relative inefficiency of the visual modality in orienting attention in time, compared to the auditory modality. Moreover, this pattern of results might also points to the higher cost associated to the cross-modal processing (indeed, as reported in the section 4.2.2 of the Method, auditory target stimuli could be preceded by visual other than auditory cues), as compared to the unimodal processing of an auditory cue signaling the occurrence of an auditory target.

#### **4.5.2.3 Steady State Evoked Potentials**

Previous studies showed that the neural synchronization to the beat and related harmonics can be captured by the occurrence of a steady-state evoke response at the frequency of the beat (Nozaradan et al., 2011, 2012, 2013; Celma-Miralles et al., 2016). Nevertheless, there is no information to date on potential differences between musicians and non-musicians in this steady response. Here, I examined whether musicians show a different modulation of the SS-EPs relative to non-musicians.

Perceiving a rhythm with a specific metric structure adds to the beat frequency additional periodicities that correspond to harmonics and subharmonics of the beat frequency. Thus, I evaluated the SS-EPs to the beat frequency (2.17 Hz, corresponding to the two strong beats in the rhythmic sequence used in this experiment), its first harmonic ( $2f$ , 4.34 Hz) and subharmonic ( $f/2$ , 1.08 Hz, corresponding to the first strong beat in the rhythmic sequence used in this experiment).

In line with recent research demonstrating larger SS-EPs amplitude at the beat and meter frequencies (Nozaradan et al., 2011, 2012), I observed a stronger response at these frequencies compared to the other frequencies of the rhythm. Importantly, differences between musicians and non-musicians were captured at the beat frequency (2.17 Hz) as showed by a smaller SS-EPs amplitude for musicians than non-musicians. I suggest that this might reflect the ease with which musicians processed some specific components of rhythm. As a result, the stronger response of non-musicians at the beat frequency (2.17 Hz), which

corresponded to the two strong beats of the rhythmic sequence used in this experiment, may indicate their need for a more effortful tracking of rhythm in time (and, as such, of a synchronization to both strong beats of rhythm).

#### **4.6 Conclusion**

In summary, orienting attention in time improves stimulus processing, modulate neural correlates associates to its perception and have an impact on both behavioral and neural dynamics before its actual perception. In particular, temporal orienting seems to be highly influenced by the cue-target interval length, with short intervals eliciting the highest efficiency, triggering a trend for augmented neural (CNV) and behavioral anticipations, and resulting in a reduced novelty effect (reduced P300). Additionally, I found different effects deriving from auditory and visual cues to direct attention to target occurrence, which was always a sound omission in this experiment. In this respect, auditory cues were more effective as visible from the behavioral and neural patterns. This finding could be framed within previous reports of augmented temporal acuity of the auditory modality compared to the visual one, which is more apt for spatial tasks (Posner et al., 1980); and it further corroborates the higher efficiency of a unimodal relative to a cross-modal stimulus elaboration. Finally, although other abilities other than musical training one might have influenced on the orienting of attention to points in time, behavioral results revealed a stronger influence of long-term musical training compared to short-term one with the instantaneous cues, on the ability to orient attention in time; in particular, this was even heightened when auditory cues were used thus suggesting a specific effect of long-term musical training with stimuli delivered in the mostly trained modality for a musician (auditory).





## Chapter 5

### Conclusions

The research presented in this thesis focused on two main aspects.

First, I examined the effects of long-term musical training on the processing of auditory temporal structures (rhythmical, non-rhythmical and metrical aspects of rhythm) using a task that did not imply music knowledge and in contexts where rhythm/meter were not task-relevant (Chapter 2 and 3). To highlight this aspect, I designed a series of behavioral and EEG experiments in which musicians and non-musicians were tested on the detection of auditory deviant stimuli inserted within simple auditory temporal structures (rhythmical, non-rhythmical). All these aspects were addressed using non-musical stimuli to analyze if the effects of long-term musical training are visible also when moving out from a non-musical context.

The approach I used in these experiments was to analyze behavioral (reaction times- RTs- and signal detection measures- SDT) and ERP (MLR, MMN – Experiment 3) responses to deviants as function of rhythm/meter elaboration. The choice of adopting SDT measures was aimed to dissociate whether long-term musical training influences more the sensitivity of participants' auditory system (as measured by  $d'$  values) or the response bias (as measured by the criterion). The choice of simultaneous behavioral and EEG recordings during task performance (Experiment 3) additionally allowed for a detailed understanding of the processes involved in long-term musical training. Indeed, the high temporal resolution of EEG measures provides a straightforward mean to investigate on-line the temporal neural dynamics of stimulus analysis. Specifically, the extraction of early (MLR) and late (MMN) latency ERP responses associated with deviance

detection provided a more complete picture of the specific (if any) level of stimulus analysis that is modified by long-term musical training.

Experiment 1, described in Chapter 2, compared musicians and non-musicians looking at RTs, sensitivity and criterion associated with the detection of two types of auditory deviant stimuli (frequency and temporal deviants) inserted within rhythmical and non-rhythmical temporal structures. Results revealed an overall superior performance of musicians compared to non-musicians, indicating that long-term musical training leads to changes in the way auditory stimuli are elaborated (Herholz & Zatorre, 2012; Klein et al., 2016; Munte et al., 2002), even when these are non-musical items. Interestingly, as the SDT results revealed, this superiority was linked to sensitivity and was not the result of a different response criterion used to detect deviant stimuli. Moreover, the data pointed to a large-scale processing of rhythm independent of musical expertise. In fact, all subjects were faster, showed a higher sensitivity and a more liberal response criterion when detecting deviants embedded within rhythmical as compared to non-rhythmical temporal structures. This corroborates previous findings highlighting the spontaneous ability to perceive rhythm, which for example has been observed also in newborns (Winkler et al., 2009; Zentner & Eerola, 2010).

In Experiment 2, reported in Chapter 2, I focused on the effects of long-term musical training on a peculiar aspect of rhythm: the meter. The point here was to highlight if differences between musicians and non-musicians would have been more evident in a superior processing of the metrical structure of rhythm. Indeed, it could be thought that because of their training with various rhythms musicians process rhythm differently from a non-musicians, for example using different strategies to group sounds of a rhythmical sequence. In particular, it might be speculated that musicians form bigger and more complex groups of sounds in which different level of relevance is attributed to each sound; contrariwise, non-musicians might find difficult to consider elaborated groups of sounds and rather prefer a sequential processing of sounds. In this context, if from one side the literature reports contrasting findings, with some studies pointing to a superior meter elaboration in musicians compared to non-musicians (Jogsma et al., 2003; 2004; Geiser et al.,

2010), while others failing to report such differences (Geiser et al., 2009); on the other side, it is not always clear whether behavioral and neural differences found could be considered a real signature of a different elaboration of rhythm/meter as function of long-term musical training, or if instead these simply reflected the ease with which musicians elaborated complex musical stimuli/tasks. In this experiment, I attempted to overcome these limitations manipulating deviants' occurrence across metrical positions and using non-musical stimuli that were not relevant to the task at hand. Hence, I looked at RTs, sensitivity and criterion used by subjects to detect deviant stimuli (frequency deviant) at either strong (more salient) or weak (less salient) metrical positions of the rhythmical structure.

An effect of long-term musical training was found for reaction times, which were shorter for deviant stimuli at salient and highly expected metrical positions (strong) compared to the less expected ones (weak). This is in line with previous evidence suggesting a hierarchical processing of temporal structure in musicians (Jongsma et al., 2004; Geiser et al., 2010); contrariwise, non-musicians showed the same response speed to deviants regardless of the metrical positions at which they occurred, thus pointing to a sequential, rather than hierarchical, processing of rhythm (Cooper & Meyer, 1960; Large, 2008). While the present experiment does not allow for a clear understanding of the reasons why the effects of long-term musical training were mostly visible in reaction time data, results might suggest that training-related effects are mostly evident when considering motor related measures. One possibility accounting for this result is that throughout a long-term training musicians master precise movements' organization, high movement synchronization and rigorous control of movements; this results in structural and functional changes of musicians' motor brain areas leading to more efficient motor responses.

At the same time, according with earlier findings (see Zatorre, Chen and Penhune, 2007; Herholz and Zatorre, 2012 for reviews) the faster response speed observed might also reflect a more efficient communication between auditory-motor areas of the brain in musicians, which might explain the enhanced auditory processing and the prompt organization of a motor act required for the deviant detection.

Experiment 3, reported in Chapter 3, provided a neural investigation of the behavioral effects delineated for the experiments described above. Here, the use of combined behavioral and EEG measures allowed for an on-line monitoring of the neural responses linked to several stages of analysis of deviance perception. Specifically, with this experiment I explored: (1) the effects of long-term musical training on deviance detection as a function of rhythm and meter (strong - weak comparison); and (2) the effects of long-term musical training at early stages of deviance detection.

Behavioral results replicated the results observed in the first experiment (Experiment 1 and 2). Additionally, they indicated that musicians were faster than non-musicians when detecting deviants within rhythmical structures as compared to non-rhythmical ones. A possible explanation for the discrepancy between these results and those of the first experiment (Experiment 1) is that here deviant stimuli in the rhythmical structure appeared on strong or weak metrical positions, while in first experiment I did not manipulate their metrical position (Experiment 1), or I used only the rhythmic condition (Experiment 2). Here, the combined manipulation of rhythm and meter likely boosted the effect of musical expertise, leading to an interaction between group and temporal structure.

EEG results pointed to an effect of rhythm on the MMN regardless of musical expertise, with deviant stimuli within rhythmical structure eliciting a larger MMN amplitude compared to those inserted in non-rhythmical one. Consistent with previous evidence showing a generally enhanced neural response to rhythm elaboration regardless of musical training (Ladinig & Honing, 2009; Winkler et al., 2009), this result could be considered the neural signature of the ability of the auditory system to organize events in time (Bregman, 1994), and it further underlined that a stronger rhythm processing influences the neural responses associated to the perception of a stimulus inserted within rhythm.

Moreover, I observed larger MMN amplitude in musicians relative to non-musicians in all conditions; this reflects their enhanced auditory perception (Koelsch et al., 1999) that was visible also when non-musical auditory stimuli were processed. Additionally, effects of long-term musical training were particularly evident when focusing on the metrical aspect of the rhythmical structure wherein musicians and non-musicians displayed a different modulation of the MMN amplitude. Although this effects has not been reported by all

previous studies (see Bouwer et al., 2014 for an example), a different meter processing as function of musical expertise was reported by Geiser et al., 2010, although showing an opposite pattern of results (enhancement of the MMN amplitude for metrically incongruent conditions in musicians compared to non-musicians) with respect to the one reported here. These discrepancies in the literature might be attributed to many aspects; for example, different experimental paradigms have been adopted by previous studies. For instance, while Bouwer et al., 2014 used acoustically varied stimuli that might have influenced per se the neural response registered to an acoustically richer stimulus; Geiser et al., 2010 compared the MMN amplitude of musicians and non-musicians directly using temporal changes that might have captured musicians' attention more than non-musicians resulting in the stronger neural activation of the first relative to the second. Here, using acoustically simple stimuli and without inserting any temporal manipulation, I observed larger MMN amplitude to deviant stimuli at strong positions in musicians relative to non-musicians suggesting that stimuli coinciding with strong metrical positions are more relevant for musicians than for non-musicians.

This interpretation is linked to the secondary goal of this experiment that, as I have mentioned in the Introduction, was to disentangle the relative contribution of a dynamic attending mechanism as theorized by DAT, from a predictive process as framed within the PC theory in play during rhythm processing. While is difficult to conceive this result within the PC theory (according to which long-term musical training should minimize the prediction error, and therefore the MMN amplitude), it could be accounted for by the DAT, which posits that high attentional levels are allocated to events coinciding with highly expected points in a regular structure (e.g., metrically strong positions of a rhythm). Thereby, this result suggests that when involved in meter processing, musicians seem to use a dynamic allocation of attention to temporally salient events more than predictive mechanisms.

Finally, the analysis of the EEG data displayed an effect of long-term musical training at the early cortical stages of auditory deviance detection (Middle Latency responses - MLRs), where I found a stronger effect of rhythm for musicians as compared to non-musicians. Indeed, deviants within rhythmical structure elicited a larger Na amplitude in musicians relative to non-musicians. The present result is placed between previous observations of long-term musical training effects on the processing of auditory stimuli at (1) the brainstem

level (Musacchia et al., 2007) and at (2) the MMN level (Koelsch et al., 1999), and it further highlights that these effects at the level of the MLRs mirror a stronger influence of long-term musical training specifically linked to rhythm processing compared to non-rhythmical sequences of sounds.

Hence, although long-term musical training effects transfer also to non-musical auditory context (as demonstrated here), it seems that a relevant role is played by the early ability of the auditory system to process stimuli organized rhythmically.

EEG results further indicated that rhythm and meter processing emerged at two stages (early -MLR and late -MMN) in musicians, this likely reflects that long-term musical training influence the way and the stage at which rhythm and meter are elaborated. Indeed, while effects of rhythm influenced deviant detection already at an early level of deviance analysis (e.g., MLR); meter effects were visible only at later stages of deviance processing (e.g., MMN). I speculate that modulations of early neural components are a sign of the ease with which our auditory system encode rhythmically organized stimuli; contrariwise, the extraction and processing of a more structured organization, as the metrical structure of rhythm, implies a more complex processing and as such appear only later on the time line from deviant onset. Yet, given that no previous studies in the literature described a different modulation of these early components as function of long-term musical training, future studies may consider exploring further this aspect for example inserting deviant stimuli within different types of rhythms (simple, complex, binary or ternary). This in order to understand if neural changes derives from the processing of different rhythmical patterns, or are linked to the complexity of the structure; indeed, it could be that more complex rhythmical structures, requiring a more complex elaboration, modulate later neural responses only.

The relatively low robustness of the statistical outcomes of the electrophysiological results (see section 3.4.2) recommends caution in the interpretation of this pattern in line with predictive or attending mechanisms. Additionally, although the experimental paradigms were not designed to highlight which theory best explain rhythm processing in musicians and non-musicians, the Dynamic Attending theory seems to best describe

the MMN results obtained from musicians and non-musicians. Indeed, musicians displayed a larger MMN amplitude than non-musicians, suggesting that a higher attending level was used to process deviant stimuli. Contrariwise, considering the behavioral results alone is not clear which theory (DAT or PC) could have better explained the effects observed. This aspect need to be further investigated in future studies using a paradigm that allow to separate predictive from attending mechanisms in play while detecting a stimulus inserted within rhythmical sequences of sounds.

The first part of this thesis (Experiment 1, 2 and 3) showed a general effect of rhythm and meter in both musicians and non-musicians as visible from behavioral and neural modulations associated to deviant detection. However, musicians displayed a stronger and earlier effect of rhythm as well as a refined meter processing compared to non-musicians, as it was evident from faster reaction times, higher  $d'$  level and from modulations of early and late latency neural components. The observation that these effects were present when using simple auditory series and in a task that did not require an explicit processing of temporal structures suggests that long-term learning experience can influence cognition beyond the specific domain of training (Boot et al., 2008; 2011; Brochard et al., 2004; Klein et al., 2016; Lee & Noppeneay, 2014; Munte et al., 2002). However, the studies described above were not designed to observe if effects of long-term musical training transfer to sensory modalities other than the mostly trained one for a musicians (auditory).

The second aim of the thesis (Chapter 4) was to investigate the influence of long-term musical training on the ability to orient attention in time using auditory and visual cues. Indeed, although there is now consensus in the literature on the enhanced processing and sensorimotor synchronization to auditory rhythmical structures in musicians (see Zatorre, Chen and Penhune, 2007 and Herholz & Zatorre, 2012 for reviews), a crucial aspect that has remained largely unexplored is whether musicians handle these abilities more efficiently when orienting attention to specific points in time as compared to non-musicians. Furthermore, by means of auditory and visual cues I explored whether effect of long-term musical training are evident only when considering the auditory modality (which is strongly trained in musicians), or if these could be

generalized to vision as well. To investigate both aspects (temporal orienting ability and sensory modality effects), I used a modified version of a cross-modal temporal cueing paradigm where participants listened to a continuous auditory rhythmical sequence wherein auditory or visual cues informed them about the occurrence of a target stimulus (auditory omission) presented after short, long or mixed (either short or long) intervals of time. I explored these aspects by looking at behavioral and electroencephalographic correlates associated to the anticipation and perception of target stimuli.

As reported already in the Chapter 4, a recent study by Vassena and colleagues (2016) attempted to test this aspect using a cross-modal cueing paradigm in which subjects indicated if a target stimulus was compatible (target stimulus and cue with the same pitch or correct spatial-pitch association) with a previously presented cue, and reported an overall improved performance in musicians. Nonetheless, the paradigm adopted in their study did not directly address subjects' ability to extract relevant information from cues to orient attention in time, but rather tested predictive abilities for stimulus identity/location. Therefore, although accumulating evidences suggests a stronger cross-modal processing in musicians compared to non-musicians (Kussner et al., 2014; Vassena et al., 2016) because of their training on the integration of multimodal stimuli, no studies have examined if musicians use this enhanced cross-modal processing ability to better orient attention in time. Indeed, this aspect could offer new insight on the acknowledgment of the benefits that a long-term musical training might have also in everyday tasks.

The behavioral and neural results of Experiment 4 revealed that auditory cues were more effective in orienting attention in time. This is in line with earlier works pointing to the auditory modality as better tuned for temporal information compared to the visual one (more apt to orient attention in space) (Posner et al., 1976), and also reflects the facilitation of an unimodal processing (anticipation of a deviant in the same sensory modality of the cue) relative to the cross-modal processing of visual cues – auditory deviants. In contrast, visual cues facilitated deviance detection only when deviants were preceded by short intervals. Additionally, I observed that stimuli preceded by short intervals triggered improved behavioral performance and modulated neural responses for both musicians and non-musicians. The behavioral result is in line with



the existing literature explaining this facilitation in terms of the temporal certainty peculiar of a short interval length, which has been showed to enhance target stimuli (Niemi & Naatanen, 1981; Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010).

The ERP results indicated that short cue-target intervals prompted a higher level of expectancy for upcoming stimuli, which boosted the neural response before target onset (CNV). Furthermore, the high certainty peculiar of short intervals modulated also the neural response to deviant perception resulting in the smallest P300 amplitude compared to all other intervals. This is compatible with previous studies demonstrating that temporal orienting has larger effects for target occurring after short rather than long cue-target intervals (Miniussi et al., 1999; Correa et al., 2006; see Nobre, 2001 for a review). Here, I speculate that the short and predictable interval elicited the highest level of certainty about the target occurrence, reducing the need to update the stored information about the auditory context as well as reducing the perception of “novelty” when the deviant stimulus was detected.

It is worth noting that the neural components considered did not reveal any difference between musicians and non-musicians. In fact, effects of long-term musical training were evident only at the behavioral level and were magnified when auditory cues were used. As it was speculated before, one possibility to explain the discrepancy observed between behavior and neural results is that differences between musicians and non-musicians seen in the behavior were strongly associated to response execution, which at the neural level would be better captured in specific neural responses associate to the execution of an action (e.g., Lateralized Readiness Potential; beta power modulations). However, the paradigm used in the present study was not designed to specifically address effects related to the preparation of movements, given that it required responses to only one type of stimuli, and because only the right hand was used to respond. Future work might consider a paradigm in which subjects are required to continuously produce motor outputs, like in the form of finger tapping, synchronously to a rhythmical sequence of sounds.

A second possibility is that differences due to musical expertise might be captured in the strength of the neural coupling between auditory and motor cortices of the brain. Indeed, although it has been showed that auditory–motor brain areas are coupled in all individuals regardless of musical training, there is now

consensus in the literature that their interaction is stronger and more efficient in musicians compared to non-musicians. Future studies could directly address this aspect by investigating differences in the strength of auditory-motor coupling between musicians and non-musicians using a temporal orienting task. For this purpose, the magnetoencephalography (MEG) might be considered as a more efficient neuroimaging tool because of its higher spatial resolution with respect to EEG.

In keeping with an improved rhythm processing in musicians, a secondary goal of Experiment 4 was to examine differences between musicians and non-musicians by looking at a particular pattern of neural response associated to rhythm processing: the steady state evoked potential (SS-EP). Although previous studies looked at rhythm processing as a function of specific modulations of the SS-EPs (Nozaradan et al., 2011, 2012, 2013; Celma-Miralles et al., 2016), whether long-term musical training influences this response has remained largely unexplored. Our data showed a reduced SS-EP amplitude associated to the processing of the strong beats of rhythm in musicians compared to non-musicians. I suggest that this might reflect the ease with which musicians processed some specific components of rhythm. As a result, the stronger response of non-musicians at the frequency of the strong beats (2.17 Hz) may indicate their need for a more effortful tracking of rhythm in time (and, as such, of a synchronization to both strong beats of rhythm). Follow-up investigations should consider rhythms of increasing difficulty to examine if there is a correlation between years of musical training and the complexity level beyond which rhythm does not require an effortful processing. Also, whether these aspects are specifically linked to a specific musical training remain to be investigated comparing musicians trained on different types of instruments, and with singers (who do not undergo an extensive motor training). Indeed, it could be hypothesized that for example a drummer, who is extensively trained on the precise coordination of movements to a rhythm, displays a stronger rhythm perception compared to a singer who, although trained musically, is not trained on the synchronization of motor outputs to rhythms. This specific training results in modification of the structure and function of many brain areas, such as auditory and motor areas as well as in their stronger interaction. The higher efficiency of

these brain areas could facilitate rhythm processing, resulting, for instance, in the amplitude reduction of the steady state evoked potentials derived from rhythm processing.

In conclusion, on the basis of the evidences obtained it could be stated that effects of long-term musical training seem to transfer to non-musical contexts, as long as the auditory modality is considered, and that attentive mechanisms seem to guide an online monitoring and an efficient processing of information in the time domain.



## References

- Aagten-Murphy, D., Cappagli, G., & Burr, D. (2014). Musical training generalises across modalities and reveals efficient and adaptive mechanisms for reproducing temporal intervals. *Acta psychologica*, 147, 25-33.
- Alho, K., Grimm, S., Mateo-León, S., Costa-Faidella, J., & Escera, C. (2012). Early processing of pitch in the human auditory system. *European Journal of Neuroscience*, 36(7), 2972–2978.
- Allan, L. G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation*, 22(1), 39-58.
- Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *European Journal of Neuroscience*, 29(7), 1514-1521.
- Bendixen, A. (2014). Predictability effects in auditory scene analysis: a review. *Frontiers in neuroscience*, 8, 60.
- Berlyne, D. E. (1971). *Aesthetics and psychobiology* (Vol. 336). New York: Appleton-Century-Crofts.
- Besson, M., Faïta, F., & Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience letters*, 168(1), 101-105.
- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in psychology*, 2.
- Bogacz, R., Wagenmakers, E. J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed–accuracy tradeoff. *Trends in neurosciences*, 33(1), 10-16.
- Bolger, D., Trost, W., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta psychologica*, 142(2), 238-244.
- Bolger, D., Coull, J. T., & Schön, D. (2014). Metrical rhythm implicitly orients attention in time as indexed by improved target detection and left inferior parietal activation. *Journal of cognitive neuroscience*, 26(3), 593-605.

- Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology*, 6(2), 145-238.
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, 129(3), 387-398.
- Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology*, 2, 226.
- Borgmann, C., Roß, B., Draganova, R., & Pantev, C. (2001). Human auditory middle latency responses: influence of stimulus type and intensity. *Hearing Research*, 158(1), 57-64.
- 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.
- 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.
- Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The “ticktock” of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, 14(4), 362-366.
- Brochard, R., Dufour, A., & Despres, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain and Cognition*, 54(2), 103-109.
- Celma-Mirallas, A., de Menezes, R. F., & Toro, J. M. (2016). Look at the Beat, Feel the Meter: Top-Down Effects of Meter Induction on Auditory and Visual Modalities. *Frontiers in Human Neuroscience*, 10.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). 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.
- 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(4), 412.
- Cooper, G., & Meyer, L. B. (1963). *The rhythmic structure of music*. University of Chicago Press.

Correa, Á., & Nobre, A. C. (2008). Spatial and temporal acuity of visual perception can be enhanced selectively by attentional set. *Experimental brain research*, 189(3), 339-344.

Correa, Á., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing: A review and new evidence from event-related potentials. *Brain research*, 1076(1), 116-128.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, 18(18), 7426-7435.

Creelman, C. D. (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, 34(5), 582-593.

Daikhin, L., & Ahissar, M. (2012). Responses to deviants are modulated by subthreshold variability of the standard. *Psychophysiology*, 49(1), 31-42.

Donchin, E. (1981). Surprise!... surprise?. *Psychophysiology*, 18(5), 493-513.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, 270(5234), 305.

Elliott, M. T., Wing, A. M., & Welchman, A. E. (2014). Moving in time: Bayesian causal inference explains movement coordination to auditory beats. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1786), 20140751.

Elmer, S., Hänggi, J., Meyer, M., & Jäncke, L. (2013). Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex*, 49(10), 2812-2821.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of cognitive neuroscience*, 10(5), 590-604.

Escera, C., & Malmierca, M. S. (2014). The auditory novelty system: An attempt to integrate human and animal research, 51, 111–123.

Escera, C., Leung, S., & Grimm, S. (2014). Deviance detection based on regularity encoding along the auditory hierarchy: electrophysiological evidence in humans. *Brain topography*, 27(4), 527-538.

Fitch, W. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in systems neuroscience*, 7, 68.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, 25(4), 355-373.

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Gunter, T. C. (2004). ERPs reflect lexical identification in word fragment priming. *Journal of Cognitive Neuroscience*, 16(4), 541-552.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1456), 815-836.

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.

Gaillard, A. W., & Näätänen, R. (1976). Modality effects on the contingent negative variation in a simple reaction-time task. In *The responsive brain* (pp. 40-45).

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, 23(27), 9240-9245.

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, 45(1), 93-102.

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception—training increases hierarchical metre processing. *European Journal of Neuroscience*, 32(11), 1979-1985.

Goswami, U. (2003). How to beat dyslexia. *The Psychologist*.

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, 29(23), 7540-7548.

Griffin, I. C., Miniussi, C., & Nobre, A. C. (2002). Multiple mechanisms of selective attention: differential modulation of stimulus processing by attention to space or time. *Neuropsychologia*, 40(13), 2325-2340.

Griffin, I. C., Miniussi, C., & Nobre, A. C. (2001). Orienting attention in time. *Frontiers in Bioscience*, 6, 660-671.



Grimm, S., Escera, C., Slabu, L., & Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology*, 48(3), 377–384.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in cognitive sciences*, 11(11), 466-472.

Helmbold, N., Rammsayer, T., & Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *Journal of Individual Differences*, 26(2), 74-85.

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, 76(3), 486-502.

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.

Honing, H., Ladinig, O., Háden, G. P., & Winkler, I. (2009). Is Beat Induction Innate or Learned? *Annals of the New York Academy of Sciences*, 1169(1), 93-96.

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.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *Journal of Neuroscience*, 29(10), 3019-3025.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological review*, 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.

Jones, M. R., & McAuley, J. D. (2005). Time judgments in global temporal contexts. *Attention, Perception, & Psychophysics*, 67(3), 398-417.

Jones, M. R. (2009). Musical time. *The handbook of music psychology*, 81-92.

Jones, M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology-Human Perception and Performance*, 23(3), 693-709.

Jongsma, M. L., Desain, P., Honing, H., & RIJN, C. M. (2003). Evoked potentials to test rhythm perception theories. *Annals of the New York Academy of Sciences*, 999(1), 180-183.

Jongsma, M. L., Quiroga, R. Q., & van Rijn, C. M. (2004). Rhythmic training decreases latency-jitter of omission evoked potentials (OEPs) in humans. *Neuroscience Letters*, 355(3), 189-192.

Jongsma, M. L., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biological psychology*, 66(2), 129-152.

Klein, C., Liem, F., Hänggi, J., Elmer, S., & Jäncke, L. (2016). The “silent” imprint of musical training. *Human brain mapping*, 37(2), 536-546.

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *Journal of experimental psychology*, 51(3), 179.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6), 1309–1313.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599-605.

Kussner, M. B., Tidhar, D., Prior, H. M., & Leech-Wilkinson, D. (2014). Musicians are more consistent: Gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Frontiers in psychology*, 5.

Ladinig, O., Honing, H., Hááden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception: An Interdisciplinary Journal*, 26(4), 377-386.

Lange, K., Rösler, F., & Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: An event-related potential study. *Psychophysiology*, 40(5), 806-817.

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and cognition*, 69(1), 127-137.

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.

Lappe, C., Steinsträter, O., & Pantev, C. (2013). Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection. *Frontiers in human neuroscience*, 7, 260.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review*, 106(1), 119.

Large, E. W. (2000). On synchronizing movements to music. *Human movement science*, 19(4), 527-566.

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. *The psychology of time*, 189-232.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169(1), 46-57.

Lecaignard, F., Bertrand, O., Gimenez, G., Mattout, J., & Caclin, A. (2015). Implicit learning of predictable sound sequences modulates human brain responses at different levels of the auditory hierarchy. *Frontiers in human neuroscience*, 9.

Lee, H., & Noppeney, U. (2011). Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proceedings of the National Academy of Sciences*, 108(51), E1441-E1450.

Lee, H., & Noppeney, U. (2014). Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech, and music.

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception: An Interdisciplinary Journal*, 1(2), 229-252.

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

Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, 20(3), 1817-1829.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization* (No. 8). Oxford University Press on Demand.

Malmierca, M. S., Sanchez-Vives, M. V., Escera, C., & Bendixen, A. (2014). Neuronal adaptation, novelty detection and regularity encoding in audition. *Front Syst Neurosci*, (June), 111.

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human perception and performance*, 17(4), 1057.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychol. Rev.* 79, 487–509.

McAdam, D. W., Knott, J. R., & Rebert, C. S. (1969). Cortical slow potential changes in man related to interstimulus interval and to pre trial prediction of interstimulus interval. *Psychophysiology*, 5(4), 349-358.

Mento, G. (2013). The passive CNV: carving out the contribution of task-related processes to expectancy. *Frontiers in human neuroscience*, 7.

Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.

Miniussi, C., Wilding, E. L., Coull, J. T., & Nobre, A. C. (1999). Orienting attention in time: Modulation of brain potentials. *Brain*, 122(8), 1507-1518.

Mouraux, A., Iannetti, G. D., Colon, E., Nozaradan, S., Legrain, V., & Plaghki, L. (2011). Nociceptive steady-state evoked potentials elicited by rapid periodic thermal stimulation of cutaneous nociceptors. *Journal of Neuroscience*, 31(16), 6079-6087.

Muller-Gethmann, H., Ulrich, R., & Rinkenauer, G. (2003). Locus of the effect of temporal preparation: Evidence from the lateralized readiness potential. *Psychophysiology*, 40(4), 597-611.

Munte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, 3(6), 473-478.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, 104(40), 15894-15898.

Naatanen, R., Muraenen, V., & Merisalo, A. (1974). Timing of expectancy peak in simple reaction time situation. *Acta Psychologica*, 38(6), 461-470.

Naatanen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, 42(4), 313-329.

Naatanen, R. (1992). *Attention and brain function*. Psychology Press.

Naatanen, R., 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.

Naatanen, R., Astikainen, P., Ruusuvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: An expression of the sensory–cognitive core of cognitive processes. *Brain research reviews*, 64(1), 123-136.

Naatanen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, 42(1), 25-32.

Naatanen, R., 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.

Nazimek, J. M., Hunter, M. D., Hoskin, R., Wilkinson, I., & Woodruff, P. W. (2013). Neural basis of auditory expectation within temporal cortex. *Neuropsychologia*, 51(11), 2245-2250.

Niemi, P., & Naatanen, R. (1981). Foreperiod and simple reaction time. *Psychological bulletin*, 89(1), 133.

Nobre, A. C. (2001). Orienting attention to instants in time. *Neuropsychologia*, 39(12), 1317-1328.

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Current opinion in neurobiology*, 17(4), 465-470.

Nobre, A. C. K., & Rohenkohl, G. (2014). Time for the fourth dimension in attention. In *The Oxford handbook of attention*.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, 31(28), 10234-10240.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, 32(49), 17572-17581.

Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2013). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex*, 25(3), 736-747.

Ono, K., Altmann, C. F., Matsushashi, M., Mima, T., & Fukuyama, H. (2015). Neural correlates of perceptual grouping effects in the processing of sound omission by musicians and nonmusicians. *Hearing research*, 319, 25-31.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811-814.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, 12(1), 169-174.

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

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.

Patston, L. L., Hogg, S. L., & Tippett, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality*, 12(3), 262-272.

Peter, V., McArthur, G., & Thompson, W. F. (2010). Effect of deviance direction and calculation method on duration and frequency mismatch negativity (MMN). *Neuroscience letters*, 482(1), 71-75.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological review*, 78(5), 391.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychological review*, 83(2), 157.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, 32(1), 3-25.

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.

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception: An Interdisciplinary Journal*, 24(1), 37-48.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1).

Regan, D. (1989). *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine*.

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic bulletin & review*, 12(6), 969-992.

Ritter, W., & Vaughan, H. G. (1969). Averaged evoked responses in vigilance and discrimination: a reassessment. *Science*, 164(3877), 326-328.

Rockstroh, B. (1982). *Slow brain potentials and behavior*. Urban & Schwarzenberg.

Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS One*, 6(1), e14620.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83(2), 164-175.

Russeler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Münte, T. F. (2001). Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neuroscience letters*, 308(1), 33-36.

Sanders, A. F. (1977). Structural and functional aspects of the reaction process. *Attention and performance VI*, 3-25.

Schlaug, G. (2001). The brain of musicians. *Annals of the New York Academy of Sciences*, 930(1), 281-299.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature neuroscience*, 5(7), 688-694.

Schroger, E., & Wolff, C. (1996). Mismatch response of the human brain to changes in sound location. *Neuroreport*, 7(18), 3005-3008.

Schroger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience*, 8(6), 527-539.

Schroger, E. (1998). Measurement and interpretation of the mismatch negativity. *Behavior Research Methods, Instruments, & Computers*, 30(1), 131-145.

Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *European Journal of Neuroscience*, 33(1), 189-196.

Schwartz, 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.

Schwartz, M., & Kotz, S. A. (2015). The timing of regular sequences: production, perception, and covariation. *Journal of cognitive neuroscience*.

Slabu, L., Escera, C., Grimm, S., & Costa-Faidella, J. (2010). Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials. *European Journal of Neuroscience*, 32(5), 859-865.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, 31(1), 137-149.

Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. *Journal of Neuroscience*, 33(3), 1282-1290.

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hearing research*, 261(1), 22-29.

Tecce, J. J. (1972). Contingent negative variation (CNV) and psychological processes in man. *Psychological bulletin*, 77(2), 73.

Tervaniemi, M., Castaneda, A., Knoll, M., & Uther, M. (2006). Sound processing in amateur musicians and nonmusicians: event-related potential and behavioral indices. *Neuroreport*, 17(11), 1225-1228.

Thaut, M. H., Trimarchi, P. D., & Parsons, L. M. (2014). Human brain basis of musical rhythm perception: common and distinct neural substrates for meter, tempo, and pattern. *Brain sciences*, 4(2), 428-452.

Tierney, A., & Kraus, N. (2013). Neural responses to sounds presented on and off the beat of ecologically valid music. *Frontiers in systems neuroscience*, 7.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs: General and Applied*, 77(13), 1.

Van Zuijen, T. L., Sussman, E., Winkler, I., Naatanen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. *Journal of cognitive neuroscience*, 16(2), 331-338.

Vassena, E., Kochman, K., Latomme, J., & Verguts, T. (2016). Unimodal and cross-modal prediction is enhanced in musicians. *Scientific reports*, 6, 25225.



Vuust, P., Pallesen, K. J., Bailey, C., van Zuijlen, 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.

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.

Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., & Tervaniemi, M. (2012). The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia*, 50(7), 1432-1443.

Vuust, P., & Witek, M. A. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Frontiers in psychology*, 5, 1111.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, 20(3), 188-196.

Wallin, N. L., & Merker, B. (2001). *The origins of music*. MIT press.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature*, 203(4943), 380-384.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *The Quarterly Journal of Experimental Psychology*, 61(4), 569-587.

Weinbach, N., & Henik, A. (2012). The relationship between alertness and executive control. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1530.

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*, 106(7), 2468-2471.

Woodrow, H. (1909). *A Quantitative Study of Rhythm: The Effect of Variations in Intensity, Rate and Duration*. (No. 14). Science Press.

Woodrow, H. (1914). The measurement of attention. *The Psychological Monographs*, 17(5), i.

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.

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PloS one*, 9(6), e99868.



# Appendix

## Experiment 1 and 2

### Notes

1. For completeness, here I report the results of the post-hoc analyses on the Deviant Type x Temporal Structure x Group significant interaction not reported in the main text.

I conducted two post-hoc ANOVAs separately for Frequency and Temporal deviants with Temporal Structures as within-subject factor and Group as between-subject factor. The ANOVA for the Frequency deviants did not yield any significant result (Temporal Structure:  $F(1, 26) = 1.30$ ,  $p = .265$ ,  $\eta_p^2 = .04$ ; Temporal Structure x Group:  $F(1, 26) = 1.52$ ,  $p = .229$ ,  $\eta_p^2 = .05$ ; Group:  $F(1, 26) = .97$ ,  $p = .333$ ,  $\eta_p^2 = .03$ ). The ANOVA on the Temporal deviants indicated only a significant main effect of Temporal Structure:  $F(1, 26) = 13.9$ ,  $p = .001$ ,  $\eta_p^2 = .35$ ; Temporal Structure x Group:  $F(1, 26) = 1.45$ ,  $p = .239$ ,  $\eta_p^2 = .05$  and the main effect of Group  $F(1, 26) = 2.24$ ,  $p = .147$ ,  $\eta_p^2 = .07$ , were not statistically significant.

The ANOVAs conducted separately for Rhythmic and Non-rhythmic series indicated no significant effect for the Rhythmic condition (Deviant Type:  $F(1, 26) = .24$ ,  $p = .624$ ,  $\eta_p^2 = .009$ ; Deviant Type x Group:  $F(1, 26) = .44$ ,  $p = .512$ ,  $\eta_p^2 = .01$ ; Group:  $F(1, 26) = 2.9$ ,  $p = .095$ ,  $\eta_p^2 = .103$ ). In the Non-rhythmic condition there was a significant main effect of Deviant Type:  $F(1, 26) = 8.53$ ,  $p = .007$ ,  $\eta_p^2 = .24$ . The interaction Deviant Type x Group ( $F(1, 26) = 3.5$ ,  $p = .073$ ,  $\eta_p^2 = .11$ ) and the main effect of Group ( $F(1, 26) = .96$ ,  $p = .335$ ,  $\eta_p^2 = .03$ ) were not significant.

Finally, the ANOVAs conducted separately for Musicians and Non-musicians showed a significant effect of Temporal Structure for musicians:  $F(1, 13) = 8.64$ ,  $p = .011$ ,  $\eta_p^2 = .39$ ). No other significant effects were observed (Deviant Type:  $F(1, 13) = .88$ ,  $p = .363$ ,  $\eta_p^2 = .06$ ; Deviant Type x Temporal Structure:  $F(1, 13) = .005$ ,  $p = .947$ ,  $\eta_p^2 = .001$ ). The analysis on the non-musicians data revealed a significant interaction Deviant Type x Temporal Structure:  $F(1, 13) = 7.88$ ,  $p = .015$ ,  $\eta_p^2 = .37$ . Specifically, non-musicians adopted a more conservative criterion for temporal deviants in the non-rhythmical structures than in the rhythmical one ( $t = -3.59$ ,  $p = .003$ ). No significant difference emerged for the frequency deviants ( $t = .049$ ,  $p = .962$ ).

2. Participants' musical competence was tested via the Musical Ear Test (MET) (Wallentin et al., 2010). The test consists of 208 short musical phrases (104 melodic phrases played with sampled piano sounds, and 104

rhythmical phrases played with wood block sound). Musical phrases were presented in pair to each subject. Participants' task was to listen carefully to the musical phrases and to judge their identity by crossing one out of two boxes on a response sheet. The different melodies in the melodic trials contained one pitch violation. The difference in the rhythmical trials was prompted by a rhythmic change. Participants did not receive any feedback during the test. The MET lasted for 18 minutes and was administered with a laptop computer. Test scores analysis indicated that musicians responded overall better than non-musicians in both Experiment 1 and 2.

3. In Experiment 2, metrical positions were also associated with different acoustic contexts. Namely, longer silences always followed strong positions but preceded weak ones. Thus, this aspect might have an additional role in the present results (perhaps through masking effects).

4. As stated in the Introduction, due to their nature (which implied a temporal change in the auditory sequence), temporal deviants may have likely evoked an explicit form of rhythm processing, although not explicitly required by the task. Crucially, this aspect did not interact with musical expertise.

#### Experiment 4

##### ***Reaction Times***

The ANOVA indicated a significant main effect of Cue Target Interval ( $F(3, 66) = 84.919, p < .001, \eta_p^2 = .794$ ), with fastest reaction times for deviants preceded by short intervals compared to all other intervals. RTs to deviants preceded by mixed short interval were slowest than all other cue-target intervals. The main effect of Cue Modality was also significant ( $F(1, 22) = 10.642, p = .004, \eta_p^2 = .326$ ), due to faster RTs for auditory relative to visual cues. There was a significant effect of Group ( $F(1, 22) = 5.507, p = .028, \eta_p^2 = .200$ ), with faster RTs in musicians compared to non-musicians in all experimental conditions. The ANOVA showed a marginally significant interaction Cue Modality x Group ( $F(1, 22) = 3.648, p = .069, \eta_p^2 = .142$ ), a significant

interaction Cue Target Interval x Cue Modality ( $F(3, 66) = 8.534, p < .001, \eta_p^2 = .279$ ), and a marginally significant interaction Cue Target Interval x Cue Modality X Group ( $F(3, 66) = 2.517, p = .066, \eta_p^2 = .103$ ).

Post-hoc analyses (ANOVAs) on the Cue Target Interval x Cue Modality X Group interaction were conducted to investigate further effects of long-term musical training, cue modality and interval's length on the response speed associated to deviants detection. We performed two ANOVAs, with the same factors described above, separately for each level of Cue Modality (auditory, visual).

The first ANOVA for auditory cues, displayed a main effect of Cue Target Interval ( $F(3, 66) = 58.403, p < .001, \eta_p^2 = .726$ ), due to faster RTs for deviants anticipated by predictable intervals (equal RTs between predictable conditions,  $p > .515$ ) compared to mixed ones. RTs to deviants preceded by mixed short interval were slowest than all other cue-target intervals. The interaction Cue Target Interval X Group ( $F(3, 66) = 2.916, p = .041, \eta_p^2 = .117$ ) was significant. Finally, the ANOVA resulted in a main effect of Group ( $F(1, 22) = 7.150, p = .014, \eta_p^2 = .245$ ) showing faster RTs for musicians compared to non-musicians. Follow-up pairwise comparisons (t-tests) on the Cue Target Interval X Group interaction, produced faster RTs in musicians than non-musicians for all cue target intervals (short:  $t(22) = -2.124, p = .045$ ; long:  $t(22) = -2.949, p = .007$ ; mixed long:  $t(22) = -2.987, p = .007$ ) except for the mixed short interval that prompts the same RTs across groups ( $t(22) = -1.392, p = .178$ ).

The second ANOVA for visual cues, resulted in a significant main effect of Cue Target Interval ( $F(3, 66) = 64.248, p < .001, \eta_p^2 = .745$ ), revealing faster RTs for deviants preceded by short intervals compared to all other intervals. RTs to deviants preceded by mixed short interval were slowest than all other cue-target intervals. Moreover, long and mixed long intervals resulted in the same RTs. No other significant results were observed (all  $ps > .109$ ).

To explore effects of long-term musical training and cue modality as function of each interval's length on the response speed associated to deviants' detection, I carried out four additional ANOVAs with Cue Modality (auditory, visual) as within-subjects factor, and Group (musicians, non-musicians) as between subjects factor, separately for each level of Cue Target Interval (short, long, mixed short, mixed long).

The first ANOVA on the RTs for deviants preceded by short interval showed a main effect of Cue Modality ( $F(1, 22) = 8.509, p = .008, \eta_p^2 = .279$ ), due to faster RTs for deviants preceded by auditory cues as compared to visual ones. The interaction Cue Modality X Group ( $F(1, 22) = 4.413, p = .047, \eta_p^2 = .167$ ) was also significant. The main effect of Group was not significant ( $p = .092$ ). Follow-up pairwise comparisons (t-tests) on the Cue Modality X Group interaction, revealed a significant difference between musicians and non-musicians only for deviants preceded by auditory cues, because of faster RTs for musicians than non-musicians ( $t(22) = -2.124, p = .045$ ). Moreover, while musicians were faster in deviants detection anticipated by auditory cues compared to visual ones ( $t(22) = -5.905, p < .001$ ), non-musicians did not show such differentiation ( $p = .661, t = .451$ ).

The second ANOVA on the RTs for deviants preceded by long interval showed a main effect of Cue Modality ( $F(1, 22) = 15.350, p = .001, \eta_p^2 = .411$ ), due to faster RTs for deviants preceded by auditory cues as compared to visual ones. The interaction Cue Modality X Group was marginally significant ( $F(1, 22) = 4.105, p = .055, \eta_p^2 = .157$ ). The main effect of Group was also significant ( $F(1, 22) = 6.771, p = .016, \eta_p^2 = .235$ ), showing faster RTs for musicians compared to non-musicians. I analyzed the Cue Modality X Group interaction by means of follow-up pairwise comparisons (t-tests), results displayed a significant difference between musicians and non-musicians only for deviants preceded by auditory cues, because of faster RTs for musicians than non-musicians ( $t(22) = -2.949; p = .007$ ). Additionally, while musicians were faster in deviants detection anticipated by auditory cues compared to visual ones ( $t(22) = 4.910, p < .001$ ), non-musicians did not show such differentiation ( $t(22) = 1.188, p = .260$ ).

The third ANOVA on the RTs for deviants preceded by mixed short interval did not show any significant effect (all  $ps > .107$ ).

Finally, the last ANOVA on the RTs for deviants preceded by mixed long interval yielded a main effect of Cue Modality ( $F(1, 22) = 5.279, p = .031, \eta_p^2 = .194$ ), due to faster RTs for deviants preceded by auditory cues as compared to visual ones. The main effect of Group was also significant ( $F(1, 22) = 7.073, p = .014, \eta_p^2 = .243$ ), because of faster RTs for musicians compared to non-musicians. The interaction Cue Modality X Group was not significant ( $F(1, 22) = 1.188, p = .288, \eta_p^2 = .051$ ).

### ***Accuracy***

The ANOVA indicated a significant main effect of Cue Target Interval ( $F(3, 66) = 21.540, p < .001, \eta_p^2 = .495$ ), due to higher percentage of correct detections for deviants anticipated by mixed intervals' lengths compared to predictable ones. I found a statistically significant difference in the accuracy between mixed intervals, due to higher percentage of correct detection for deviants anticipated by mixed short interval as compared to long ones ( $p = .004$ ). Furthermore, I did not find statistical significant differences in the accuracy to deviants preceded by predictable intervals (short, long) (all  $ps > .1$ ). The ANOVA did not show any other significant effect (all  $ps > .401$ ).





