This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Neural correlates of beat and metre perception:
the role of the inferior frontal gyrus

Sujin Hong

A thesis submitted for the degree of Doctor of Philosophy
The University of Edinburgh
December 2014
Declaration

This PhD research was supported by Marie Curie Initial training Network Fellowship, “Europe, Brain, and Music: New Perspectives for stimulating cognitive and sensory processes (‘EBRAMUS’, EC Grant Agreement No: 238157)” from 2010 to 2013.

I declare that this thesis submitted to the University of Edinburgh is entirely my own work and has not been submitted for other degree or professional qualification.

Sujin Hong
Acknowledgements

Most of all, I would like to thank my two supervisors, Dr. Katie Overy and Prof. Neil Roberts, for their invaluable advice and encouragement for my PhD research in Edinburgh over the last four years. I thank Dr. Alexa M. Morcom for her precise and helpful advice on fMRI studies.

I am grateful to all staffs at the Clinical Research Imaging Centre (CRIC) of the University of Edinburgh who helped me to conduct the fMRI experiments. In particular, I thank all the radiographers who kindly assisted with the MR scanning, and Julian Sparrow for his work to transfer and anonymise the MR images. I thank Dr. Calum Gray and Dr. Tom MacGillivray for accommodating my use of the image analysis laboratory at CRIC whenever possible, and thank Dr. Martin Connell for your kindness during my working in the image analysis laboratory. I thank Paul Kennedy, for your time and assistance through numerous pilot sessions and in the use of Presentation software for the fMRI experiments. I also thank all the volunteers who participated in the two fMRI studies.

I thank Prof. Sonja A Kotz and Dr. Daniela Sammler for hosting me at the University of Manchester and Max Plank Institute in Leipzig during my research exchange within the Europe Brain and Music (EBRAMUS) project. I also thank all colleagues of EBRAMUS - it was a unique group journey for music and brain research, where I have enjoyed rich and diverse scientific discussions.

I would like to thank Dr. Armin Heinecke and Prof. Rainer Goebel for their kind technical help and advice regarding BrainVoyager software. In particular, Dr. Armin Heinecke provided advice on the imaging processing method to import a maximum probability map from SPM to BrainVoyager, which is described in Chapter 3.
I also thank Emma Moore and Jenny Triggs for their proofreading of the text of my thesis and advice on use of English.

I would like to thank all my PhD colleagues and office mates in Edinburgh, Ana, Kirsteen, Rebecca, Eric, Paul, Ellen, Karen, George, Clare, Jill, and Emma, who have shared their wisdom and positive thinking from trivial to important issues of my research life. I also thank my previous colleagues and supervisors at the Neuroscience Research Institute in South Korea, who shared my early scientific experience on music and brain research.

Finally, I dedicate this thesis to my family, who may have shared my PhD research done joyfully in the beginning, and probably patiently at the end.
Abstract

Temporal regularity and metrical organisation are important factors in beat and metre perception. The current thesis aims to investigate neural correlates of beat and metre perception in healthy non-musician volunteers, using functional magnetic resonance imaging (fMRI). In particular, the thesis focuses on determining the role of the Inferior Frontal Gyrus (IFG, in particular BA44/45) in beat and metre perception. The IFG has been proposed to be involved in higher order cognitive processes during various temporal sequencing, such as speech, movement, and music. Previous studies have shown that the temporal processing of rhythm activates auditory and sensorimotor areas, but the role of the IFG in rhythm perception has not yet been fully understood.

Study 1 investigated beat perception in complex rhythms, in which the addition of volume accents either enhanced or weakened the beat perception, resulting in Unaccented, Beat Accented or Non-Beat Accented rhythms. Participants (N=14; 6 males) listened to rhythm pairs across all three conditions, and judged whether each rhythm pair was the same or different. Results showed that left IFG activation (BA44) was significantly greater during the Non-Beat Accented condition compared to Beat Accented condition, whereas the right IFG activation showed no significant difference between the two conditions. Study 2 investigated metre perception of a series of isochronous sequences, of which metrical organisation was grouped by 2/4 (C2), 3/4 (C3), or 4/4 (C4) using pitch accents, or remained without metrical grouping (or 1/4, C1). The same participants (N=15; 6 males) listened to the stimuli and indicated the perceived metrical grouping level. Results showed that the activation of bilateral IFG parametrically increased from C2 to C3 to C4. Interestingly, the activation was found to
be significantly greater in C1 relative to C2, suggesting that involuntary subjective in C1 may increase the brain response.

Converging results from both Study 1 and Study 2 demonstrated, firstly, that the bilateral IFG is involved in rhythm perception in addition to the auditory and sensorimotor areas, including primary and secondary auditory areas, supplementary motor areas, premotor areas, insula, and basal ganglia. Secondly, the left IFG (BA44) in particular was significantly modulated by the rhythmic complexity relating to both temporal regularity and metrical organisation, while showing the suppression during the Beat Accented rhythm condition of Study 1 and the binary pattern (C2) of Study 2. This thesis argues that the left IFG (BA44) may have the role the higher order cognitive processing, such as attention and prediction, in the perception of hierarchical structures in metric rhythms.
# Table of Contents

Declaration .......................................................................................................................... i
Acknowledgments .............................................................................................................. ii
Abstract ............................................................................................................................... iv
Table of Contents ............................................................................................................... vi
List of Figures ...................................................................................................................... viii
List of Tables ...................................................................................................................... ix
Abbreviations .................................................................................................................... x

Chapter 1. Introduction ...................................................................................................... 1

2.1. Theoretical understanding of rhythm processing ...................................................... 4
  2.1.1. Defining musical terms of rhythm, pulse, beat, and metre perception ........ 4
  2.1.2. Rhythmic grouping ......................................................................................... 8
  2.1.3. Attention to rhythm ....................................................................................... 13
  2.1.4. Prediction in rhythm perception ..................................................................... 20
  2.1.5. Summary ....................................................................................................... 25

2.2. Behavioural studies of rhythm processing ............................................................... 25
  2.2.1. Developmental aspects of rhythm processing .............................................. 26
  2.2.2. Positive effects of metric rhythms on cognitive processes ......................... 33
  2.2.3. Summary ....................................................................................................... 35

2.3. Neuroimaging studies of rhythm processing .......................................................... 36
  2.3.1. Auditory and sensorimotor areas in rhythm perception ................................. 36
    2.3.1.1. General introduction to auditory and sensorimotor areas ...................... 37
    2.3.1.2. Rhythmic complexity in auditory and sensorimotor areas .................... 43
    2.3.1.3. Sensorimotor areas in speech and action observation ......................... 47
  2.3.2. Inferior frontal gyrus in rhythm perception ..................................................... 50
    2.3.2.1. General introduction to IFG ................................................................... 51
    2.3.2.2. IFG in musical rhythm ......................................................................... 59
    2.3.2.3. Rhythmic complexity in IFG ............................................................... 63
  2.3.3. Summary ....................................................................................................... 70

2.4. Aims of this thesis ...................................................................................................... 72
TABLE OF CONTENTS

Chapter 3. Methodology .................................................................................................................. 73
  3.1. Introduction to functional Magnetic Resonance Imaging (fMRI) ........................................... 73
  3.2. Experimental designs of fMRI .............................................................................................. 78
  3.3. Region of interest (ROI) analysis ........................................................................................... 80
  3.4. Developing a method for ROI analysis ................................................................................... 84

Chapter 4. Beat perception in complex rhythms ........................................................................... 90
  4.1. Introduction .......................................................................................................................... 90
  4.2. Methods ................................................................................................................................ 94
  4.3. Results .................................................................................................................................. 107
  4.4. Discussion ............................................................................................................................ 115

Chapter 5. Metre perception in isochronous sequences ................................................................. 124
  5.1. Introduction ........................................................................................................................ 124
  5.2. Methods ................................................................................................................................ 128
  5.3. Results .................................................................................................................................. 136
  5.4. Discussion ............................................................................................................................ 141

Chapter 6. General discussion .................................................................................................... 146
  6.1. Evidence of IFG activation in beat and metre perception ..................................................... 146
  6.2. The role of IFG in beat and metre perception ........................................................................ 155
  6.3. Limitations of the current thesis ........................................................................................... 159
  6.4. Implications and future directions ......................................................................................... 161
  6.5. Conclusions .......................................................................................................................... 163

References ..................................................................................................................................... 165
Appendices ...................................................................................................................................... 184
List of Figures

Figure 2.1. Schematic of rhythm, pulse, beat, and metre perception ......................... 8
Figure 2.2. Schematic of non-isochronous subdivisions of beat ............................. 12
Figure 3.1. Image preprocessing of Colin27T1 and MPM templates ....................... 87
Figure 3.2. Segmentation of MPM-based ROIs in bilateral IFG (BA44/45) .............. 88
Figure 4.1. Music notations and the corresponding sound waves of Unaccented, Beat
   Accented, and Non-beat Accented rhythm conditions .................................... 97
Figure 4.2. Music notations of all rhythmic stimuli of Unaccented, Beat Accented,
   and Non-beat Accented rhythm conditions .................................................... 99
Figure 4.3. Schematic of fMRI scan ..................................................................... 101
Figure 4.4. Behavioural results in a discrimination task during fMRI scan ............ 107
Figure 4.5. Brain activation during all rhythm conditions of Unaccented, Beat Accented, Non-beat Accented rhythm conditions ......................................................... 108
Figure 4.6. The mean percent change in BOLD response in the bilateral IFG
   activation during each rhythm condition .......................................................... 109
Figure 4.7. Significantly greater activation in the left STG activation during Non-beat
   Accented rhythm condition compared to Unaccented rhythm conditions .......... 110
Figure 4.8. The left lateralised brain activation during Non-beat Accented rhythm
   condition compared to Beat Accented rhythm condition .................................. 111
Figure 4.9. The left IFG (BA44) activation defined ROI analysis during Non-beat
   Accented rhythm condition compared to Beat Accented rhythm condition .......... 112
Figure 4.10. The mean percent change in BOLD response in the left IFG (BA44)
   activation ....................................................................................................... 113
Figure 4.11. The mean percent change in BOLD response in the left STG and left
   IFG (BA44) during each rhythm condition of the same and different rhythm
   pairs .................................................................................................................. 114
Figure 5.1. Schematic of rhythmic stimuli ............................................................. 131
Figure 5.2. Schematic of fMRI scan ..................................................................... 132
Figure 5.3. Brain activation that were parametrically modulated from C1 to C2 to C3
   to C4 .............................................................................................................. 137
Figure 5.4. The mean percent change in BOLD response in the activation areas that were parametrically modulated from C1 to C2 to C3 to C4 ........................................ 138
Figure 5.5. The bilateral IFG (BA44/45) activation in ROI analysis ...................... 139
Figure 5.6. The mean percent change in BOLD response in bilateral IFG activation ................................................................................................................................. 140

List of Tables

Table 4.1. Participant demographics ........................................................................... 95
Table 4.2. Stereotaxic coordinates of brain regions defined by both whole brain and ROI analyses ................................................................................................................. 115
Table 5.1. Participant demographics ........................................................................... 129
Table 5.2. Stereotaxic coordinates of brain regions defined by both whole brain and ROI analyses ................................................................................................................. 140
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Auditory cortex</td>
</tr>
<tr>
<td>STG</td>
<td>Superior temporal gyrus</td>
</tr>
<tr>
<td>STS</td>
<td>Superior temporal sulcus</td>
</tr>
<tr>
<td>PMC</td>
<td>Premotor cortex</td>
</tr>
<tr>
<td>SMA</td>
<td>Supplementary motor areas</td>
</tr>
<tr>
<td>BG</td>
<td>Basal Ganglia</td>
</tr>
<tr>
<td>IFG</td>
<td>Inferior frontal gyrus</td>
</tr>
<tr>
<td>MFG</td>
<td>Middle frontal gyrus</td>
</tr>
<tr>
<td>fMRI</td>
<td>functional Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>BOLD</td>
<td>Blood oxygenation level-dependent</td>
</tr>
<tr>
<td>MEG</td>
<td>Magnetoencephalographic</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
</tr>
<tr>
<td>rTMS</td>
<td>repetitive Transcranial Magnetic Stimulation</td>
</tr>
<tr>
<td>MPM</td>
<td>Maximum probability map</td>
</tr>
<tr>
<td>PFC</td>
<td>Prefrontal cortex</td>
</tr>
<tr>
<td>VLPFC</td>
<td>Ventral lateral prefrontal cortex</td>
</tr>
<tr>
<td>DLPFC</td>
<td>Dorsal lateral prefrontal cortex</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>EPI</td>
<td>Echo-planar images</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>VBM</td>
<td>voxel-based morphometry</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction

Rhythm perception is an ability to process temporal information, similar to recognising faces or learning languages. In rhythm research, it has been shown that the ability to perceive rhythmic patterns develops at an early age (Demany, McKenzie, & Vurpillot, 1977; Chang & Trehub, 1977; Hanon & Johnson, 2005), and can be further developed during the lifetime (e.g., through music training or exposure to a certain culture). In addition, rhythmic processing has been found to be related to the improved or facilitated behavioural performance during perception, cognition, and motor tasks in both music and non-music domains. For example, perceiving artificial speech syllables was significantly improved when metrical temporal patterns were offered during the listening (Cason & Schöen, 2012), and the rhythm reproduction task was significantly improved when the rhythms were metric relative to non-metric patterns (Essens & Povel, 1985). In clinical populations, rhythmic components have been utilised in a therapeutic method. For example, stuttering was improved when metronome cues were heard by stutters (Azrin, Jones, & Flye, 1968), and the melodic intonation therapy (Albert, Sparks, & Helm, 1973) improved the speech production of aphasic patients.

Rhythm perception in music often accompanied by the gestural movements of tapping, clapping or nodding, of which rhythm perception becomes explicitly visible through such rhythmic behaviors. Numerous brain researches focus on neural mechanisms of this rhythm-induced movement. On the other hand, the neural correlates relating to the cognitive processes of rhythm perception received relatively less attention, although perceiving beat and organising metrical grouping of rhythms can often drive the movements. In addition, neural correlates during the perception of even simple rhythmic patterns are still largely unknown.
This thesis defines, therefore, firstly what the key factors of rhythmic structures relating to the above-described rhythm abilities are, and secondly, how these key factors of rhythmic structures are perceived in the brain. Discussion on rhythm production or movement is beyond the scope of this thesis, but its literature is partly discussed to understand the neural correlates in rhythm perception. The rhythms that this thesis discusses are also the rhythmic structures of western music\(^1\), of which temporal structures are largely dependent on both temporal regularity, such as a steady beat, and metrical organisation, such as metre. The temporal structures of beat and metre are easily identified in everyday music listening, thus the rhythm perception that this thesis refers to is the ability to perceive the temporal patterns of beat and metre underlying a series of rhythm stimuli.

Rhythm processing is highly dependent on a time series of sound information, thus perceptual expectation over time can be strongly involved in the cognitive processes underlying rhythm perception. For example, when listeners perceive salient temporal regularity of beat and metre during rhythm perception, the temporal regularity may lead the listeners to predict the upcoming rhythms. This may consequently introduce the listeners a particular way to listen to rhythms and music, suggesting that hierarchical structures in metrical rhythm sequences involve higher order cognitive processing.

IFG has been suggested to be involved during, not only rhythm tasks (Platel, Price, & Baron et al, 1997; Grahn & Brett, 2007), but also during rule-based processes, such as syntactic language processing (Obleser, Meyer, & Friederici, 2011), harmonic sequencing in music (Maess, Koelsch, Grunter, & Friederici, 2001; Tillmann, Janata, & Bharuch, 2003), and hierarchical organisation of action plans (Koechlin & Jubault, 2006).

---

\(^1\) Music discussed in the current thesis is specified as western tonal music.
CHAPTER 1. Introduction

The current thesis focuses on the cognitive processing of rhythm perception in the brain, which may involve IFG relating to temporal processing of sequencing.

The current thesis begins with examining theoretical work on rhythm perception. In Chapter 2, the musical terms that are used in this thesis are defined, and theoretical approaches to understanding cognitive processes underlying rhythm perception, including grouping, attention, and prediction, are discussed. The review of behavioural studies is concerned with rhythmic aspect involved in improving behavioural task performance, where the conceptual design for the fMRI studies in the thesis is developed. Following this, the review of neuroimaging studies of rhythm perception overviews neural correlates of rhythm processing in general, and supplement neural understanding regarding the cognitive processes of temporal regularity and metrical organisation in rhythm perception. At the end of Chapter 2, the neural hypotheses of the thesis are set up.

Chapter 3 explains the methodology of fMRI used, in that a particular attention is given to region of interest (ROI) analysis to investigate the IFG region. Chapter 4 and Chapter 5 report on two fMRI studies: Chapter 4 discusses beat perception in complex rhythms, and Chapter 5 discusses metrical organisation in isochronous sequences. Chapter 6 converges the research findings from the two fMRI studies and discusses the role of IFG in beat and metre perception, followed by the implications and future directions of this thesis.
Chapter 2. Literature review

Chapter 2 is organised into three parts: theoretical, behavioural, and neuroimaging studies. Firstly, a theoretical approach examines the definition of musical terms and the potential cognitive processes involved in rhythm perception. Secondly, behavioural studies are discussed from two perspectives: the developmental aspects and positive functions of rhythmic perception on cognitive processes. Thirdly, a review of neuroimaging studies examines the neural correlates of beat and metre perception, specifically the auditory and sensorimotor areas and the IFG. Finally, neural hypotheses on beat and metre perception for this thesis are outlined.

2.1. Theoretical understanding of rhythm perception

This section begins by defining musical terms relating to rhythmic structures, which are used in this thesis, followed by theoretical consideration to understand potential cognitive processes involved in rhythm perception. As rhythmic structures between beat and metre are in hierarchical relations, the cognitive processes during rhythm perception may include higher order cognitive processing in hierarchical structures of metric rhythms as well. This section discusses three potential cognitive processes involved in rhythm perception, which are rhythmic grouping, attention, and prediction.

2.1.1. Defining musical term: rhythm, pulse, beat, and metre perception

Musical terms, such as rhythm, pulse, beat, and metre, are frequently used in the empirical research of rhythm, but these terms sometimes have different meanings in different studies. For example, in one study the term ‘rhythm’ describes an increased number of subdivisions of beat compared with isochronous sequences (Bengtsson &
Ullén, 2006), whereas in another, the same term of ‘rhythm’ describes temporal regularity compared with temporal irregularity (Geiser, Zaehle, Jäncke, & Meyer, 2008). In the former definition of rhythm, rhythmic complexity concerns the increased information in rhythm perception. Meanwhile, in the latter definition, rhythmic complexity explains the increased temporal irregularity during rhythm perception, which previous studies have also described as ‘periodic or non-periodic’, ‘regular or irregular’, or ‘metric or non-metric’. It is, therefore, important to define any musical terms used in a rhythm study, particularly ‘rhythm’ and ‘rhythmic complexity’.

In music theory, Cooper and Meyer (1960) define rhythm as the process of interpretation in sequencing, suggesting that rhythm processing is a psychological and subjective mental representation. Accordingly, rhythm perception of pulses and beats can be further distinguished depending on a given rhythmic context. For example, in their defining metre it has been suggested that pulse and beat can be conceptually differentiated (Cooper & Meyer, 1960, p. 4):

“Meter is the measurement of the number of pulses between more or less regularly recurring accents. Therefore, in order for meter to exist, some of the pulses in a series must be accented – marked for consciousness – relative to others. When pulses are thus counted within a metric context, they are referred to as beats. Beats which are accented are called “strong”; those which are unaccented are called “weak”.”

The above theoretical distinction between pulse and beat by Cooper and Meyer (1960) suggests that perceiving temporal regularity can be called pulses in general, and when pulses are accented with strong and weak patterns, it can be called beats. Metre is a series of beats, thus, beat perception of binary (i.e., Strong and weak beat) and ternary (i.e., Strong, weak, and weak beat) patterns can be also regarded as metre perception. If
pulse perception does not induce perceptual accents, it does not necessarily result in metre perception.

This distinction between pulse and beat perception can be useful in particular when subjective grouping is studied in empirical research. For example, Brochard, Abecasis, and Potter et al (2003, see also the details in 2.2.1. Developmental aspects of rhythm perception) showed that subjective accenting happened during the perception of isochronous tone sequences, suggesting that attention modulated the subjective grouping. When isochronous tone sequences are heard by listeners, some may not group the tone sequences, whereas other may involuntarily group the isochronous tone sequences with various patterns of beats, such as a duple metre of ‘strong-weak (2/4)’ or a triple metre of ‘strong-weak-weak (3/4)’.

The difference between pulse and beat can be understood that beat perception can be further organised into higher grouping levels of rhythmic structures, such as metre. Huron (2006) defines metre as a repeating pattern of beats, and explains the hierarchy between beat and metre in a larger period of metre, which is called hypermeter. The subjective grouping, in which isochronous tone sequences can be heard as either pulse, without metric grouping, or beat with metric accents, can be difficult to measure precisely, as the psychological interpretation of metric accents in isochronous sequences cannot be explicitly controlled. If the isochronous tone sequences are manipulated by phenomenal metric accents\(^2\), this can make the metric accents explicit while suppressing subjective groupings.

---

\(^2\) Jackendoff & Lerdahl (1981) explain phenomenal, structure, and metrical accents in music, in that phenomenal accents are physically emphasised accents, such as pitch change, or loudness. Structure accents are made from musically meaningful phrases, and metrical accents indicate a series of beats that has strong and weak metric patterns.
Interestingly, in music theory research it has been suggested that there is no consensus by which to define rhythms (London, 2001). The rhythms in western tonal music have been suggested to be a unique temporal structure of music (London 2005), in which the definition of rhythm indicates a temporally organised structure including beat and metre, which can be comprised of either isochronous or non-isochronous temporal patterns. Theoretical music concepts about rhythms thus give a more comprehensive definition in rhythm research.

The current thesis uses the terms of pulse, beat, metre, and rhythm. The distinction between pulse and beat, which depends on perceiving metric accents, is taken from Cooper and Meyer (1960), which can be further understood by hierarchical rhythmic structures (Huron, 2006). There exists a hierarchy between beat and metre perception, where metre perception is dependent on differently accentuated metric groupings in beat perception, and the beat perception is based on pulse perception. The cognitive processes involved in rhythm perception thus rest on these hierarchical groupings among pulse, beat, and metre within rhythmic structures.

Despite the psychological interpretation or mental process as the definitions of rhythm from both music theory and empirical research imply, such as beat induction or metre perception, it can be often found that everyday music listening has phenomenal accents to facilitate beat and metre perception (e.g., popular music or children’s song). These salient and explicit phenomenal accents in beat and metre perception can also make it easier for listeners to control the unwanted variables, which can be potentially caused by subjective groupings of individual, in experiments.

In summary, the current thesis defines rhythm as a comprehensive musical term, which describes a series of complex subdivisions of beat in this thesis. Pulse perception is defined as perceiving temporal regularity in rhythms, and beat perception as inferring
strong and weak patterns of pulses. (Figure 2.1.). In the figure, both pulse and beat perceptions exist at the same level in rhythmic structures, which are to describe hierarchical rhythmic structures due to subjective groupings of strong and weak beat patterns. Metre perception is to organise a series of beats, which can have various metrical grouping patterns, such as 2/4, 3/4, or 4/4. The current thesis particularly focuses on the cognitive processes underlying beat and metre perception.

Figure 2.1. Schematic of rhythm, pulse, beat, and metre perception

2.1.2. Rhythmic grouping

As discussed in the previous section, rhythmic grouping is a mental process (Cooper & Meyer, 1960). For example, pulse and beat perception can be differentiated depending on the subjective grouping patterns on metric accents. When phenomenal accents for the salient metrical organisation of rhythms are lacking, the perceptual differences between the two can be large. Rhythmic grouping has been also discussed, which is distinguished
from metric grouping or metre by musical context (Lerdahl & Jackendoff, 1981, 1983). Lerdahl and Jackendoff (1981) suggest that grouping in rhythm combines musically meaningful units together, which can be, for example, driven by melodic phrases that can have a longer period than an underlying metrical grouping unit. In this sense, rhythmic grouping is not necessarily consistent with a metrical grouping, and the interaction between rhythmic and metric grouping (Lerdahl & Jackendoff, 1981) may happen. Similarly, Todd (1994) distinguishes metrical grouping from phrasal grouping by defining that the phrasal grouping is subjective and expressive interpretation of individuals, which can be determined by the temporal proximity of rhythmic structures.

Discussion on the differences between metric and rhythmic grouping here is to introduce the idea that the rhythmic grouping may consist of musically meaningful units, whereas metric grouping is highly dependent on hierarchically organised rhythmic structures over time, which may be a unique temporal processing of musical rhythm. For example, metric grouping providing the salient temporal accents of strong and weak beats keeps listeners informed of the repetitive temporal frames during music listening. The metric grouping may also enhance either segmenting the long phrase of rhythmic grouping into several smaller units, or combining the short phrases of rhythmic grouping by longer phrases of rhythmic grouping, which are musically meaningful units.

*Metrical organisation in beat and metre perception*

As discussed in the previous section, beat and metre exist in a hierarchical relationship, and the different processing of hierarchical structures in rhythm may be found in the perceptual and cognitive processes of beat and metre perception. For example, Drake (1998) proposed two levels of rhythm processing: a low level processing that involves
segmenting complex rhythms into groups at the level of beat, and a high level processing involving the segmentation of rhythms over a longer time span at the level of metre or the higher level of metre. The low level processing of rhythm has been suggested to be basic and universal, whereas the high level processing is considered to be an acquired and learned ability. In the subsequent experimental illustrations, the theoretical framework of the low and high level processing in rhythm perception was explained, as the time span of high level processing widened in accordance with age or musical training, and low level processing showed no significant difference by those variables (Drake, 1993). Both the theoretical framework and experimental illustrations in Drake (1998) suggest that rhythm perception can have different cognitive processes depending on the levels of rhythmic structures, and that the cognitive processing in the two different levels can be flexibly interactive. Thus, this theoretical proposal suggests that cognitive processes of metrical organisation are an acquired ability, and that neural correlates of beat and metre may be different.

More recently, Fitch (2013) proposed that pulse is music specific, whereas metre is observed in other domains, such as speech, suggesting that hierarchical metrical grouping may relate to unique hierarchical cognitive processes in human language ability. For example, the study, comparing nonhuman primates with human participants in an artificial grammar-learning task, showed that hierarchical cognitive processing was only learned by humans (Fitch & Hauser, 2004). In this experiment, two types of grammatical structures were examined, which are finite state grammar (FSG, e.g., ABABAB) and phrase structure grammar (PSG, e.g., AAABBB). It was found that humans could learn both grammar structures, whereas nonhuman primates (e.g., Tamarin) could only learn FSG, as they failed to learn the rules of the hierarchical structures of PSG, which required the ability to store the hierarchical structures of PSG for a longer time span than
FSG. Hierarchical processing relating to language ability in humans has been suggested to be unique (Fitch & Hauser, 2004), of which hierarchical processing is similar to that of musical metre (Fitch, 2013). The idea that a hierarchy in rhythmic structures is not only music specific, but can also be found in other cognitive domains, is interesting, particularly, in terms of the capacity to store information for a longer time span.

Theoretical frameworks for rhythm processing proposed by Drake (1998) and Fitch (2013) commonly suggest that rhythm perception has different two levels in a hierarchical manner, which may result in different cognitive processes. Regarding the ability to perceive temporal regularity and to organise metrical organisation in rhythmic structures, the higher order cognitive processing during hierarchical structures of metrical organisation has been suggested to be further developed by musical training (Drake, 1998), and the hierarchical cognitive processing has been suggested to be unique in human language faculty (Fitch, 2013).

*Interaction between beat and the subdivisions of beat*

A further hierarchy between beat and metre can be found further between the subdivisions of beat and beat level, in which beat perception may not always require the perception of metrical organisation. For example, metre perception is highly dependent on the perception of temporal regularity or pulse, but the sense of pulse in rhythms may not need the cognitive processing of metrical organisation (Parnutt, 1994). It will therefore be more difficult to perceive pulse or temporal regularity when the subdivisions of beats are in non-isochronous rhythmic patterns (e.g., rhythms comparing both triplet and duplet in Figure. 2.1.) than when subdivision of beats consists of duplets only. Although metrical organisation of integer units in subdivisions of beats has been
suggested to be important in perceiving temporal regularity (Essens & Povel, 1985), this particular case is based on ‘duration accents’ to induce beat perception. If phenomenon accents of the beat are added on a series of rhythms, the rhythms may induce either pulse or beat perception regardless of the combination of subdivisions of the beat. For example, duplet, triplet, quadruplet, and quintuplet can be combined together, which makes it difficult to induce beat perception due to the lack of repetition of the same subdivisions of beat, and the phenomenal accents can introduce beat perception. In this manner, phenomenal accents are needed to induce beat perception in the rhythms consisting of non-isochronous subdivisions of beat (Figure 2.2).

![Diagram of Non-Isochronous Subdivisions of Beat](image)

**Figure 2.2.** Schematic of non-isochronous subdivisions of beat

In the above example if phenomenal metrical accents were added in the beat-based, but non-isochronous subdivisions of beat rhythms, it might be easier to induce a sense of pulse. A sense of pulse in this case, however, does not necessarily induce the metrical organisation of strong and weak beats, which can be a higher rhythmic grouping interacting with an underlying metrical grouping. When isochronous sequences are heard by listeners, it has been suggested that the listeners group the isochronous sequences into a higher level of metrical grouping (Brochard, Abecasis, & Potter et al., 2003), whereas if the rhythmic complexity is increased by non-isochronous subdivisions of beat, the role of phenomenal beat accents will be important for pulse or beat perception. This suggests that rhythms consisting of non-isochronous subdivisions of beats can be useful for
optimising the role of phenomenal beat accents to perceive temporal regularity or metrical organisation between the subdivisions of beats and the beat level.

2.1.3. Attention to rhythm

Perceiving sequential information requires one to pay attention continuously. In auditory attention, perceiving sound information can often be altered by accentuation. For example, volume or pitch accents result in greater attention during rhythm perception. This section discusses attention relating to accentuation for temporal regularity and metrical organisation in rhythm perception.

Attention to temporal regularity and metrical organisation

Joint accent structures (Jones, 1987) describe time hierarchy, with the interrelationship between melodic and rhythmic accents, in that listeners selectively attend to rhythmic patterns based on temporal regularity, and their dynamic attending to rhythms can be shifted to higher levels of time hierarchies. Beyond attending rhythms of both temporal regularity and irregularity, attention to time hierarchies between beat and meter brings several considerations. Firstly, the hierarchical relations between beat and metre may require different attending modes. Secondly, attention to rhythm may depend on the subjective groupings of metrical organisation. For example, when listeners hear isochronous tone sequences, attending to the metrical groupings in the isochronous sequences can be different from an individual’s subjective metrical grouping (e.g., metrical groupings of 2 beats (2/4), 3 beats (3/4) or 4 beats (4/4) can be interpreted by individuals). Attending to the temporal regularity of beat may not vary with listeners, but
attending to different hierarchical levels of metric grouping in rhythms may much vary with the listeners.

Jones and Boltz (1989) have proposed two attending modes: ‘future-oriented attending’ and ‘analytic attending’. The future-oriented attending mode anticipates information when the information is based on high temporal coherence, and in this case, the attention induces temporal expectation. The analytic attending mode groups or counts local information when the information has low temporal coherence. The former can be understood as top-down processing, whereas the latter can suggest bottom-up processing. For example, musical rhythms often have a high level of predictability due to their regular and repetitive temporal structures, such as beat and metre. Thus in accordance with the two attending modes, metric rhythms may generally increase temporal expectation of subsequent rhythm events. On the other hand, if the predictability of rhythms becomes low due to increased temporal irregularity or low temporal coherence, attention to these rhythms may require cognitive processing of grouping or counting local rhythmic components. Thus, this theory suggests that these two different attention modes can be dynamically chosen depending on the degrees of temporal coherence within rhythmic structures.

Regarding the hierarchical levels in rhythmic structures, Drake, Jones, and Baruch (2000) suggested that the tempo and temporal hierarchy in the dynamic attending theory (DAT) can be developed by age and musical training, and that the time points in which listeners can attend the most is a referential level or tactus, which can be shifted into higher or lower levels during music listening. In this study, it was demonstrated that the referential level or tactus, which can be understood as the subjective sense of pulse, became slower rates or higher hierarchical levels depending on the increase of age or musical training in children. The slower rates of the referential level during attending
rhythms, thus, suggest a widened time span of metrical grouping in rhythm perception possibly within the future oriented attending mode.

That attending rhythms can be subjective, in terms of its time hierarchy, and vary depending on age or musical training suggests that other exogenous factors may also affect the two attending modes in rhythms. For example, different tasks requiring different cognitive strategies for rhythm perception and production can place differing demands on attending modes in rhythm processing. For example, if metric rhythms are given to listeners and the task is to judge the same or different rhythms during rhythm perception, it may require more analytic attending mode, compared with just listening to the rhythms, despite the high temporal coherence of the metric rhythms used, which would normally involve future oriented attending mode. On the other hand, if a task is to tap in synchrony with the beats in the metric rhythms, more future-oriented attending mode may be required, whilst the attention to subdivisions of beats is suppressed. If a task is to reproduce the rhythms, which requires to memorising the subdivisions of rhythms to reproduce, more analytic attending mode may be required. It is, therefore, the individual task, which determines the type of rhythm processing is another exogenous factor to determine the types of dynamic attending modes employed regardless of rhythmic structures.

Two future-oriented and analytic attending modes as well as a referential level in rhythm processing shows its flexibility depending on individuals or given conditions and suggests that attending temporal regularity or metrical organisation in rhythm can also be flexible depending on the different intensities given to beats. For example, strong and weak beats in metre perception may have the different intensities on attention as well. A recent study showed that the imagining metre produces additional neural activity in the strong beats compared with the weak beats (Nozaradan, Peretz, Missal, &
Mouraux, 2011), providing evidence of different intensities of attention between strong and weak beats. Although the metre was imagined, rather than heard in musician group in this experiment, this mechanism corresponds well with attending rhythm theories.

With respect to the intensity of attention, suppression or enhancement during attending rhythms can also be considered. Broadbent (1956) suggested the filter theory in auditory attention, in which sensory processing appreciates the information that captures attention, whilst other information that is not attended is suppressed. Although this filter theory has since been modified due to its limitation in being unable to explain the exceptional cases, such as when inattentive information can be recalled, enhanced or suppressed attention depending on the accentuation of sound information can be applied to the attending mode to perceive temporal regularity or metrical organisation. For example, attention on strong beats is enhanced, whilst the attention on weak beats is suppressed, and information on strong beats can thus be more easily perceived than that on weak beats.

Large and Jones (1999) proposed that ‘attending rhythms’ is a rhythmic behaviour to generate internal oscillation to entrain future events, and to use attentional energy to target information. According to this theory, entrainment is the process that combines both external rhythms (rhythmic structures that heard) and internal rhythms (temporal expectancy in mind). Although internal oscillation is created or altered by the external events of rhythmic structures, once listeners create the internal oscillation, it acts as a time frame in which to perceive future information, resulting in top-down processing. Similarly, Winkler, Denham, and Nelken (2009) proposed that detecting temporal regularity in auditory perception happens in the low levels of the auditory system, where the detected or perceived temporal regularity has been suggested to have a role in information prediction as well as parsing complex sequences. The role of
temporal regularity to group or parse information suggested by Winkler et al. is in line with the referential level observed during hierarchical processing in rhythm perception (Drake, 1998). Thus, the attention to temporal regularity may contribute to the appreciation not only of the accented information of strong beats or temporal regularity, but also the unaccented information of weak beats or subdivisions of temporal regularity.

Taken together, the relationship between attending mode and the intensity of attention, particularly to temporal regularity or metrical organisation in rhythms, can be modulated by exogenous factors, such as rhythmic structures, the music expertise of individuals, or the specific tasks given. This suggests that in terms of the temporal aspects of attention, the future-oriented attending mode for top-down processing interacts with the analytic attending mode for bottom-up processing in rhythm perception. The different intensities of attention on attentive and inattentive information may show enhanced attention on strong beats and suppressed attention on weak beats in general. However, the increase of attention on strong beats may not always suppress the attention on weak beats, and may in fact facilitate the perception of the weak beats, as grouping or parsing using strong beats as a framework implies sufficient cognitive efficiency to process sequential information in rhythm perception.

Neural mechanisms regarding attention

Large and Snyder (2009) proposed that neural resonance is involved in rhythm perception, and that beat perception has a spontaneous neural oscillation. In addition, high-frequency oscillatory activity in the brain, which is beta activity in neural oscillations during the anticipatory rhythmic burst during rhythm perception and production, has been suggested to synchronise between auditory and motor brain regions.
‘Internal self-sustained oscillations’ proposed by Large and Palmer (2002) can be understood as entrainment. In this study, the oscillations correlated with beats in rhythms and perceiving temporal regularity were created by the self-sustained oscillations, which were dependent on temporal expectancy and adapted to temporal fluctuations. This computation model on metre perception proposed multiple internal oscillations occur at different periods, which in turn affects the rhythm perception.

Neural mechanisms regarding attention in rhythm can be also considered in terms of the intensity of attention, in which neural activity can be related to the neuronal anticipation during rhythm attending. For example, Driver and Frith (2000) proposed the idea that attention modulates baseline neural activity in the sensory areas. They introduced and discussed an fMRI study, which demonstrated that perceptual expectation activates the higher sensory areas before actual perception activates the primary sensory areas (Chawla, Rees, & Friston, 1999). This fMRI study investigated the perception of colour and motion processes during a detection task, and found the significant activations in higher visual cortices of V4 for colour and V5 for motion even before the presentation of visual stimuli activated the primary visual cortex, thus suggesting the attention to visual stimuli modulates the activation in higher cortex.

The modulated baseline due to attention before the perception may be common in general sensory processing, including auditory processing. For example, perceiving temporal regularity induces temporal expectation or entrainment during listening to rhythms, and the baseline activation in higher sensory areas can be modulated by the expectation created by the temporal regularity. Presumably, the effect of modulated baseline may present not only in the higher sensory area but also in the primary and secondary sensory areas. Driver and Frith (2000) pointed out that the shift or modulation of baseline activity due to attention suggests that brain activity responds to the
CHAPTER 2. Literature review

endogenous factors, such as attention, rather than exogenous factors, such as external events, although the general problems with neural coding related to the difficulty to dissociate exogenous and endogenous factors at a neural level remains.

The dissociation of exogenous and endogenous factors becomes even more complicated when the stimuli relate to temporal processing, such as rhythm perception. Rhythmic structures as exogenous factors are usually continuous temporal sequences, which simultaneously induce the endogenous factor of attention during the rhythm perception, thus to dissociate the two factors in rhythm perception is challenging. The amplitude of neural activity in rhythm perception is consequently assumed to include the interactions between exogenous and endogenous factors. Meanwhile, Fritz, Elhilai, David, and Shamma (2007) argue that measuring attention by its signal of duration, intensity, or selectivity makes it difficult to exclude potential confounds, such as task design or subjectivity. For example, most physiological studies of EEG and MEG have shown an increase in neural activity in auditory cortex when attention to sound stimuli increases (e.g., mismatch negative (MMN) paradigm reflects that an oddball event increases the amplitude of neural activity). However, temporal expectation or entrainment has been also suggested to suppress neural activity in the auditory cortex (Martikainen, Kaneko, & Hari, 2005).

In summary, neural mechanisms regarding attention to rhythms regarding can be understood in terms of both the temporal aspect and intensity (or magnitude) of neural activity. The correlation between increased attention and increased neural activity may be difficult to confirm experimentally due to the difficulty with dissociating exogenous and endogenous factors in rhythm processing, however, it is important to note that expectation or entrainment created by the perception of temporal regularity or metrical
organisation in rhythms, which induces top-down processing in rhythm perception, modulates the neural activity in higher sensory areas.

2.1.4. Prediction in rhythm perception

This section discusses prediction during rhythm processing. Due to the temporal regularity of metric rhythms, prediction is one of important cognitive processing that can often be expected during rhythm listening. In particular, the cognitive processes of temporal regularity and metrical organisation in rhythmic structures resulting in future-oriented attending mode may relate to prediction, or possibly top-down processing in rhythmic sequences. Predicting following rhythms due to the established temporal expectation is not only important cognition in music listening, but also suggesting how we learn a model to process temporal information and perceive new information with the established model. In this sense, neural mechanisms of rhythm perception can be understood with a predictive coding model in terms of this learning aspect. In addition how perceptual expectation can affect (i.e. enhance and suppress) brain activity is discussed with the studies dealing with repetition suppression in fMRI research.

Predictive coding model to understand rhythm perception

The predictive coding model assumes that there are hierarchical relations among the cortical areas in the brain, and that cause and effect relations exist during the processing of sensory information, which modulate the optimisation of prediction and the minimisation of prediction error at the perceptual level (Friston, 2008; Friston & Kiebel, 2009). According to the theory of free energy in the brain (Friston, 2010), when prediction is optimised during perception, action, and learning, the free energy in the
brain becomes decreased. This free energy principle has been also suggested at synaptic levels. For example, predictions occurred in presynaptic level and prediction errors occurred in postsynaptic level increase the connection strength between the two, in which the predictions efficiently minimise the prediction errors with the assumption that sensory information in the world is consisted of dynamic and hierarchical ways. Thus, the theory of free energy discussing on the synaptic levels can be useful to account for the predictive brain at the global aspect of neural networks, rather than to explain the local changes of the brain activation.

The theoretical framework of the predictive coding model can be useful to explain rhythm perception. For example, the sensory information in rhythm perception is mainly comprised of beat and metre that are laid in a hierarchical manner, resulting in top-down processing and prediction. Temporal regularity during beat and metre perception may optimise prediction to subsequent rhythms, and minimise predictive errors, which consequently reduces the free energy in the brain. Thus, it can be predictable that more predictable rhythmic structures will result in the less free energy required in the brain.

The two attending modes of future-oriented attending and analytic attending (Jones & Boltz, 1989) can also be understood regarding predictions and predictive errors in relation to top-down and bottom-up processing, respectively. The top-down and bottom-up processing may interact each other during rhythm perception, suggesting that two processes might be differently weighted depending on a given rhythmic structure. For example, top-down processing may largely employ the future-oriented attending mode in metric rhythms, which optimises predictions while minimising prediction errors. On the other hand, bottom-up processing may result in the analytic attending mode, which can be the process to learn from prediction errors to optimise predictions. Thus,
the interactions between the two cognitive processes and two attending modes in rhythm perception can result in the differently weighted relations between optimising predictions and minimising prediction errors in rhythmic structures.

The idea of free energy principle regarding the predictive coding model presumably suggests that the optimised predictions or minimised prediction errors may suppress neural activities. Thus, it is predictable that temporal regularity in sequences will minimise prediction errors, requiring less energy in the brain. The increased or decreased neural activity due to endogenous factors (e.g., attention, prediction, or expectation), however, is still controversial as discussed in the previous section. For example, the increased predictions during the perception of regular rhythms may minimise prediction errors to subsequent perception of the rhythmic patterns, resulting in reducing neural activity. However, the increased prediction during the appreciation of familiar music, such as emotional response to music, may not necessarily reduce, but enhance neural activity. The latter case may relate to the aesthetic experience (Overy & Molnar-Szakacs, 2009; Schaefer, Overy, & Nelson, 2013), whilst the former case can be related to the cognitive efficiency to resolve rhythmic complexity.

Nevertheless, additionally evoked neural activity can be happened to minimise prediction errors (or by the increased attention during entrainment) in higher sensory cortices before the actual perception of sensory input in low primary sensory cortices. This additional neural activity in higher sensory cortices may potentially modulate the neural activity in lower primary cortices as well, or at least the neural networks regarding connectivity strength (as free energy principle in global neural networks assumed). For example, sensory processing in primary sensory areas may need less energy to perceive information when prediction errors are minimised, which can be understood as a cognitive benefit.
CHAPTER 2. Literature review

Taken together, the predictive coding model supports a rule-based, hierarchical cognitive processing, such as rhythm perception, in which where metrical organisation may minimise predictive errors, and at the same time, optimise predictions in its hierarchical temporal structures. The free energy principle is especially considered in relation to neural mechanisms in beat and metre perception. However, the neural activity in higher sensory areas in accordance with the increased predictions and the decreased prediction errors in metric rhythms can be theoretically dependent on a way to appreciate rhythms (e.g., emotional response to music versus cognitive processes of resolving rhythmic complexity)

Neural repetition suppression

Whether or not the increased predictions suppress neural activity, it is highly predictive of perceptual expectation that either advanced cues or repetitions of the same stimuli would modulate the involved neural activity even before perceiving actual sensory information. Todd (1994) explained a masking phenomenon at the cellular level, in which a tree model of ‘rhythmogram’ describes rhythmic grouping. This theory is based on biological mechanisms in auditory memory system, in that a forward masking happens in the hair cell at the level of the nerve cell in the auditory system. For example, if a pulse is repeated within 400ms, the second pulse reduced its peak response or neuronal activity. If the gap between pulses is smaller, the reduction becomes larger. Similarly, Ogawa, Lee, Stepnoski, Chen, Zhu, and Ugurbil (2000) demonstrated that when a visual stimulus of 10ms duration was repeated with an inter-stimulus interval (ISI) of between 200ms and 1000ms, the BOLD contrast in primary visual cortex (V1) was suppressed. It will be important to examine both whether or not these low-level
sensory processing can be related to higher cognitive functions of prediction and whether higher-order cortical areas can also be modulated by the repetition of sensory information.

Regarding the activation of the higher-order cortical areas involved in prediction and attention to the prediction, Büchel, Coull, and Friston (1999) demonstrated that the more repeated trials reduced more fMRI response in secondary visual areas during learning of the associate visual objects with their location, and the estimation of effective connectivity of neural networks involved in this learning process was shown to be strengthened. Summerfield, Tritschuh, and Monti et al. (2008) also demonstrated that the activation in the fusiform gyrus (FFA) was decreased by the repetition of visual stimuli that consisted of a pair of the same or different face images.

On the contrary, research from single neuron recordings in primary auditory cortex in rats showed that neural activity was enhanced when expectation happened during auditory perception (Jaramillo & Zador, 2011). In this study, neural activity in rats were measured during both anticipation and effect periods using detection and discrimination tasks, in which the anticipation period before showing target stimuli, compared with the condition showing the stimuli without anticipation, enhanced local field potentials (LEPs).

Taken together, repetition suppression in neural activity due to perceptual expectation has been shown in several fMRI studies, and this suppression can be further understood by the neural mechanisms of a free energy principle in relation to predictions. The evidences showing that when expectation was created, neural activity in both low-levels and higher-order cortical areas was suppressed suggest that the perceptual expectation for predicting subsequent information may be well suited to the predictive
coding model and free energy theories to account for the cognitive processing during the perception of hierarchical, metric rhythmic structures.

2.1.5. Summary
Theoretical examinations of the use of the musical terms, rhythmic grouping, attention, and prediction have been considered for understanding rhythm perception. Hierarchical metrical rhythmic structures have been found to be important as a factor to modulate the cognitive processes between different levels of beat and metre. In particular, higher-order cognition in metrical grouping, such as attention and prediction, has been discussed to modulate the brain activation in both lower levels and higher-order cortical areas.

2.2. Behavioural studies of rhythm processing
The current thesis was initiated by questioning what the key factors of rhythmic structures are, and how the key factors of rhythmic structures are perceived in the brain. This section deals with the first question of the key factors of rhythmic structures, which result in the positive effects on the various types of behavioural performance, such as improving the accuracy of perceptual and motor tasks and increasing cognitive capacity, such as working memory. These positive effects in rhythmic behaviour, however, can be advanced with understanding the underlying theoretical concepts of musical terms and relevant cognitive processes in the previous section. In this section, behavioural studies in rhythm perception focusing on the higher cognitive processing of metrical organisation are reviewed. As it has been suggested that different levels in rhythmic structures, such as beat and metre, may have different types of cognitive processing (e.g.,
universal ability to perceive temporal regularity, and the acquired or learnt ability for perceiving higher-level metrical organisation), the developmental aspects in rhythm perception has been also discussed. Thus, behavioural studies of rhythm perception summarise two parts: firstly, the developmental aspects of rhythm perception are reviewed, and secondly, the positive effects of metrical rhythmic structures are discussed.

2.2.1. Developmental aspects of rhythm perception

Rhythm perception is one of cognitive skills that humans acquire from the birth of life. The evidence to support the developmental aspects of rhythm perception has several aspects.

Early development of rhythm perception

Numerous studies have shown that rhythm perception develops at early age. For example, it has been shown that infants discriminate the different types of rhythmic grouping in auditory perception (Demany, McKenzie, & Vurpillot, 1977; Chang & Trehub, 1977), of which rhythm sequences were not necessarily musical, but organised by the gestalt principle of proximity. Infants inferred beats during rhythm perception (Winkler, Háden, & Ladinig et al., 2009), in that when infants heard the omitted down beats in rhythm perception, the neural activity relating to the violated sensory expectation was evoked. In this study downbeats suggest the accented metrical organisation based on the temporal regularity of pulse. Thus, this study can also be further understood that infants may detect both the temporal regularity of pulse and the metrical patterns of strong and weak beats.
Hanon and Johnson (2005) demonstrated that infants perceived metrical organisation, in which 7-month-old infants discriminated between duple and triple metre during rhythm perception. Similarly, Bergeson and Trehub (2006) have shown that infants detected duple metre, but not triple metre, suggesting that a duple metre pattern may be a predisposition in rhythm perception. That infants discriminate between different types of metrical organisation (i.e., duple and triple metre) suggests that the cognitive processes involved in metre perception develop at an early age. A study about statistical learning of sequential sound patterns supports this. For example, Saffran, Aslin, and Newport (1996) demonstrated that 8-month-old infants could learn language processing based on statistical learning. In this study, 8-month-old infants could statistically learn to segmenting words through a short period of exposure for 2 minutes. The mechanism of statistical learning of neighbouring speech sounds may be extended to general learning of auditory perception, such as beat and metre perception.

Not only the perceptual discrimination of rhythmic structures, but also the ability to interact between rhythm perception and production has been found to develop early. Zenter and Erola (2010) showed that infants can move to music rhythmically, and that they can adjust their movements to the tempo of sound stimuli. This study suggests that the cognitive processing to interact between rhythm perception and their movements is early developed in infants. Phillips-Silver and Trainor (2005) also demonstrated that when infants heard either a binary or ternary rhythmic pattern, they listened to the rhythmic pattern that are congruent to their bouncing movements longer than the rhythmic pattern which are not congruent. When infants were not bounced prior to listening, the preference to certain rhythmic patterns was not shown, and this study suggested a multisensory effect on the auditory rhythm perception such as a vestibular-auditory interaction, suggesting that the rhythmic movement affects rhythm perception.
Both studies suggest that the interactions between rhythm perception and production are developed early.

It has been shown that rhythm perception is associated with higher cognitive process at an early age. Fujioka, Mourad, and Trainor (2011) showed that the neural activity of infants showed different patterns depending on 4 month and 12 month old infants while they were listening to an isochronous single piano tone. The EEG results showed that 4Hz neural activity in temporal areas was found in 4 month infants, whilst 6Hz neural activity in temporal-frontal areas was found in 12 month old infants. Fujioka et al. suggest that the perception of rhythmic stimuli induces neural activity in bilateral temporal areas as well as the additional brain activity of the frontal areas of 12 month old infants, suggesting that this additional neural activity may represent the higher cognitive process of rhythm perception, which is projected from thalamo-cortical networks to higher cortical networks during auditory perception.

In children aged 5 to 7 years old it has been shown that they not only perceived temporal regularity in rhythms, but also used the temporal regularity for grouping and reproducing the rhythms (Drake & Gérard, 1989). The ability to use temporal regularity of rhythmic structures as a referential level to group rhythms was significantly improved by the age in 7 year old children compared with 5 year old children. The result suggests that the cognitive skill to perceive temporal regularity and group the subdivisions of beats into the beat level develops early age, even without a special musical training.

Taken together, perceiving temporal regularity and metrical organisation is developed early, and a cognitive skill to use the temporal regularity as a referential time level for organising rhythms into hierarchical metric patterns is also developed early, suggesting neural correlates of metrical organisation in rhythm perception might be well developed in healthy adults without formal musical training.
The effect of musical training on rhythm perception

Rhythm is an essential part of music learning in that the musical skills relating to rhythm perception and production are practiced and trained repetitively. As a result, there can be significant differences between musicians and non-musicians in terms of rhythmic skills. In musician, the positive correlations between musical training and the rhythm skills relating to rhythm perception and production can be shown, although the perceptual skills of simple rhythmic patterns of metrical organisation have been suggested to be early developed without training at an early age in the previous studies. Interestingly, it has been suggested that rhythm skills may relate to non-music cognitive skills, such as spatial-temporal reasoning (Rauscher, Shaw, & Levine et al., 1997), or phonological and reading skills (Anvari, Trainor, Woodsides, & Levy, 2002). Despite the relations between rhythm skills and non-music cognitive skills, the rhythm skills might be mostly affected by musical training (the positive effect of music training on other non-music cognitive skills will be discussed later in the next section).

Drake (1993) examined a reproduction rhythm task performance using both simple and complex metric rhythms in children, adult musicians, and adult non-musicians. The accuracy of task performance showed that all groups showed a similar level of accuracy in their task performance, in which rhythmic stimuli comprised of subdivisions of the beat based on a binary pattern. However, the accuracy in the rhythm reproduction task proportionally increases with the increased period of musical training, in particular, when more complex, hierarchical rhythmic structures are tested. The results of this study suggest that working memory capacity in the processing of metric rhythms might be relating to musical training, in that metrical organisation is a key factor to show
the effect of musical training. Drake, Penel, and Bigand (2000) also showed that musically trained listeners compared with non-musically trained listeners organised rhythmic structures using a time frame for a longer span. This study also supports that working memory capacity of musically trained listeners may have better cognitive efficiency to process the metrical organisation in rhythms.

Regarding attention in rhythm perception, it has been shown that musicians compared with non-musicians showed the enhanced attention to metrical accents when the perceptual sensitivity to rhythmic grouping in differently accented isochronous tone sequences was examined (Kung, Tzeng, Hung, & Wu, 2011). In this study, participants subjectively interpreted metre perception with isochronous tone sequences depending on their musical training, and the musicians showed superior performance in the rhythm tasks.

Brochard, Abecasis and Potter et al. (2003) showed that subjective grouping of a binary pattern (i.e., the pattern of strong and weak beats) was a default metric grouping in their ERP study. A paradigm of violating expectancy using isochronous sequences showed the result that a significantly late positive parietal ERP response when a strong beat was violated compared with when a weak beat was violated. The violation was not disrupting the temporal regularity of rhythmic structures, and a strong beat became softer by the decreased volume for altering the metrical grouping. The large signal difference was found when the strong beats were violated, and that was interpreted as suggesting that strong beats receive more attention than weak beats, which supports the dynamic attending theory (Jones & Boltz, 1989). In addition, musicians compared with non-musicians showed greater signal intensity when strong beats were violated.

Taken together, the cognitive processes during rhythm perception are dependent on musical training as well as the levels of metrical organisation in rhythms. The
cognitive strategies employed in the use of temporal regularity as a referential time level for grouping metric rhythms improved rhythm task performance while expanding the time span or capacity to memorise rhythms, of which higher cognitive processing may also be found in general cognitive domains beyond music.

*Cultural influences on metre perception*

Numerous studies have shown that culture is an important factor, which affects rhythm perception. For example, Hanon and Trehub (2005) showed that infants of 11-12 months age perceived foreign rhythmic structures, whilst adults could not distinguish the foreign rhythmic structures, suggesting that there exists a sensitive period to learn or to acquire certain rhythmic patterns of culture. Iversen, Patel, and Ohgushi (2008) also showed that two different native speakers of English and Japanese grouped rhythmic sequences in different ways. This study interpreted the results that the different rhythmic patterns embedded in their own language resulted in different rhythmic grouping. More interestingly, the results in this study showing that linguistic rhythms affected rhythmic grouping may suggest that learning language can be specifically dependent on learning rhythmic aspects of language, particularly on foreign language processing in adults. In this study, duration was more affected than loudness by individuals’ subjective rhythmic grouping based on cultural differences (i.e., the duration patterns of short-long or long-short were differently perceived between English and Japanese native speakers). This suggests conversely that volume or pitch accents might less affect subjective grouping of individuals in rhythm perception.

Comparative studies between African and western tonal music in theoretical research discussed considerable similarity and differences in rhythmic structures
between two cultures, where the temporal regularity in terms of accent patterns and the metre in terms of grouping in rhythms were shown to be important differences between the two music styles (Temperley, 2000; Kvifte, 2007).

Musical training on rhythm perception suggests positive effects in rhythm tasks, such as the increased memory capacity, better task performance in rhythm tasks. On the other hand, cultural differences in rhythm perception suggest that the sensitivity to discriminate subtle differences of rhythmic patterns can be decreased with age, which are affected by certain cultures, suggesting that beat and metre perception can be distinctive rhythmic structures in western tonal music.

Deficit in rhythm processing
Numerous studies have shown a deficit in the ability to perceive rhythms, for example, due to a lesion in the brain (e.g., stroke patients) or to degenerative diseases (e.g., Parkinson’s disease or Huntington’s disease). The ability to discriminate duration was impaired in patients with lesions in either left or right hemispheres, while metrical interpretation of musical sequences were intact in the patients (Peretz, 1990). The patients with Parkinson’s disease (PD), in particular, with dysfunctions in the basal ganglia had been found to be impaired in discriminating beat-based rhythms (Grahn & Brett, 2009). By examining various rhythm tasks in dyslexic children, it was shown that dyslexic children compared with controls had difficulties in rapid temporal processing (Overy, Nicolson, Fawcett, & Clarke, 2003). Deficits in rhythm perception can be due to various reasons and occur in various levels of rhythmic structures, suggesting rhythm deficits may range from sensory processing to higher cognition.
CHAPTER 2. Literature review

Taken together, cognitive processing in beat and metre perception develops at an early age, where rhythm skills relating to metrical organisation can be affected by musical training, aging, and cultures.

2.2.2. Positive effects of metric rhythms on cognitive processes

In the previous discussion on the developmental aspects of rhythm processing, music training has been suggested to further develop the perception of metrical organisation. This section discusses the relationships between rhythm perception and the improved task performance.

Studies using rhythm tasks have shown that the perception of rhythmic complexity is improved by temporal regularity or metrical organisation. Essens and Povel (1985) showed that metrically interpreted patterns compared with a non-metrically interpreted pattern improved the accuracy in a rhythm reproduction task. Following this, Povel and Essens (1995) proposed a theory that an internally induced clock improves the accuracy of a reproduction task. Rhythmic cues of the temporal regularity of rhythm have been suggested to improve perceptual task performance on a range of tasks (Mathias, Palmer, Perrin, & Tillmann, 2014; Rothermich, Schmidt-Kassow, & Kotz, 2012; Escoffier, Sheng, & Schirmer, 2010; Cason & Schön, 2012), where the temporal regularity of rhythms may enhance perceptual expectation for predicting subsequent information. Huss, Verney, and Fosker et al. (2011) showed that the perception of metrical structures was related to the auditory perception of rise time\(^3\), phonological processing, and literacy skills in children. In this study, they tested the comprehensive categories of perceptual skills, suggesting that the metrical structures of music might not

\(^3\) The amplitude of sound can be divided into rise time, steady state, and fall time. Rise time or amplitude envelope onset in their study was logarithmically modulated from 15msce to 300msce.
be a key factor to be linked to language processing. The results were interpreted as suggesting that temporal processing in music might be relevant to the cognitive processing in sequencing of various auditory information.

Temporal regularity of rhythms also has been suggested to improve motor behaviour. For example, a rhythmic stimulus, such as the regular, metronomic sound, immediately reduced stuttering (Arzin, Jones, & Flye, 1968), and this positive effect lasted for a few minutes, suggesting that speech production in synchrony by a metronome sound can have positive effects in the syndrome of stuttering. A task to tap to the beat in synchrony in metric rhythms compared with isochronous sequences showed that even if metric rhythms did not improve the accuracy, but influenced the higher-level patterning during tapping (Patel, Iversen, Chen, & Repp, 2005). The positive effects of rhythms can presumably be considered as two aspects: the temporal regularity of pulse may affect initiating motor action, whereas the metrical organisation regarding higher order rhythmic structures may affect the higher-order cognition, such as motor planning.

The interactive, hierarchical relations between beat and metre, and its top-down processing due to metrical organisation suggest that rhythm perception might be further extended to other higher order cognition, such as memory. Dowling (1973) demonstrated that rhythmic patterns of simple melodies were differently grouped by subjective chunking and memory span of individuals, suggesting that rhythmic grouping are related to higher order cognition. As discussed in the previous section, young children used the temporal regularity of pulse to group metric rhythms (Drake & Gérard, 1989), where the number of pulses rather than the number of rhythmic elements was key to determine on children’s memory load.

Meanwhile, the accuracy of task performance in both rhythm perception and production may have a limit to explain human rhythmic behaviour, because it can be the
subsidiary result of the rhythmic behaviour, rather than the aim of rhythmic behaviour. More fundamental questions regarding rhythmic behaviour can be why certain rhythmic structures are more preferred, shared, or attended by listeners, which are possibly related to affection or appreciation in music. To be free from the limitation of measuring accuracy of rhythmic behaviour, thus, measuring brain activity can be useful to account the nature of rhythm processing, which may be different from the rhythm processing actively engaged by cognitive tasks. Nevertheless, it is a dilemma to set a completely passive listening condition in empirical research. We often observe rhythmic behaviours in daily life, such as tapping or nodding, but the fundamental understanding of how beat and metre perception are processed in the brain remains largely unknown, although the rhythm perception leads to such rhythmic behaviours.

2.2.3. Summary

Evidence from behavioural studies reveals that the perception of beat and metre develops at an early age, develops further by training, and even can be impaired by lesions in the brain. This review also supplements the theoretical understanding of rhythm processing through defining key of rhythmic structures, which are temporal regularity and metrical organisation. The positive effects relating to temporal regularity and metrical organisation in rhythms can be found in various tasks of perception, cognition, and motor behaviour across music and non-music domains. In particular, the evidences supporting positive effects of rhythmic components (i.e. the improved motor movements in synchrony with a beat, and improved rhythm task performance using metrical organisation relating to higher order cognitive processing) supplement the need of theoretical approaches to distinguish temporal regularity and metrical organisation.
2.3. Neuroimaging studies of rhythm processing

This section focuses on exploring the neural basis of rhythm, beat and metre perception, with particular consideration of metrical organisation. Firstly, the auditory and sensorimotor areas, which have been found to be involved in rhythm perception, are discussed. Secondly, IFG (in particular BA44 and 45) region is discussed to examine its role in rhythm perception. In order to examine the neural correlates of metrical organisation in rhythm perception, rhythmic complexity regarding temporal regularity and metrical grouping are discussed.

2.3.1. Auditory and sensorimotor areas in rhythm perception

As discussed in the review of behavioural studies, perceiving temporal regularity in rhythm affects motor movement in synchrony with the temporal regularity. Thus, co-activation between auditory and motor regions seems highly predictive for the neural correlates during both rhythm perception and production. In spite of that, several questions are promptly brought up: does rhythm perception involve the same regions that are involved in rhythm production? Does rhythmic complexity modulate the activations in auditory and sensorimotor areas during rhythm perception? To begin with a review for the role of auditory and motor areas in rhythm perception, two sections are consequently elaborated: (1) auditory and sensorimotor areas involved in rhythm perception and production, (2) auditory and sensorimotor areas involved in rhythmic complexity.
2.3.1.1. General introduction to auditory and sensorimotor areas

Numerous brain imaging studies have shown that rhythm perception activates auditory and sensorimotor areas, including the primary and secondary auditory areas, premotor cortex (PMC), pre-supplementary areas and supplementary areas (pre-SMA/SMA), insula, the basal ganglia (BG), and cerebellum (Penhune, Zatorre, & Evans, 1998; Sakai, Hikosala, & Miyauchi et al., 1999; Bengtsson, Ullén, & Ehrsson et al., 2009; Grahn & Brett, 2007, Grahn & Rowe, 2009; Chen, Penhune, & Zatorre, 2008a, 2008b; Chapin, Zanto, & Jantze et al., 2010; Kornysheva, von Cramon, Jacobsen, & Schubotz, 2010; Chen, Zatorre, & Penhune, 2006). In addition, the researches into auditory and motor areas in tapping tasks during a steady beat or isochronous sequences have been also considered, although they were not motivated as music-related researches (Rao, Harrington, & Haaland et al., 1997; Lewis, Wing, & Pope et al., 2004).

Importantly, rhythm perception has been found to activate auditory and sensorimotor areas despite a lack of actual movement. The activation areas during rhythm perception are similar to the sensorimotor regions that are involved in rhythm production (Bangert, Peschel, & Schlug et al., 2006; Grahn & Brett, 2007; Chen, Penhune, & Zatorre, 2008a, 2008b; Chapin, Zanto, & Jantze et al., 2010), including primary and secondary auditory areas, pre-SMA/SMA, PMC, and cerebellum. The activation of primary motor areas without actual movement was also shown when professional musicians imagined music performance (Haueisen & Knösche, 2001; Meister, Krings, & Foltys et al., 2004).

Chen, Penhune, and Zatorre (2008b) compared neural correlates during rhythm perception and production using fMRI, and demonstrated that sensorimotor areas including SMA, PMC, and cerebellum are commonly activated in both rhythm perception and production conditions. In this study, three conditions of passive listening,
listening with anticipation to tap, and listening with tapping in synchrony were examined. Activation in the PMC was further delineated into dPMC, midPMC, and vPMC, in which vPMC and dPMC were significantly activated during tapping in synchrony condition, whereas midPMC was activated across all three conditions. They suggested that the motor system including midPMC, SMA and cerebellum as important neural regions during rhythm perception.

Studies of both rhythm perception and production commonly have shown that auditory and sensorimotor areas, including auditory areas, SMA, primary motor cortex, PMC, cerebellum, and BG. Subregions in the auditory and sensorimotor areas during rhythm perception are specified in below. The studies in music as well as music-relevant research are included for the broaden understanding about the functions of auditory and sensorimotor areas.

**Auditory areas**

In auditory areas, the structural asymmetry has been discussed with the range from fetus to adult brain, suggesting a language dominant left hemisphere is larger than right hemisphere in most of the population who are right-handers. For example, Chi, Dooling, and Gilles (1977) demonstrated that the fetus had thicker auditory structures in the left hemisphere than right hemisphere. In the adult brain, the size of temporal speech regions in the left hemisphere was larger than right hemisphere (Geschwind & Levitsky, 1968; Galaburda, Sanides, & Geschwind, 1978; Dorsait-Pierre, Penhune, & Watkins et al., 2006). The structural asymmetry of auditory areas from fetus to adult brain suggests that neural plasticity relating to language processing may occur in the left hemisphere.
CHAPTER 2. Literature review

The evidence of the structural asymmetry of auditory areas also suggests a potential functional asymmetry between the two hemispheres. Several neuroimaging studies have shown that left auditory areas have the functions of speech or language processing. For example, Belin, Zilbovicius, Crozier, Thivard, and Fontaine (1998) demonstrated that the auditory areas were modulated in different ways during the perception of non-verbal sound of rapid (40msec) frequency and slow (200msec) frequency. The rapid sound processing was lateralised in left auditory area, and the activation of right auditory areas was significantly decreased while rapid frequency was processed, in which the activation in left auditory areas was not changed much. The results were interpreted as suggesting that the left auditory areas is specialised for rapid and complex sound sequences, such as speech sound.

The asymmetries found in both structure and function in auditory areas show commonly that the left auditory areas were lateralised during the perception of speech or speech-like sound. The processing of rapid and slow sound in left auditory areas (Belin et al.) suggests that the perception of rapid, complex rhythms may be processed in left auditory areas. Similarly, Zatorre, Belin, and Penhune (2002) proposed that left auditory areas was specialised in temporal resolution, such as rapid, complex auditory perception, and spectral processing was suggested to be better processed in the right auditory areas, where relatively lateralised auditory function was commonly found in both music and speech. In addition, the left lateralised auditory activation during metre perception in musicians (Vuust, Pallesen, & Bailey et al., 2005; Abecasis, Brochard, & del Rio et al, 2009) suggests that the left auditory areas may also be related to proficiency due to musical training.

In addition to the left-right asymmetries in auditory areas, the different functions between the dorsal and ventral auditory streams with respect to a hierarchy of
cortical processing was suggested (Rauschecker, 1998), in that the ventral stream of the auditory cortex in non-human primates was found to have a role in the processing of auditory patterns. In particular, the auditory temporal paths, which are connected to the prefrontal cortex, was suggested to process higher cognitive auditory information during auditory perception, such as language and music.

Moreover, Peretz and Zatorre (2005) proposed that temporal processing in musical rhythms might be more widespread and bilateral than pitch processing in auditory areas. For example, that the perception of a metrical simple rhythm activating the bilateral anterior Superior Temporal Gyrus (STG, Grahn & Brett, 2007), supports that the perception of rhythmic patterns may activate the auditory areas bilaterally.

Premotor cortex (PMC)
The premotor cortex (PMC) can be further divided into ventral, mid, and dorsal PMC. The vPMC was, for example, suggested to have a role of feed forward computations while improving the accuracy of auditory and motor timing, of which region is anatomically connected to the basal ganglia (BG) and supplementary motor areas (SMA) (Kornysheva & Schubotz, 2011). As discussed in the previous section, Chen, Penhune, and Zatorre (2008b) showed that the PMC was involved in both rhythm perception and production, in that the midPMC was found to be activated in passive rhythm perception, and v/dPMC was found to be activated during rhythm perception with a synchronous tapping task. Schubotz and von Cramon (2002a) showed that the right vPMC is involved in attention and prediction in visual-motion sequential processing, suggesting that the activation in vPMC with co-activation in frontal lobes relates to perceptual complexity in terms of prediction during the perception and motor planning of sequences. In their
subsequent study (2002b), they showed that the right premotor areas had a dominant role in both sequential production and perception, suggesting the role of right hemisphere in analysing the sequences perceived.

The basal ganglia (BG) and Supplementary motor area (SMA)

Numerous studies have suggested that BG is involved in both motor and cognitive functions including timing, attention, sequencing, and chunking.

Regarding timing and attention, the attention in temporal processing was suggested to activate BG as well as SMA (Coull, Vidal, Nazaria, & Macar, 2004). Predictive or internally generated movements was also found to activate BG and SMA (Cunnington, Windischberger, Deecke, & Moser, 2002). Activation in the BG also was proposed to have a role in sustaining attention during time perception (Ravizza & Ivry, 2001; Nieoullon, 2002). For example, a spaced finger tapping task, which requires both cognitive and motor functions, in a continuation condition (without an auditory cue sound) activated greater activation in the BG than a synchronisation condition (in synchrony to auditory cue sound) (Rao, Harrington, & Haaland et al., 1997; Lewis, Wing, & Pope et al., 2004). Brown and Marsden (1998) proposed a neural hypothesis that the BG would be involved in attentional mechanisms, binding information while facilitating neuronal synchronisation to cortical areas, selecting as well as sequencing the movements. In addition, time perception involves cortico-subcortical areas, in particular BG when a task is time perception compared with other tasks, such as pitch or control conditions (Harrington, Zimbelman, Hinton, & Rao, 2010).

Regarding sequencing and chunking, temporal sequencing tasks were found to activate the BG (Ivry & Spencer, 2004; Dreher & Grafman, 2002; Coull, Vidal, Nazaria,
The BG and pre-SMA circuits were proposed to play a role in temporal sequencing of non-motor functions such as music or speech perception (Kotz, Schwarte, & Schmidt-Kassow, 2009). Research findings from animal studies also supported the proposal that the BG and SMA are involved in temporal sequencing tasks (Shima & Tanji, 2000; Brotchie, Iansek, & Horne, 1991). Brotchie et al. suggested that the BG has a cognitive function during sequential movements, in that neural activity in the BG in monkeys is evoked by predictability, not by motor movement itself during motor tasks.

BG was proposed to have roles in the encoding, learning, and chunking of action sequences (Graybie, 1998), in which BG was involved in slow and implicit learning with respect to cognitive and motor action patterns, and anatomically anterior striatum (caudate nucleus) were projected from the prefrontal cortex within cortico-basal ganglia circuits. The BG in sensorimotor system was suggested to have a role of chunking information. Thus, attention and grouping in rhythm perception may be the neural mechanisms involving the activation of BG.

**Cerebellum**

The cerebellum has been suggested to have a role of timing (Ivry & Keele, 1989; Dreher & Grafman, 2002). Buhusi and Meck (2005) suggested that the cerebellum was involved in the temporal processing of millisecond time scale and precise timing, such as music and speech. Absolute, duration based-timing was also found to activate cerebellum (Teki, Grube, Kumar, & Griffiths, 2011). Auditory motor synchronisation in rhythmic sounds was found to activate cerebellum and cortico-cerebellar circuits (Thaut, 2003). In particular, timing durations with an integer ratio or regular perceptual accents in rhythm
perception, which induce internal beats, have been found to activate bilateral motor areas including cerebellum (Sakai, Hikosaka, & Miyauchi et al., 1999; Lewis, Wing, Pope et al., 2004, Schubotz & von Cramon, 2002a, 2002b; Penhune, Zatorre, & Evan, 1998).

Similarly, Ivry and Spencer (2004) proposed that time processing with respect to predicting exact timing would involve cerebellum, and the timing system would involve the activation in BG.

### 2.3.1.2. Rhythmic complexity in auditory and sensorimotor areas

Complexity modulates brain activity in general, and numerous neuroimaging studies examined rhythmic complexity during rhythm perception using various tasks.

Firstly, aSTG was found to be involved in beat perception, in that simple metric rhythms compared with complex metric rhythms activated greater activation in aSTG (Grahn & Brett, 2007). This result suggests that aSTG may have a role in recognising the metrical patterns of beat. Patients having undergone the cortectomy of unilateral temporal lobe were examined in various music tasks, including measuring the recognition of pitch intervals, rhythm, and metre, suggesting that the posterior STG (pSTG) had a role for melody processing and the aSTG for metric processing (Liégeois-Chauvel, Peretz, & Babaï et al., 1998). The two studies suggest that the aSTG may have a role in the perception of metric patterns.

Rhythmic complexity also modulated the regions traditionally known as motor areas. For example, the BG was activated when a regular beat is internally induced by metrical simple rhythms (Grahn & Brett, 2007). In addition, internally generated beat perception either by auditory or visual stimuli has been found to activate the BG (Grahn, Henry, & McAuley, 2011). The dPMC also was activated to rhythmic complexity (Chen,
 CHAPTER 2. Literature review

Penhune, & Zatorre, 2008a). The pre-SMA/SMA was found to be activated in passive listening and modulated by rhythmic complexity (Begntsson, Ullén, & Ehrsson et al., 2009).

Brain subregions involved in rhythmic complexity, however, showed different activation patterns depending on different tasks given (e.g., a rhythm discrimination or a rhythm reproduction task). For example, the activations in aSTG, pre-SMA / SMA, BG, and left IFG were modulated by rhythmic complexity during a rhythm discrimination task (Grahn & Brett, 2007), whereas the activation of dPMC and the right IFG in musicians group was modulated by the rhythmic complexity during a rhythm reproduction task (Chen, Penhune, & Zatorre, 2008a).

The discrepant activation in the auditory and sensorimotor areas in accordance with the rhythmic complexity of the same metric simple, metric complex, and non-metric rhythms might possibly be due to different cognitive strategies employed for different tasks. Frontal areas activated by rhythmic complexity also showed discrepancies, suggesting presumably different strategies of higher-order cognition for encoding rhythms during rhythm perception may result in the discrepancies. For example, both a rhythm discrimination task and a rhythm reproduction task may require the cognitive processing to memorise subdivisions of beats during the period of rhythm perception, however, however, a rhythm reproduction task may not be able to exclude completely the cognitive process for motor planning.

In addition, the choice of presentation style of rhythmic stimuli may result in the different degrees of familiarity and subsequently result in different brain activation patterns within auditory and sensorimotor areas regardless of the rhythmic complexity. For instance, each rhythm trial used in Grahn and Brett (2007) was different, whereas one specific rhythm was presented for each metric simple, metric complex, and non-
metric rhythm condition in Chen, Penhune, and Zatorre (2008a). Different degrees of exposure to the rhythmic stimuli may be a potential factor to cause discrepancies in brain activation in accordance with rhythmic complexity.

Kung, Chen, Zatorre, and Penhune (2013) found that rhythmic complexity increased activations in the STG and ventral lateral prefrontal cortex (VLPFC) during a tapping task, but the regions were not found in a beat-finding condition. In this study rhythmic complexity was parametrically modulated from metrical simple to metric complex. Neural correlates during both beat finding and tapping to beats conditions included the activation in the VLPFC, STG, superior temporal sulcus (STS), pre-SMA/SMA, v/dPMC, and dorsal lateral prefrontal cortex (DLPFC), using regression analyses with variables such as the metricality of the rhythms, subjective rating, and tapping performance. The results showed that no region was significantly activated by the decreased rhythmic complexity. This suggests that, within metric rhythms, there was no significantly enhanced activation by the increase of rhythmic complexity during beat perception. As the participants of this study were musicians group, neural correlates of non-musicians group in responding to the rhythmic complexity will need to be investigated, because musicianship can be an important factor to determine the subjectivity in interpreting rhythmic complexity. Other regions beyond auditory and sensorimotor areas include, a rhythm discrimination task in Grahn and Brett (2007) showed the activation of bilateral BG, bilateral aSTG, and left IFG, whereas a rhythm reproduction task in Chen, Penhune, and Zatorre (2008a) showed the activation of dPMC, and right IFG, in musician group.

The relations between musicianship and the perception of rhythmic complexity are also discussed in the above rhythm studies. For example, the perception of metric simple rhythms compared to metric complex or non-metric rhythms has been found to
show significant activation of the pre-SMA/SMA in both musicians and non-musicians (Grahn & Brett, 2007), in that musicians compared with non-musicians showed significantly greater activation of pre-SMA/SMA. Similarly, the musicians compared with non-musicians showed greater activations of right DLPFC, right IFG, and cerebellum during resolving rhythmic complexity, but these activations were not shown in non-musicians (Chen, Penhune, & Zatorre, 2008a). Thus, although both musicians and non-musicians showed activation in the same regions during rhythm perception, BOLD response in the activation differed according to musical training, supporting the previous theoretical and behavioural studies on the relationship between musicianship and the rhythm perception. During beat-based rhythm perception there was no significant difference in brain activation depending on rhythmic complexity in musician group (Kung, Chen, Zatorre, & Penhune, 2013), suggesting that the rhythmic complexity and task determine the cognitive processing of rhythm perception between musician and non-musicians groups in different ways.

It has also shown that the activations of sensorimotor areas are differently modulated during rhythm perception and production depending on musicians and non-musician groups (Bangert, Preshel, & Schlaug et al., 2006; Baumann, Koeneke, & Schmidt et al., 2007). Bangert et al. found that the sensorimotor areas, including the auditory and sensorimotor areas of SMA, PMC, and STG, and the IFG, were activated during listening in musicians. In a playing condition of this study, the activations of SMA and dPFC were shown in only musicians’ group, suggesting that a music-relevant task may increase brain activity in musicians compared with non-musicians, which also implies that a passive listening condition (without a task) may not necessarily be consistent with the results in the task-based listening condition. Baumann et al. compared musicians and non-musicians during both listening and playing conditions,
and found that the listening condition activated auditory areas, v/dPMC, SMA, BG, cerebellum, and IFG, whereas the playing condition showed activation in SMA, vPMC, BG, cerebellum, and IFG. Both listening and playing conditions commonly involved sensorimotor areas, but musicians compared with non-musicians showed additional brain activation of SMA, IFG and pSTG. These greater activation in musicians compared with non-musicians suggest that musicians were more engaged during music tasks than non-musicians, and the complexity of listening and playing conditions was not a factor to increase brain activation. This case can be understood the alternative explanation about the free energy principle that was discussed in the previous section, in that, for example, not only musical complexity but also music appreciation may increase neural activity by different mechanisms.

Taken together, although rhythmic complexity as an exogenous factor generally modulates neural activity in sensorimotor areas, musical training or musicianship places differing demands on cognitive processing in rhythmic complexity, affecting neural activity during rhythm perception. In this sense, rhythm research in musicians as well as non-musician healthy volunteers is important for referential understanding, to some extent, of neural correlates during rhythm perception in most populations.

**2.3.1.3. Sensorimotor areas in speech and action observation processes**

The involvement of motor and sensorimotor areas in perception has been extensively discussed during the perception of other cognitive domains as well. For example, the research on speech perception and action observation has shown the activation in sensorimotor areas.
Auditory and sensorimotor areas during speech perception

The motor theory of speech (Liberman & Mattingly, 1985) suggested the involvement of motor areas during speech perception. For example, the PMC was proposed to have a role of speech perception, where the regions were interpreted as having a gestural origin for speech production (Iacoboni, 2008). Pulvermüller, Huss, and Kherif et al. (2006) showed that the articulation of phoneme sound activated somatosensory cortex, and these regions were further distinguished by lips and tongue movements. In addition, Pulvermüller and Fadiga (2010) proposed that speech perception involved sensorimotor circuits, and suggested that the frontocentral action system including IFG were related to higher cognitive processes in language.

Speech perception was found to activate auditory as well as motor areas (Wilson, Saygin, Sereno, & Iacoboni, 2004; Pulvermüller, Huss, & Kherif et al., 2006). Wilson et al. demonstrated that passive listening to meaningless phoneme sounds activated vPMC bilaterally; the regions were also activated during a speech production task. The activation of vPMC in both speech perception and production suggests that speech perception involves the motor areas without actual movements, suggesting a role of vPMC during cognitive processing, rather than the motor functions of speech production. Although if the activation of vPMC in speech perception is due to either motor-relevant cognition (e.g., imagining or preparing for speech production), or auditory perception of phoneme sounds has not been fully understood, but this activation suggests a significant coupling of the neural correlates between speech perception and production.

Hickok, Buchsbaum, Humphries, and Muftuler (2003) compared between the perception and production of novel music and speech sound stimuli to investigate auditory-motor interaction in speech processing. They found that the area of the
posterior Sylvian fissure at the parietal–temporal boundary (Spt) was commonly activated during both perception and production conditions in music and speech. This result suggests that the Spt is involved in articulation tasks of both speech and music.

Numerous studies have shown that sensorimotor areas are involved in cognitive processing during speech processing. For example, Sato, Trembly, and Gracco (2009) demonstrated that left vPMC was involved in phoneme segmentation. In this repetitive transcranial magnetic stimulation (rTMS) experiment, when rTMS applied on vPMC, participants became slower in their phoneme discrimination, which required the processing of phoneme segmentation. The two tasks were phoneme identification and syllable discrimination, and the task performance was not affected by rTMS being applied on vPMC. Similarly, the cognitive processes of vPMC, such as grammatical processing in language sequences (Opitz & Kotz, 2012) and speech segmentation (Cunillera, Câmara, & Toro et al., 2009), were shown, suggesting that sensorimotor areas are involved in higher order cognitive processing.

Scott, McGettingen, and Eisner (2009) have proposed that the activation of sensorimotor areas is not speech-specific. Although evidence that numerous studies in speech processing showed that motor areas were involved during speech perception, this study proposed so-called ‘a joint behaviour’, which is conversational behaviours in speech, to account for the sensorimotor area involved in speech perception. As the conversation is shared by two or more speakers, the timing and rhythmic aspects during the conversational speech between the speakers has been suggested to activate sensorimotor areas. ‘Timing’ in speech behaviour that this study proposed is higher-order cognitive processing than the auditory perception of physical sound information. This suggests that if the former is the case to activate sensorimotor areas, sensorimotor areas are involved in several discrete functions corresponding to each higher cognitive
processing of timing, auditory perception of phonetic sounds, or gestural origin for speech production.

*Sensorimotor areas involved in action observation*

Evidence that sensorimotor areas are involved during visual perception has been shown. For example, action recognition activates the PMC (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), and action observation of touching activates somatosensory areas (Schaefer, Xu, Flor, & Cohen, 2009). In line with the human mirror neuron system, action observation or imagination has been suggested to activate vPMC (Buccino, Binkoski, & Fink et al. 2001; Binkofski & Buccino, 2006). Pre-SMA is involved in the organisation of action sequences (Kennerley, Sakai, & Rushworth, 2004). The activation of auditory and motor areas is involved in music perception, speech perception and action observation. In particular, sensorimotor areas are involved in both perception and production during sensory processing.

Taken together, the involvement of sensorimotor areas in perception has been found not only in rhythm perception, but also speech perception and action observation. Rhythmic complexity relating to metrical organisation has shown large discrepancies in the activation patterns in the sensorimotor areas and frontal areas, including IFG, although auditory and sensorimotor areas have been found to be commonly activated during rhythm perception.

**2.3.2. Inferior frontal gyrus in rhythm perception**

IFG is the major part of the ventrolateral prefrontal cortex of human brain, consisting of Brodmann Areas 44, 45 and 47 (Petrides, 2005), where the subdivisions of areas are
cytoarchitectonically different. In the review by Petrides, the research of monkey’s brain showed the corresponding human brain regions of area 45, in which retrograde fluorescent tracers were projected from superior temporal gyrus in auditory system to the area 45. Area 47 is related to visuotemporal and paralimbic areas and consists of large subdivisions, which are connected to orbitofrontal areas. Broca’s area (BA44/45) particularly has long been discussed with regard to the motor functions in speech production, speech perception, speech comprehension, and due to emerging evidences from neuroimaging studies, Broca’s area also has more recently been suggested to have multiple functions of cognitions (Amunts, Weiss, & Mohlberg et al., 2004).

As sequencing speech and music information has long been considered as analogous processing, it is highly predictive that the Broca’s area and its right homologues may be involved in rhythm perception, which employs cognitive processes such as attention, expectation, and prediction. Thus, this section focuses on the delimited frontal regions of Broca’s area and its right homologue for examining the review of neural correlates of rhythm perception. The role of bilateral IFG (BA44/45) in rhythm perception has not been well discussed, although the activation in these regions was often reported. This review consists of two sections: firstly, introduction to known functions of bilateral IFG (BA44/45), and secondly, the activation of IFG (BA44/45) in rhythm perception were discussed. The latter were reviewed by comparing between music and non-music stimuli as well as between rhythm and non-rhythm tasks.

2.3.2.1. General introduction to IFG

Due to extensive research into language processing in the left IFG, the known functions of IFG can be roughly divided into language and non-language functions, where
lateralisation between left and right IFG is a distinctive issue. In addition, the function can be further discussed with language processing and non-language processing. Discussion about the analogy between language and music processing is beyond the scope of this thesis, but rhythm processing of language is included because predictive brain in sequential processing in IFG may not be domain specific.

To begin with, brain structure of Broca’s area can be considered in line with its function. For example, Nishitani, Schumann, Amunts, and Hari (2004) discussed that asymmetries between the left and right IFG (BA44 and 45) in that BA44 was more asymmetric than BA45, suggesting that vocalisation in relation to speech may develop the lateralised activation in the left hemisphere. Cantalupo and Hopkins (2001) propose that asymmetry in IFG (BA44) is resulted by evolution for more than million years in that they have examined the size of IFG (BA44) of apes, which show that left IFG is larger than the right hemispheric homologue. Broca’s area can be further dissociated by Brodmann Area 44 and 45 (BA44/45) in terms of different cytoarchitectonical structures (Amunts, Schleicher, & Bürgel et al., 1999, Amunts, Weiss, & Mohlberg et al., 2004) in that each region of BA44 and 45 has been suggested to have different functions.

On the other hand, Keller, Crow, & Foundas et al. (2009) have suggested that there is no significant lateralisation between left and right IFG. They examined the variance of the structures of Broca’s area and suggested that there is no strong evidence of structural asymmetry relative to functional asymmetry between Broca’s area and its right hemisphere homologue.

Also, it has been shown that grey matter of musicians is relatively increased than non-musicians in Broca’s area (Sluming, Barrick, & Howard, 2002), in which the results using voxel-based morphometry (VBM) analyses showed that the grey matter of
Broca’s area in musicians group is less reduced in accordance with aging compared the age matched non-musicians group.

*Language function of left IFG*

The left IFG has been suggested to be involved in various aspects of language processing, including speech production, speech perception, syntactic and semantic processing, phonological processing, and complexity. First, recent neuroimaging studies have shown that motor function related to speech production activates the left IFG. For example, internal speech by generating words activates Broca’s area (Hinke, Hu, & Stillman et al., 1993). Watkins and Paus (2004) also have demonstrated in their PET with TMS experiment that the excitability of the motor system of speech production correlates with the activation of Broca’s area, suggesting that Broca’s area is the interface between speech perception and motor systems. The results of these studies support that left IFG is involved in motor function of speech processing, which are in line with the activation of sensorimotor areas in speech perception. Left IFG can be one of sensorimotor area networks between perception and production. Phonological processing involved the left IFG (BA44) (see review Bookheimer 2002; Hagoort, 2005; Heim, Eickhoff, & Amunts, 2008). Heim, Eickhoff, and Amunts (2008) have also proposed a motoric role of the left IFG (BA44).

PD patients with brain lesion in BG have been also assumed to have dysfunction of Broca’s area (BA44) (Lee, Grossman, & Morris et al., 2003). In this study, the PD patients showed a deficit in phonological processing, of which processing was measured by a detection task of phonetic errors. The result suggests the relationship between left IFG (BA44) and BG regarding motor functions of speech.
Beyond the motor function of speech production in Broca’s area, cognitive processes of syntactic and semantic processes have been shown to activate left IFG. Semantic tasks have also been found to activate the left IFG (BA45) (Amunts, Weiss, & Mohlberg et al., 2004; Bookheimer, 2002; Hagoort, 2005; Friederici, Opitz, & von Cramon, 2000). Bookheimer (2002) proposes a subsystem of IFG rather than multiple functions of IFG in terms of the role of IFG in language processing. Friederici, Opitz, and von Cramon, (2000) found dissociated activation of Broca’s area in semantic and syntactic tasks in their fMRI experiment, in that the violation of syntactic processing activates left posterior frontal operculum adjacent to Broca’s area, with greater activation of STG.

Syntactic processing has been suggested to activate Broca’s area (Embick, Marantz, & Miyashita et al., 2000) in that grammatical errors compared to spelling errors showed greater activation in Broca’s area, suggesting that Broca’s area is specialised in syntactic processing. In syntactic processing, more complex stimuli have been shown to increase the activation of Broca’s area (BA45) (Caplan, Alpert, Waters, & Olivieri, 2000). Obleser, Meyer, and Friederici (2011) have also demonstrated that the activation of left IFG was increased by the increased syntactic complexity and the degraded speech perception. Rogalsky, Matchin, and Hickok (2008) also showed that the activity of Broca’s area was increased by syntactic complexity. Bokde, Tagamets, Friedman, and Horowitz (2001) compared words and word-like stimuli in that semantic processing activates left IFG (BA45), whereas phonological processing activates left IFG (BA44). Similarly, when phonological and semantic tasks were compared, phonological processing activates the left STG (Wernicke’s area), weaker activation of right STG, and Broca’s area (Demonet, Chollet, & Ramsay et al., 1992).
Non-language functions of the left IFG

Broca’s area also has been suggested to have a role of non-language function, such as sequencing, working memory, or complexity. Regarding higher cognitive processes in frontal areas in general, Duncan (2001) proposes adaptive coding in PFC, and points out that the function of frontal areas in terms of working memory, attention and cognitive control is not yet fully specified.

Huettel, Mack, and McCarthy (2002) demonstrated that PFC is involved in violation of both repeating and alternating pattern conditions during visual sequencing. The increase of length of sequences increased the percent signal change in BOLD response of neural activation. In their study, however, violations of repeating patterns, not alternating patterns activated BG, which may suggest that BG is particularly dependent on regular pattern recognition, so that when the regularity is violated, the activity of BG increased. In line with the study showed PD patients’ phonological deficit (Lee, Grossman, & Morris et al., 2003), in which the task could be also understood as the violation of expectation. It can be inferred that left IFG (BA44) and BG may have the role of recognition of regular patterns.

Hargoot (2005) proposes that left IFG would have a role of unification in language processing. Although this paper mainly discusses language processing, unification is not limited to the language domain, and can be generalised to other non-language cognitive domains. The unification that he suggests is time-dependent information processing in language beyond processing across syntax, comprehension, semantic, and production. Broca’s area and its adjacent cortex in the left IFG is proposed as a crucial node for the processing of unification. It has been also emphasised that specificity of subregions in the left IFG is important, which can also be in line with the multisystem of IFG (Bookherimer, 2002). Nishitani, Schürmann, Amunts, and Hari
(2004) have suggested that Broca’s area may have an interactive role for multiple
cognitive functions within neural networks in sequencing.

Broca’s area (BA44) also has been considered to have non-language motor
function. For example, Broca’s area has been suggested to have a role of imitation of
action (Heiser, Iacoboni, & Maeda et al., 2003). Using rTMS, they demonstrated that
finger movement was disrupted by a stimulation of TMS in Broca’s area, suggesting that
Broca’s area is essential for finger movements. A non-language motor function, such as
complex hand movements, has been suggested to activate Broca’s area BA44,
suggesting that the region is a part of the motor network (Binkofski & Buccino, 2004;
Binkofski & Buccino, 2006).

However, neuropsychological research with the patients who have lesion in the
left hemisphere showed that they had a difficulty to learn novel complex hand movement
sequences. This suggests that the left hemisphere is involved in sequencing movements,
rather than motor function (Kimura & Archibald, 1974). Nishitani, Schürmann, Amunts,
and Hari (2004) argued that Broca’s area is involved in action understanding and
imitation, rather than motor function, in that action sequencing and parsing (or
segmenting) in terms of perceptual processing have been suggested to involve Broca’s
area. They speculate that the right hemisphere may have a role of nonverbal sequencing,
whereas the left hemisphere is involve in verbal sequencing.

An action execution and observation paradigm investigating the mirror neuron
system has shown repetition suppression in IFG bilaterally (Kilner, Neal, & Weiskopf et
al., 2009) in that the execution or observation of the same action suppressed the
activation of IFG compared to the execution or observation of different action.
Compared to the execution of hand movement, the observation of hand movements has
also been found to activate the IFG (Rizzolatti, Fadiga, & Matelli et al., 1996; Grafton,
Arbib, & Fadiga et al., 1996; Buccino, Binkoski, & Fink et al., 2001), suggesting that the IFG has a role of sequencing in time regarding cognitive processing.

Timing has been suggested to involve the IFG. Coull and Nobre (2008) have suggested that explicit and implicit timing can have different neural correlates. Explicit timing involves the activations of IFG with BG, SMA, and cerebellum, and implicit timing involves the left lateralised premotor and parietal cortices.

Working memory has also been suggested to activate PFC, in particular left IFG (BA44/45) and right MFG (BA9) (Braver, Cohen, & Nystrom et al., 1997) during an n-back task in fMRI experiment.

Hierarchical processing has been suggested to involve left IFG (BA44). When behavioural sequences were hierarchically organised, the subchinking of individual elements of sequences and chunking of superimposed larger units in hierarchical relations of action execution activated BA45, BA44 and the premotor cortex of BA6 (Koechlin & Jubault, 2006). In their study, complexity modulated in selective motor sequencing has been shown to activate bilateral IFG, suggesting that hierarchical processing in motor sequencing may involve the IFG. Broca’s area has also been suggested to have a role of hierarchical processing (Tettamanti & Weniger, 2006), in which hierarchical processing regardless of a specific modality has been suggested as a key factor to activate Broca’s area. In addition, it has been proposed that both language and non-language stimuli may involve the left IFG (BA44/45) in that the regions may be involved in the processing of general sequences (Friederici, 2006). In this paper, the role of the IFG has been further dissociated, in that violating expectation in sequencing activates posterior part of left frontal operculum (BA6/44), whereas complexity in semantic processing involves the anterior part of the left IFG (BA45).
Overall, left IFG (BA44/45) may have a motor relevant function regardless of language and non-language processing as well as higher cognitive functions related to sequencing. The complexity of sequences can be a key factor to modulate the activation of IFG. As discussed IFG may have multiple functions, however, it seems more reasonable to understand a multisystem within IFG, such as hierarchical processing regardless of the various types of information (e.g., speech, music, or action movement).

**Functions of right IFG**

Firstly, attention and time perception have been suggested to involve right IFG. For example, perceiving intervals in time perception has been suggested to activate the right IFG (Rao, Mayer, & Harrington et al., 2001). Right IFG has also been suggested to have a role of timing and attention (Hedden & Gabrieli, 2006, Weissman, Roberts, Wisscher, & Woldorff, 2006). Posner and Dehaene (1994) propose that attention networks within fronto-parietal lobes based on their review of fMRI data regarding visual experiment. Hedden and Garieli (2006) have proposed that the lapses of attention in relation to flow, which are attentional control processes as executive functions, involve prefrontal and parietal lobes, including right IFG. Slower reaction time (RT) was associated with the reduced activation of prefrontal regions (i.e., right IFG, ACC), and the reduction of neural activity happened before the presentation of stimuli, whereas the activity of right IFG was increased again with a faster RT, which they called, reorientation of attention after the lapse. Right IFG and sensorimotor areas are proposed as the core neuroanatomical circuits in timing (see review Coull & Nobre, 2008).

Secondly, inhibition within attentional networks also has been suggested to activate right IFG. Stevens, Calhoun, and Kiehl (2005) have demonstrated that auditory
oddball tasks, such as novel and infrequent target, activate the right hemisphere greater than the left hemisphere. Hampshire, Chamberlain, Monti, Duncan, and Owen (2010) have also suggested that the right IFG has the functions of inhibition and attentional control. Regarding attention by inhibition, Chikazoe, Konishi, & Asari et al (2007) have shown that right IFG is activated during response inhibition across different modalities. Aron, Fletcher, and Bullmore et al. (2003) have demonstrated that that right IFG (BA45) was activated by inhibition of ‘no-go’ condition, of which activation was significantly correlated high stop signal reaction time (SSRT) in that they examined a stop-signal task (or a go/no-go task) with patients with lesions of the right frontal lobe.

Thirdly, the recovery of speech processing after stroke has been suggested to involve the function of the right hemisphere homologue of Broca’s area. For example, Cao, Vikingstad, George, Johnson, and Welch (1999) have shown that stroke patients with a lesion in left hemisphere show the activation in the right IFG during their recovery, of which recovery is measured by semantic tasks in that partially damaged left IFG is also recovered together, suggesting the interaction between left and right hemispheres. Right hemisphere after stroke has been also assumed to assist the recovery of speech production in aphasic patients (Albert, Sparks, & Helm, 1973) in that melodic intonation therapy, which is based on musical processing, showed the intact ability of speech production in patients.

2.3.2.2. IFG in musical rhythm

In this section, the activation of IFG in rhythm processing is discussed with four sub-parts: (i) music versus non-music stimuli, (ii) rhythm versus non-rhythmic tasks, (iii)
musicians versus non-musicians, and (iv) rhythmic complexity in terms of temporal regularity.

Music versus non-music stimuli

To define music specific activation in IFG, studies comparing music and non-music stimuli are discussed, even if the aims of the studies are not focusing on music. In particular, due to the analogous relationship between musical rhythm and speech rhythm, numerous studies in speech processing are discussed.

Brown, Martinez, and Parsons (2006) compared music and language tasks using PET with amateur musicians. The task was to improvise the melodies of music and sentences of language. Neural networks of both conditions showed similar, shared activation patterns in that direct comparison between two music and language conditions showed greater activation of Broca’s area (BA44) in language condition, and right IFG (BA44) in music condition. They have proposed a model of motor phonology and syntax for the left (BA44) in language processing, but the right IFG (BA44) for melodic improvisation has not sufficiently explained.

Abrams, Bhatara, Ryali, and Balaban et al. (2011) have demonstrated that the activation of IFG (BA44/45) is involved in temporal processing of both music and speech. The comparisons between natural and scrambled music and speech sound stimuli have shown that the activation of bilateral IFG (BA44/45) is significantly dissociated depending on music and speech based on the classifying analysis of probabilities. However, in their study, how the subregions of IFG are dissociated depending on music and speech has not been fully elucidated. The results suggest that the right IFG may be
reliably activated by music stimuli compared with other non-music stimuli, i.e., speech in this case.

Although temporal sequences of sound stimuli were not rhythmic, Broca’s area was activated when hummed notes and syllables of human vocal sound were compared (Gelfand & Bookheimer, 2003). Left posterior IFG (or Broca’s area) was activated by both syllables and hummed notes in that syllable sound additionally activated the left supramarginal gyrus, suggesting that the co-activation of left IFG and left supramarginal gyrus might consist of phonological loop. The role of left IFG in their study has been suggested to have a role of sequencing in general beyond language-specific information.

**Rhythm versus non-rhythm task**

When a rhythm task is compared to a non-rhythm task, the activation of IFG has been reported. For example, compared to pitch and timbre tasks, a rhythm task significantly activated left IFG (Broca’s area, BA44) along the left insula (Platel, Price, & Baron et al., 1997). Here, the rhythm task was that participants pressed a left button when rhythm sequences were irregular and a right button when regular. Although regular rhythms were not compared to irregular rhythms in the study (Platel et al.) in that whether left IFG was differently activated by different degrees of temporal regularity or not was not clear, the fact that left IFG responded to a rhythm task suggests that cognitive processing of temporal grouping significantly activates left IFG.

A non-rhythm task in music studies has shown the activation of Broca’s area. When musical expectation in harmonic sequencing is violated, the activation of Broca’s area has been shown (Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch, Gunter, & Cramon et al., 2002). The stimuli in their studies only modulate harmonic relations, so
the rhythmic component or temporal aspect of music stimuli is not the factor to modulate Broca’s area. However, in terms of violating expectation in sequencing, it is reasonably predictable that rhythmic sequences may also have the same mechanism.

In synchronisation tapping to the rhythm compared to self-paced tapping to the noise (Kornysheva & Schubotz, 2011), and isochronous tapping to the rhythm compared to listening to rhythm (Thaut, 2003) showed the greater right IFG activation.

Berkowitz and Ansari (2008) have shown that when musicians improvise sequences of their motor movements, left IFG is activated. In their study, trained musicians participated in improvisation of both rhythmic and melodic motor sequences in that musicians read music scores and played the music during fMRI scan, in which the rhythmic structures for both rhythm and melody conditions are either isochronous sequences or beat-based rhythms. Significant activation of the left IFG (BA44/45) has been shown when rhythm processing (i.e., grouping in time domain) is compared with melody processing (i.e., grouping in frequency domain).

Alcock, Wade, Anslow, and Passingham (2000) have shown that pitch is processed in the right hemisphere, and rhythm is processed in the left hemisphere, with brain-damaged patients. In a perceptual rhythm discrimination task, patients heard rhythm sequences (stimuli used in Peretz, 1990), and judged the same or different rhythms by indicating the cards. Results showed that the scores of rhythm discrimination task were significantly lower in the patients with left hemisphere lesions compared to the patients with right hemisphere lesions. This study proposes a neural hypothesis that left hemisphere may have a role of rhythm discrimination task.

Broca’s area (BA44) has been suggested as a core brain network in musical structure with the ventral premotor area and anterior superior temporal gyrus (Koelsch, 2006). Koelsch (2006) suggests that IFG (BA44) as one of musical syntactic processing
networks, of which syntax has been discussed with violation of harmonic expectancy. Prediction and expectation in terms of semantic processing in auditory perception have been suggested to involve IFG (Koelsch, Kasper, & Sammler et al., 2004). Musical phrase structures have been suggested to be analogous with syntax in language in terms of the activation of Broca’s area (Patel, 2003).

2.3.2.3. Rhythmic complexity in IFG

In this section, the activation of the IFG is further discussed with rhythmic complexity in relation to metrical organisation.

Rhythmic complexity of metrical organisation

Rhythmic complexity is often varied using two or three levels with the two levels being temporal regularity and irregularity (or isochronous and non-isochronous sequences), and the three levels often being referred to metric simple, metric complex, and non-metric rhythms.

Firstly, rhythmic complexity can be varied with metric simple, metric complex, or non-metric rhythms using duration accents (Chen, Penhune, & Zatorre, 2008a, 2008b; Grahn & Brett, 2007). Metric simple rhythms consist of groups of integer units (e.g., 112-22-4-31 or 13-1111-4-22), of which duration of grouping patterns is regular, so that the listeners can infer beat perception from the rhythm sequences. In particular, metric simple rhythms induce not only temporal regularity, but also metrical organisation with a simple binary pattern. Metric complex rhythms also consist of integer units, but integer units are not grouped with the same duration (e.g., 14-12-32-111 or 22-311-4-111), so rhythmic complexity increases. Non-metric rhythms consist of non-integer units (e.g.,
1.4, 2.5, or 4.5), where each unit as well as the groups of units produces temporal irregularity. Therefore, metric simple rhythms introduce beat (or a binary pattern of metrical organisation), whereas metric complex rhythm and non-metric rhythm may not introduce metrical organisation, which can be largely dependent on individual interpretation.

As discussed in the previous section of behavioural studies, the development of metre perception can be varied in accordance with cultural experience, musical training, or brain lesions. Nevertheless, perceiving a simple binary pattern of metrical organisation can be learned at an early age, such as infants and children, suggesting that the binary pattern underlying metric rhythms can be very familiar to listeners. The binary pattern can also have a role to induce top-down processing due to its salient, predictable temporal regularity in rhythm perception. Contrary to metric simple rhythms, metric complex rhythms or non-metric rhythms hardly induce this binary pattern of metrical organisation, as both of which temporal structures do not offer a salient referential time level of beat to group rhythms, and the processing to count the rhythms would be required instead. Therefore, listening to metric complex and non-metric rhythms can be an analytic attending mode with more bottom-up processing than top-down processing using beat. As a result, the comparison between metric simple rhythms and non-metric rhythms may imply more cognitive processes, such as attending modes, prediction, and different weights of top-down or bottom-up processes, compared to the discrimination between temporal regularity and irregularity. To investigate temporal regularity in rhythms, therefore, that rhythmic structures are more controlled can have benefit to specify the neural correlates of perceiving temporal regularity.

Complexity has been also suggested to modulate left IFG in various cognitive domains, such as sequencing of visual stimuli, speech processing, and violating
harmonic expectancy. It is therefore reasonably predictable that rhythmic complexity will activate left IFG. For example, metric simple relative to metric complex and non-metric rhythms activates left IFG (BA45), which are co-activated with BG and STG (Grahn & Brett, 2007). In Grahn and Brett (2007), participants listened to the same rhythms twice, and then when they heard the third rhythm, they indicated whether the third rhythm was the same or different from the first two rhythms. To perform a discrimination task, the participants may listen to details of rhythms, which can be for example the subdivision of beats in metric simple rhythms, rather than listening to holistic pattern of rhythms, which are for example beats or segmentation of rhythms. Therefore, listening to rhythms with a discrimination task may require bottom-up processes of grouping in that the generation of beats in metric simple rhythms can weight top-down processing compared to metric complex or non-metric rhythms. In this sense, the greater activation in left IFG, BG, and aSTG may suggest that encoding rhythms is accompanied with not only bottom-up, but also top-down cognitive processes of metrical organisation.

Then, why the activation in the IFG becomes greater in simpler complexity? Other non-music studies have suggested that the increase of complexity increases the activation of left IFG, however, in Grahn and Brett (2007) although metric simple rhythms as the least complex condition increase the activation of left IFG. As the aSTG has been suggested to be important for metre processing (Liégeois-Chauvel, Peretz, & Babaï et al., 1998), of which results are consistent with Grahn and Brett (2007), the activation of the left IFG in metric simple rhythms may not be related to the increase of complexity, and rather it can be understood as metric processing compared to metric complex and non-metric rhythms, which do not induce metric processing. More specifically, whereas metric simple rhythm induces the temporal regularity based on the
metrical organisation of subdivisions of beat, metric complex and non-metric do not induce the temporal regularity. The increased rhythmic complexity between metric complex and non-metric rhythms, therefore, may be similar to oddball sound experience or mismatch experimental paradigm if listeners are not familiar with the rhythmic stimuli. The proportionally increased rhythmic complexity with metric simple, metric complex, and non-metric rhythm conditions may require sufficient familiarity to rhythmic stimuli.

With the same rhythmic complexity of metric simple, metric complex, and non-metric rhythms, only musicians group have shown that the activation of right IFG (BA44/45) is greater in metric complex rhythms relative to metric simple rhythms (Chen, Penhune, & Zatorre, 2008a). In their study, participants listened to rhythms, and performed a reproduction task. The results showed that the activity of dPMC and DLPFC was modulated by the rhythmic complexity of metric simple, metric complex, and non-metric rhythms in both musicians and non-musicians. In particular, musicians showed greater activation in right dPMC, right IFG (BA44/45), and left cerebellum. With respect to rhythmic complexity, behavioural results of the reproduction task was interpreted to suggest that musicians used a top-down process to perform the reproduction task, whereas non-musicians used grouping by the temporal proximity based on gestalt principles. Also, dPMC has been suggested to have an important role to interact the auditory and motor areas. In contrast to the activation of DLPFC, which has been suggested to have a role of retrieving information for synchronised tapping, VLPFC has been suggested to have a role of extracting temporal information.

Top-down processing of musicians during their listening may include motor plan and rehearsal of tapping, which is more than just temporal processing (Chen, Penhune, & Zatorre, 2008a). They suggest that the greater activation of right IFG is because musicians group use more working memory to perform a reproduction task.
Their subsequent study examined rhythm perception only, showed that the vPMC, of which cortices are closest areas to IFG, was activated in listening with a tapping task or at least with anticipation of a tapping task (Chen, Penhune, & Zatorre, 2008b).

As already discussed (see also 2.3.1.2. Rhythmic complexity in auditory and sensorimotor areas), the discrepancies regarding the activation of IFG in the same rhythmic complexity in the above studies may be due to differences of tasks or the presentation style of stimuli.

The two levels of rhythmic complexity also used in much rhythm research can be considered as isochronous and non-isochronous sequences, or temporal regularity or irregularity. In this case, however, if isochronous sequences are spontaneous and last for a few seconds, the perception of rhythm sequences may result in higher cognitive processes beyond temporal regularity and irregularity. As discussed (see also 2.1.2.3. Attention to rhythm), the length of rhythmic structures, which last for a few seconds, can result in higher cognitive processes regarding grouping.

For example, passive listening of isochronous, metric simple, and non-metric rhythm perception activated left IFG (Bengtsson, Ullén, & Ehrsson et al., 2009) in that the IFG activation was greater in the order of ISO > non-metric rhythms > metric rhythms. If rhythmic complexity among those rhythm conditions are conceptually considered as the parametric pattern of ISO > metric rhythms > non-metric rhythms, it may need additional explanation why non-metric rhythms actually show greater activation than metric rhythms, whereas Isochronous beats activate the left IFG most significantly. Perhaps, the length of rhythmic sequences, which last few seconds with the repetition of each short rhythm corresponding to a rhythmic complexity level, may induce the higher level of rhythmic grouping regardless of metric and non-metric rhythmic structures. Therefore, although subdivisions of rhythmic sequences are metric
simple or non-metric rhythms, a series of repetition of those short rhythm phrases creates
another level of grouping, which may be the perceived rhythmic grouping. On the other
hand, isochronous sequences can also be perceived with higher levels of metric grouping.
Although the study did not intend to investigate another level of rhythmic grouping, the
results suggest that the left IFG involve in cognitive processing during rhythmic or
metric grouping at the higher levels.

Therefore, whether rhythmic or metric grouping at different levels modulates
the activation of IFG become an important question. However, due to the lack of
neuroimaging studies regarding metrical organisation in rhythm perception, it cannot be
defined in this review. Rhythmic complexity has suggested that temporal regularity and
metrical organisation, which are important factors in rhythm perception, modulate the
activation of IFG.

*Rhythmic complexity and musicianship*

Musicianship can be an indirect factor to affect the activation of IFG regarding rhythmic
complexity in rhythm processing, which also has been discussed in the previous section
of the auditory and motor areas and rhythmic complexity. For example, Chen, Penhune,
and Zatorre (2008a) showed that only musicians group showed greater activation of the
right IFG in metric complex rhythms, suggesting working memory of musicians to
perform a reproduction task. As the musicians’ task performance was shown to be better
than non-musicians in their separate behavioural tasks in metric complex rhythms, the
increased activation of right IFG may suggest that musicians are more engaged with a
reproduction task rather than non-musicians, in that musicians have shown to use more
efficient cognitive strategy than non-musicians. Therefore, when the same rhythmic
stimuli are given, greater activation of certain brain regions in musicians compared to non-musicians does not necessarily directly suggest the easiness of task or stimuli for musicians. On the other hand, if rhythmic complexity can be easy enough for non-musicians to perform a task, it is reasonably predictable that the working memory of non-musicians also will show greater activation of the IFG.

Similarly, Sluming, Brooks, and Howard et al. (2007) have shown that orchestral musicians showed superior performance on a three-dimensional mental rotation task, which increased the activity of Broca’s area. The task was to judge match or mismatch during visual perception of visual stimuli, which were highly abstract features that can be correlated with a sight reading skill of the musicians. The results showed that the activation of Broca’s area was greater in musicians compared with non-musicians, and significantly negatively correlated with faster response time of musicians, suggesting that accurate and faster performance of musicians, which may be due to musical training, correlates with the increase of activity in Broca’s area. This increase activity of Broca’s area in musicians’ better performance is, again, not necessarily suggesting better cognitive processing. The increased brain activation of musicians or relatively decreased brain activation of non-musicians can alternatively suggest that non-musicians may not be engaged in a task as much as musicians does due to the levels of complexity.

Professional musicians compared to non-musicians have shown greater activation of Broca’s area and Wernicke’s area when they performed passive listening (Bangert, Peschel, & Schlaug et al., 2006). Although the interpretation of the greater activation of Broca’s areas cannot be specified due to the use of a passive listening condition, the results nevertheless suggest that musicians can be more involved in listening than non-musicians. Because the passive listening condition does not limit or
specify cognitive process during listening, whether musicians as an expert analyse or appreciate music during listening cannot be defined. Nevertheless, it will be important to note that the passive listening condition activates Broca’s areas in both groups.

When professional pianists improvise musical motor sequencing, the activation of IFG has been shown (Berkowitz & Ansari, 2008) and the activation was interpreted as being related to generating sequences. Regarding rhythmic complexity the results can be interpreted that the activation of left IFG is significantly greater in the condition of performing rhythm improvisation compared to performing isochronous beat rhythm condition in musicians. The increase of IFG in improvisation can be potentially interpreted as either the increase of rhythmic complexity, or the proficiency of cognitive skills in temporal processing of the musicians.

Taken together, higher cognitive process of metrical organisation in rhythm perception is a key factor to activate IFG. Previously, the discussion regarding the activation of IFG has not received attention much, however, evidence from rhythm studies suggest that the activation of IFG can be related with rhythmic complexity. In particular, whereas cognitive processes in speech and action observation have shown that the complexity increases the activation of IFG, research into rhythm perception showed that the activation of IFG was increased by metric rhythms compared to more complex rhythmic pattern, or by musician group compared with non-musician group, suggesting that the IFG may also be involved in the decreased complexity in metric rhythms.

2.3.3. Summary

This chapter reviewed neural correlates in relation to rhythm processing. Both sensorimotor areas and IFG have been found to be activated not only during rhythm
perception, but also during the perception of other cognitive domains, such as speech and action observation. Secondary auditory areas and the BG can nevertheless be predicted to be involved in rhythmic grouping in particular with respect to temporal regularity and metrical organisation, suggesting that rhythmic complexity modulates the activation of both the auditory and sensorimotor areas and the IFG, although the discrepancies across rhythm studies are discussed in relation to the tasks and rhythmic stimuli. Rhythmic complexity that is modulated by different degrees of prediction regarding metrical organisation will modulate higher sensory areas, such as secondary auditory areas. Whether the BG is also involved in different levels of predictability regarding metrical organisation is not clear.

Although it can be predicted that rhythm perception will activate IFG, whether the rhythmic complexity of metrical organisation increases or decreases the activation of IFG during rhythm perception remains questionable. For example, the activation in left IFG has been suggested to correlate with the increased complexity in language or other sequencing tasks. On the other hand, a rhythm study comparing between the metric simple and the non-metric and metric complex rhythm conditions showed that the activation of IFG increases when rhythmic complexity decreases (Grahn & Brett, 2007). In this case, the increased activation in IFG may relate to more salient temporal regularity (Grahn & Brett, 2007) or more proficiency in music tasks of the professional musician (Chen, Penhune, & Zatorre, 2008a).

Taken together, rhythmic complexity, which activates IFG, may be divided into two types. Firstly, disruption or violation of temporal expectation may increase the activation of IFG. Secondly, enhanced prediction or higher cognitive processing may also increase IFG via different mechanisms. These two different hypothetical
mechanisms regarding the activation of IFG in rhythm perception will need to be examined.

2.4. Aims of this thesis

The two research questions of this thesis are: (1) what are the key factors of rhythmic structures that affect human behaviour? And (2) what are the neural mechanisms of the key factors that are modulated in rhythm perception? Theoretical examination in the literature review reveals that the metrical organisation of rhythmic structures is a key factor in rhythmic behaviour, where the higher order cognitive processing during rhythm perception especially involves the IFG in the brain. Thus, this thesis hypothesises that IFG will be involved in higher order cognitive processing in metrical organisation in rhythm processing, and rhythmic complexity will increase the IFG activation.

The first fMRI study (Chapter 4) aimed to assess the activation of bilateral IFG during beat perception in complex rhythms, where (1) higher order cognitive processing in the hierarchical structures of metrical organisation between subdivisions of beat and the beat levels was examined, and (2) metrical organisation in the complex rhythms were investigated.

The second fMRI study (Chapter 5), which was further extended from the first fMRI study, aimed to investigate the activation of bilateral IFG during metre perception in isochronous sequences, where (1) higher order cognitive processing in the hierarchical structures of metrical organisation between isochronous beat level and metre level was examined, and (2) the relations between IFG activation and different types of metrical organisation were examined.
Chapter 3 Methodology

The current thesis used a neuroimaging method for the two studies of Chapter 4 and 5. This chapter firstly describes the principles of fMRI, in particular, of blood oxygenation level-dependent (BOLD) fMRI. Secondly, the experimental designs used for two fMRI studies are described (i.e., block and slow event related designs). Thirdly, region of interest (ROI) analysis using a maximum probability map (MPM) to investigate the region of IFG (BA44/45) is discussed, following the development of the MPM-based ROI analysis in the software of BrainVoyager.

3.1. Introduction to functional Magnetic Resonance Imaging (fMRI)

fMRI is an in vivo neuroimaging technique, in particular widely used in cognitive neurosciences, which indirectly measures neural activity. During MR scan, a strong magnetic field results the energy difference of the protons that exist in the water molecules of the blood in the brain. The protons align with the strong magnetic field in that the protons absorb energy, and then return to relaxation while releasing the energy (Edelman & Warach, 1993, see also Figure 1. p.709). MR imaging detects this energy difference in the protons, which is called MR susceptibility.

It has been found that this magnetic susceptibility change was due to paramagnetic deoxyhemoglobin which was originally found in mouse and rat brains, resulting in the oxygenation-dependent contrast (Ogawa, Lee, Nayak, & Glynn, 1990a). Since then, the principle of BOLD has been cited in numerous studies using slightly different concepts (e.g., BOLD and cerebral blood flow (CBF) was often used as the same meaning in many literatures) and terms (e.g., blood-oxygen-level-dependent instead of blood oxygenation level-dependent is used in numerous studies). Due to these
slightly different usages of the concept of BOLD contrast in fMRI research, which can be partly due to its inhomogeneous nature, it would be important and useful to go back to one of original studies, which succinctly describes both the discovery of the mechanism of BOLD contrast and the relevant principle of MR physics (Ogawa, Lee, Kay, & Tank, 1990c, p. 9868):

“It has previously been demonstrated that the presence of deoxyhemoglobin in blood changes the proton signal from water molecules surrounding a blood vessel in gradient echo MRI, producing blood oxygenation level-dependent (BOLD) contrast. BOLD contrast has its origin in the fact that when normally diamagnetic oxyhemoglobin gives up its oxygen, the resulting deoxyhemoglobin is paramagnetic. The presence of paramagnetic molecules in the blood produces a difference in magnetic susceptibility between the blood vessel and the surrounding tissue. This susceptibility difference is "felt" both by the water molecules in the blood and by those in the surrounding tissue, the effect extending significantly beyond the vessel wall. This increase in the number of spins affected by deoxyhemoglobin is a form of amplification. When the susceptibility-induced local field differences exist within an imaging voxel, there is a resultant distribution of shifts in water resonance frequencies. In the gradient-echo method, a phase dispersion of water proton signals is produced at the echo time. This dispersion reduces the signal intensity and the voxel appears dark in the image. These intensity losses, which at high magnetic fields (>= 4 T) extend significantly beyond the boundary of the blood vessel, are the source of BOLD contrast.”

Since oxygenation contrast in MR was reported (Ogawa, Lee, Nakay, and Glynn, 1990a), subsequent studies reliably replicated the same phenomenon. For example, Ogawa, Lee, Nakay, and Glynn (1990a) found that when the deoxy-haemoglobin content in red cells increases, the contrast of images was detected in the brains of live mice and rats using high-resolution magnetic fields (7 and 8.4T), and following study using in vivo a mouse brain confirmed the contrast due to paramagnetic deoxygemoglobin and demonstrated that the image contrast was dependent on the field strength of MR (Ogawa & Lee, 1990b). Subsequently, it has been demonstrated that the microvasculature of in vivo rat brain showed the image contrast by the blood oxygen
level, suggesting that BOLD contrast provide \textit{in vivo} real-time physiological information, which is blood oxygenation, using a method to accentuate the paramagnetic deoxyhemoglobin of venous blood (Ogawa, Lee, Kay, & Tank, 1990c), where they coined this phenomenon as the blood oxygenation level-dependent (BOLD) contrast, and suggested that BOLD contrast could be used to measure normal healthy brain function.

Following this, human brain function using MRI was investigated (Ogawa, Tank, & Menon et al., 1992; Kwong, Belliveau, & Chesler et al., 1992; Bandettini, Wong, & Hinks et al., 1992). Ogawa, Tank, and Menon et al. (1992) measured the MR contrast of human visual cortex and found that the MR contrast was found in the grey matter during the visual stimulation. Kwong, Belliveau, and Chesler et al. (1992) also demonstrated that blood oxygenation was detected in both animal and human brains using visual stimuli and motor paradigms, where the comparison of MR contrast between stimulation and the non-stimulation periods was calculated. This comparison of MR contrast between stimulation and non-stimulation periods becomes nowadays a standard format to report results in fMRI research. Bandettini, Wong, and Hinks et al. (1992) showed that local changes in blood oxygenation levels during a motor task in human brain, in that the time course of gradient-echo echo-planar imaging (EPI) was measured in the primary and sensory motor areas, and the MR signal was enhanced by a finger movement task.

BOLD contrast is, however, an indirect measurement of neural activity, which includes several inhomogeneous components coming from cerebral blood flow (CBF), cerebral blood volume (CBV), and cerebral metabolic rate of oxygen (CMRO\textsubscript{2}). The interpretation about neural activity measured by BOLD fMRI is not fully understood. Using positron emission tomography (PET), Fox and Raichle (1986) showed that cerebral blood flow (CBF) and cerebral metabolic rate of oxygen (CMRO\textsubscript{2}) in the resting
human brain was significantly correlated, and the regional uncoupling of CBF and CMRO$_2$ was found when somatosensory stimulation happened, resulting in the local change of CBF was much greater than tissue metabolic change. They proposed a hypothesis that CBF would be correlated to neuronal firing, but independent of CMRO$_2$. In following study, Fox, Raichle, Mintun, and Dence (1988) measured glucose uptake, oxygen metabolism, and blood flow in human brain, and on the basis of the result that neural activity increased brain glucose uptake and blood flow much more than oxygen consumption, the study suggested that the change of blood flow relating to neural activity could be due to other processes than oxygen metabolism.

Chen and Ogawa (2000) have pointed out the inhomogeneity of BOLD contrast, in that BOLD contrast includes the components from CBF, CBV, and CMRO$_2$ where an importance of the time course and BOLD contrast at the capillary level were emphasised although the correlation between BOLD contrast and CBF has been shown in various experiments. Bandettini and Ungerleider (2001) have also discussed that the biological basis and heterogeneity of BOLD contrast, in that the signal dynamics and magnitude of BOLD contrast result from the interplay of many variables, such as neural activity, metabolism, blood flow, blood volume, and oxygenation change. Raichle (2001) has suggested that BOLD signal may be increased by the processing of glutamate in astrocytes, which are non-neuronal cells after excitatory neurotransmission, while also pointing out that inhibitory neurotransmission regarding BOLD signal has not yet been investigated.

Due to these various components within BOLD contrast, the interpretation of BOLD contrast should also be carefully considered. For example, Logothetis and Wandell (2004) discussed the relationship between neural activity and BOLD contrast, suggesting that heterogeneity exists in hemodynamic response function (HRF) between
individuals, across the cortex and across sensory, motor, cognitive tasks. Despite the indirectness of BOLD contrast signal in fMRI, they proposed that fMRI method had the benefit of computational understanding and the networks of information processing in the brain. Bandettini (2009) also suggested that the interpretation of inferences of causality based on the timing differences less than 2 seconds, might include meaningless vasculature-influenced timing. Nevertheless, the correlation between BOLD contrast and neural activity has been shown in animal studies. Logothetis, Pauls, and Augath et al. (2001) demonstrated that neural activity was correlated with BOLD contrast signal in that the local field potentials (LFPs) evoked by the stimulus in the range of 30-150Hz were correlated with the BOLD response in the visual cortex of monkeys. Lee, Durand, and Gradinaru et al. (2010) also demonstrated that the specific stimulation of local CaMKIIα-expressing excitatory neurons elicited positive BOLD signals using by a novel method of optogenetic technology in living rodents.

Taken together, BOLD fMRI is an indirect measurement of neural activity, and the BOLD contrast signal in venous levels includes the inhomogeneous combinations of CBF, CBV and CMRO₂. Nevertheless, the correlations between BOLD contrast and neural activity, which have been shown in animal studies, suggests that BOLD contrast is currently the most robust and concurrent contrast mechanism to measure human brain function in terms of both spatial and temporal resolution. Thus, the interpretation of BOLD contrast should be considered with the magnitude and time series of BOLD contrast.
3.2. Experimental designs of fMRI

Linguist (2008) suggests a balanced two prerequisites in designing a valid fMRI experiment, which are optimising the statistical power of signal-to-noise (SNR) regarding BOLD contrast, and at the same time, satisfying psychological validity. Due to the signal of BOLD contrast is weaker than that of neuronal activity, it is a main challenging to detect MR signals corresponding to task relevant BOLD contrast during a task. In auditory fMRI experiment, it can be additionally difficult to get MR signals due to the loud MR scanner noise, which can be obstacles to listen to sound stimuli during MR scan. Having a temporal gap in the acquisition of MRI data to deliver sound stimuli without scanner noise, such as sparse sampling techniques in auditory fMRI (Hall, Haggard, & Akeroyd et al., 1999; Schwarzbauer, Davis, Rodd, & Johnsrude, 2006), has been suggested to overcome this problematic issue, but fMRI research on musical rhythm is highly dependent on time, therefore, it can be more informative to acquire whole time-series of fMRI data in music experiments.

This section describes conventionally used fMRI experiment designs of block and event related designs. The block design was used to present steady beat stimuli (Chapter 5), and the event-related design was used to present rapid, complex rhythm pairs (Chapter 4). Considerations of each design for each study are discussed.

**Block design**

Block design (or a boxcar design) has been widely used in both PET and fMRI experiments. As the peak of haemodynamic response of BOLD contrast delays 5 or 6s from the stimulus onset time, and declines for about 15s after the peak BOLD response), the BOLD response during the block is an average of all stimuli within the block, while
CHAPTER 4. Beat perception in complex rhythms

showing a strong statistical power. Due to this averaged BOLD response in the blocks, however, it is difficult to estimate the haemodynamic response of a single trial.

Block design can be useful in the experiment of rhythm perception when the experimental variances of continuous rhythmic stimuli are well controlled. As Amaro and Barker (2006) have explained, the block design is based on maintaining cognitive engagement in a task by presenting stimuli during blocks. For example, the perception of isochronous sequences with different types of metrical organisation can be easily designed by block conditions. Measuring emotional response or musical expectation in a listening task, which can be highly dependent on the time series of BOLD response may need to be undertaken within the framework of a block design.

Event related design

Block design has limitations in that averaging MR signal of all trials cannot show a difference between each single trial. Buckner, Bandettini, and O’Craven et al. (1996) demonstrated that a single trial of a single subject using fMRI showed a reliable signal, while measuring the BOLD response of the prefrontal cortex. In their experiment, BOLD response was acquired between the blocks and the single trial paradigm using a word-stem completion task. The comparison of MR signals between the two paradigms showed a highly correlated activation map. They suggested, that, in contrast to a block design (which offers the averaged time course of stimulus blocks), a single trial design could offer the content of the time course of a single trial, as a short temporal resolution (e.g., TR=1s, which means that EPI image is acquired at the 1-second time scale) showed the shift of the time course of a single trial between two regions.
A main benefit of event related design could be to detect the time series of BOLD response across individuals or different brain areas (Amaro & Barker, 2006). As discussed in the previous section of the principle of BOLD contrast, the time series of BOLD response can vary, with different cortices, or different tasks (Logothetis & Wandell, 2004). Although relatively short period differences (e.g., 2 seconds) can be influenced by non-neuronal inference (Bandettini, 2009), whether different brain regions or individuals show significantly different time series of haemodynamic response (e.g., more than three seconds) can be checked when the event-related design is used.

3.3. Region of interest (ROI) analysis using a cytoarchitectonic probability map

A whole brain analysis, thresholding techniques for a statistical map in accordance with experimental conditions, are conventionally used to report the results in fMRI analysis. When fMRI studies have strong neural hypotheses on certain brain regions, fMRI analysis to delimit brain regions can be performed (Poldrack, 2007; Etzel, Gazzola, & Keysers, 2009). Poldrack (2007) discussed the several purposes of ROI analysis: (1) to explore fMRI data, (2) to control Type I error that incorrectly rejects a true null hypothesis, and (3) to limit a region to test null hypotheses. It was also pointed out that ROI analysis is still developing, and that several options and methods can be varied at selecting ROIs and performing statistical analysis on fMRI data within the selected brain regions.

Caution on using ROI analysis can also be found, particularly several problems regarding functional localiser used to select local brain areas to test null hypothesis were discussed (Friston, Rotshtein, Sterzer, & Henson, 2006). Firstly, it was suggested that
ROI analysis might have a risk to overlook the true activation voxels, which were distributed in whole brain beyond the selected regions. Secondly, averaging MR signal within the ROIs might be compounded by deactivated voxels as well, resulting in weakening the signal-to-noise ratio (SNR) of stimulus condition. Finally, the anatomically constrained ROI can be problematic in terms of not having inter-subject variability, in which the brain activation of individuals would show different local maxima corresponding to tasks or stimuli. This study alternatively suggested to employ the first eigenvariate of an ROI as a temporal covariance to find coherent spatial mode in the ROI for the last issue, but generally suggested that not using ROIs would be ideal for fMRI data analyses.

The current thesis has a strong prior region of interest on bilateral IFG (BA44 and BA45), which assumed to be involved in beat and metre perception. To define ROIs of IFG, this thesis uses a cytoarchitectonic probability map, more specifically a maximum probability map. As discussed in the previous paragraphs, ROI analysis is still developing, and the methods are still flexible depending on the individual case, although several problems regarding selecting a priori region of the brain were discussed in Friston et al. the first problem brought up by Friston et al. can be supplemented by calculating the whole brain analysis before the ROI analysis, so that ROI analysis can be cooperating with a whole brain activation map without losing any significant activation sites. The second problem regarding averaging MR signal can also be supplemented by selecting a true signal within ROIs. For example, if statistical analyses are performed within the selected ROIs, and the only activated voxels corresponding to target conditions were averaged, the decreased SNR due to the noise voxels can be improved. Finally, the issues raised by this study regarding inter-subjectivity within ROIs can be supplemented by their suggestion, which uses the first eigenvariate of temporal sources.
to find spatial information of individuals. The last solution, however, seems, intuitively not be always available for fMRI analyses, as when the volumes of ROIs are not sufficiently large enough, the conventional smoothing process on the fMRI data may not allow a good spatial resolution to discrete individual subjectivity within the ROIs.

Thus, it will be important to have a reasonable approach to define ROIs for an individual study, considering both the pros and cons of the ROI analysis. This thesis uses a cytoarchitectonical probability map that is reasonably used in the current fMRI research in defining ROIs. In addition, not to have the problematic averaging issue, this thesis counts a whole brain activation map with the cytoarchitectonic map together, in that true activated voxels within the cytoarchitectonic probability map of IFG (BA44/45) are used to calculate the averaging of MR signal.

Following sections, a cytoarchitectonic probability map and its maximum probability map (MPM) are explained, and developing the methods to import MPM from SPM to BrainVoyager and to create the mask using the regions of IFG (BA44/45) of MPM for ROI analysis are described.

Cytoarchitectonic probability map

Cytoarchitectonics of the brain is the study to parse cerebral cortex into subdivisions that have structural differences at the cellular level. For example, Brodmann Areas, which refer different cytoarchitectonic areas in the cerebral cortex with numbers, is the most famous brain map (Petrides, 2014). The cytoarchitectonic brain map was based on a single adult brain, thus the brain map was dependent on the observation skills of histologists, remaining a problem to improve reliability to define the cytoarchitectonic borders in the brain.
CHAPTER 4. Beat perception in complex rhythms

Amunts, Schleicher, and Bürgel et al. (1999) reconstructed a cytoarchitectonic map of BA44 and 45 with 10 post-mortem brain structures using an observer-dependent method. This study showed the results of a statistical difference between two different laminar layers, for example, the asymmetries between the left and right BA44 region showed that the left BA44 had more density of cells compared with right BA44. The greater density of cytoarchitectonic structures in the left IFG (BA44) was suggested to be related to the functions of speech production. The idea that different cytoarchitectonic structures may have different brain functions was shown when cytoarchitectonic map was used to analyse fMRI data. For example, Amunts, Weiss, and Mohlberg et al. (2004) studied two tasks of verbal fluency and semantic tasks using fMRI, in that the results showed that the two tasks made could functionally dissociate the two subregions (BA44/45) in Broca’s area using the cytoarchitectonic probability map in fMRI analysis.

*Maximum probability map*

As discussed in the above, the individual variability on defining the borders of cytoarchitectonic probability map remains as a problem. A maximum probability map (MPM) is based on the cytoarchitectonic probability map, and offers discrete subregions of the cytoarchitectonic probability map without overlap (Eickhoff, Stephan, & Mohlberg et al., 2005; Eickhoff, Heim, Zilles, & Amunts, 2006). For example, 10% in cytoarchitectonic probabilistic map suggests that the selected cytoarchitectonic region was observed in one out of ten post-mortem brains. Similarly, 50% of cytoarchitectonic probabilistic map suggests that the selected brain region was observed in five out of ten post-mortem brains and 100% of cytoarchitectonic probability map suggests the selected brain region was observed in all ten post-mortem brains. Thus, the increase of probability means more commonly observed brain regions out of ten post-mortem
CHAPTER 4. Beat perception in complex rhythms

brains, and its size becomes smaller. The MPM embedded in SPM tool box shows all subregions of cytoarchitectonically defined brain map without overlap, of which probability is between 40% and 50%. Using MPM to define ROIs, such as IFG (BA44/45) is reasonable to select ROI as well as practically convenient.

Following section explains the developing methods to import MPM from SPM toolbox to BrainVoyager, and to create ROIs of IFG (BA44/45) using this imported MPM. In addition, potential methods to use MPM for ROI analyses are discussed.

3.4. Developing a method for ROI analysis

Preprocessing of Macrolabel images

BrainVoyager QX 2.8.0 (Brain Innovation, Maastricht, The Netherlands) was used for developing ROI analysis. Firstly macrolabel templates (Eickhoff et al., 2006), embedded in the extension of SPM toolbox, are written in an analyze format (i.e., the files including an image and header). These images were imported as DICOM format using BrainVoyager (BV) software. When the macrolabel templates were imported to BV, the information of anatomical orientation of the sagittal, coronal, and axial planes of the original macrolabel images was not matched with the default setting of the orientation of sagittal direction in BV. Thus, the directions of sagittal, coronal, and axial planes were manually rotated and adjusted to BV. It is assumed that the left and right directions of the macrolabel templates are based in a neurological way (i.e., the left indicates left hemisphere, and the right indicates right hemisphere), whereas BV conventionally uses a radiological way in the left and right directions as a default. The imported macrolabel images were transformed to the radiological direction first, in that the left side indicated right hemisphere and the right side indicated left hemisphere. To make sure the manual
CHAPTER 4. Beat perception in complex rhythms

rotation and adjustment was properly done, the structure images of Colin27, which are a high spatial resolution anatomical template from the Montreal Neurological Institute (MNI), were tested first, and then the transformation method used in Colin27 was applied to transform the macrolabel template in exactly the same way.

As a default template, Talairach coordinates was embedded in BV, and the macrolabel images were based on MNI templates. Thus, the imported macrolabel images were firstly normalised into Talairach coordinates. In the normalisation process, again Colin27 images were tested firstly, and the exactly the same method was applied to the macrolabel images. This is because that Colin27 has a high spatial resolution, which is easier to detect the transformation of brain structures, whilst the macrolabel images comprised of the large chunks coloured by the gray contrast, which is more difficult to detect the transformation of structures, such as the shift between left and right hemispheres.

For the normalisation of the Colin27 images, the structural images were aligned and transformed with a line of AC (anterior commissure) and PC (posterior commissure), using a cubic spline spatial interpolation. The aligned AC-PC Colin27 images were then normalised to a Talairach coordinate (Talairach & Tournoux, 1988), using sinc interpolation using the anatomical template installed with 12 sub-cuboids landmarks embedded in BV. The macrolabel images were normalised with the same procedure as that of Colin27. As mentioned in the previous paragraph, the macrolabel images have low resolution with rough segmentations combined, thus statistical processing was minimised in the imaging process, i.e., sinc interpolation was used.

The normalised macrolabel images showed colour-coded contrast differences between roughly segmented cytoarchitectonic areas. For example, black colour was coded as the number of 0 and white colour was coded as the number of 255, and this
gray contrast could be manually adjusted and numbered to discrete each subregions of the brain map using colour coding. BA44 and 45 of bilateral IFG could be segmented by the contrast differences coded as an integer value. After defining the contrast colour coding of each region of bilateral IFG areas (BA44 and 45), those regions were separately saved as three-dimensional volumes to be defined as ROIs (Figure 3.1.)

The Figure 3.2. showed that the transformation processing and the results of both Colin27 T1 and MPM images. The images imported showed inconsistent three-dimensional planes with BV in that the transverse, sagittal, and coronal planes of the Colin27 T1 and MPM images were corresponding to each the sagittal, coronal, and transverse planes of BV (Figure 3.1.a.). Thus, firstly the planes of the imported images were transformed to that of BV (Figure 3.1.b.). Then both images were normalised to Talairach coordinates (Figure 3.1.c.)
CHAPTER 4. Beat perception in complex rhythms

Figure 3.1. Image preprocessing of Colin 27 T1 and MPM templates (the above imaging process shows a radiological direction of left and right hemispheres, in that left side indicates right hemisphere, and right side shows left hemisphere).

The preprocessed macrolabel images showed the grey contrast differences for each subregion. BA44 and BA45 of bilateral IFG were defined as different colour coding, and each subregion was reconstructed as three-dimensional volume to be defined as ROIs. The red lines show the subregions of BA44 and BA45 in bilateral IFG (Figure 3.2. the left panel). The segmented subregions were then overlaid onto the Colin27 T1 images (Figure 3.2. the right panel) to confirm the spatial transformation, where the surface maps of a sagittal plane showed that the left IFG (BA44 and 45) on the left side, and the right IFG (BA44 and 45) on the right side.
The results showed that the MPM was successfully imported to BV and then the images were normalised to Talairach templates. The IFG (BA44/45) of MPM was successfully segmented using the different colour coding of the grey contrast, and the MPM-based ROIs for defining IFG (BA44/45) were created.

fMRI analyses can be performed in two steps. Firstly, the whole brain analysis using stimuli conditions might be calculated to get a whole brain activation map. Secondly, the same statistical analyses performed in the whole brain analysis can be applied within the selected ROIs (e.g., IFG), in which the mean percent signal change of BOLD response can be calculated with the only voxels that were activated in the whole brain analysis. For example, the selected ROIs join together with the whole brain activation map, resulting the conjunction areas, which are presumably smaller than the MPM based ROIs. Within this specified areas, mean percentage of BOLD response can be calculated, while reducing the noise voxels included.
CHAPTER 4. Beat perception in complex rhythms

Beyond the described ROI analysis using the MPM, there can be more various methodological approaches to define and use ROIs, however, the discussion on the further methodological consideration is beyond the discussion of this thesis.
Chapter 4. Beat perception in complex rhythms

4.1. Introduction

Perceiving or extracting temporal regularity during complex rhythm perception is key to understanding rhythmic behaviour in humans. In particular, as discussed in the literature review of Chapter 2, the IFG activation during rhythm perception is essential to understanding cognitive processing in the hierarchical structures of metric rhythms. fMRI studies focusing on the relationship between rhythm perception and motor movements have shown that the sensorimotor areas and IFG are involved in processing the different degrees of rhythmic complexity within metric simple, metric complex, and non-metric rhythms (Grahn & Brett, 2007; Chen, Penhune, & Zatorre, 2008a), whereas, large discrepancies has been found in both the sensorimotor areas and IFG activations. For example, compared to metric complex and non-metric rhythm conditions in a rhythm discrimination task, metric simple rhythm conditions showed the significantly greater activation in the left IFG (BA45) with the activation in STG and putamen (Grahn & Brett, 2007), suggesting that generating internal beats in a metric simple rhythm condition significantly activated the brain regions. On the other hand, rhythmic complexity in a rhythm reproduction task has been found to modulate the activation in the dPMC (Chen, Penhune, & Zatorre, 2008a). Here, a metric complex rhythm condition has been found to activate the right IFG (BA44/45) in musicians group only, suggesting that working memory may activate the right IFG due to top-down processing of musicians.

The discrepancy with respect to the relationship between IFG activation and rhythmic complexity in the two studies can be considered as resulting from several factors, such as different tasks and rhythmic stimuli. For example, a rhythm
CHAPTER 4. Beat perception in complex rythms
discrimination task (Grahn & Brett, 2007) and a rhythm reproduction task (Chen, Penhune, & Zatorre, 2008a) may require different cognitive strategies during encoding of rhythmic patterns, whereas the reproduction task may need additional cognitive processing to plan movements. Furthermore, the choice of presentation style of rhythmic stimuli can result in different brain activation patterns in the IFG. In contrast to each rhythm trial used in Grahn and Brett (2007), one specific rhythm for each metric simple, metric complex, and non-metric rhythm condition was presented in Chen, Penhune, and Zatorre (2008a), which may result in different degrees of exposure to the rhythmic stimuli.

The different tasks and presentation style of rhythmic stimuli are, however, experimental variables that are dependent on the specific studies that have been undertaken into cognitive complexity. Complexity across music and non-music domains has been suggested to increase IFG activation consistently regardless of such experimental variables. For example, the increase of syntactic complexity in sentence processing has been found to increase the activation in the left IFG regardless of the different languages, tasks, and stimuli used (Caplan, Alpert, Waters, & Olivieri, 2000; Obleser, Meyer, and Friederici, 2011; Rogalsky, Matchin, & Hickok, 2008). When harmonic musical sequences were violated, the bilateral IFG were activated (Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch, Gunter, & von Cramon et al., 2002; Tillmann, Janata, & Bharuch, 2003), suggesting that the increase of complexity due to the perception of unexpected events in music listening, such as oddball-like or mismatched sounds, increases IFG activation in both hemispheres. Overall, complexity increases the activation in IFG.

In contrast to the consistently increased activation in IFG in accordance with the increase of complexity mentioned above, either the salient metrical organisation (Grahn
CHAPTER 4. Beat perception in complex rhythms

& Brett, 2007) or the proficiency of musicians in rhythm perception (Chen, Penhune, & Zatorre, 2008a) that has been found to increase the IFG activation suggests, however, that rule-based, hierarchical cognitive processing modulates the activation of IFG. This rhythmic complexity among the three different types of rhythm conditions is perhaps incomparable. For example, the metric simple rhythm condition comprises of a series of arbitrary integer units (i.e., each unit or a subdivision of beat has a duration of between 200ms and 250ms), which are hierarchically grouped into beats based on the metrical organisation of every 3 beats (Chen, Penhune, & Zatorre, 2008a, 2008b) or every 4 beats (Grahn & Brett, 2007). The metric simple rhythm condition induces hierarchical cognitive processing in relation to the metrical organisation between the subdivisions of beat and beat level. In contrast to metric simple rhythm condition, metric complex rhythm conditions also comprise of a series of arbitrary integer units, but when the integer units are grouped the temporal patterns of metrical organisation become unclear or irregular due to the omitted accent on the beat timing, although the temporal relations between integer units at the subdivisions level of beat are isochronous. In the metric complex rhythm condition, the perceptual difference between the temporal irregularity (due to the disrupting metrical organisation at beat level) and the temporal regularity (due to the regular temporal relations between integer units at subdivisions of beat level) may cause difficulty during rhythm perception. Non-metric rhythm conditions have neither metrical organisation at the beat level nor the temporal regularity in the subdivisions of beat level.

The latter two conditions may disrupt temporal regularity, in particular, when the sense of pulse exists. This means that if the sense of pulse is not generated at all (i.e., non-metric rhythm conditions), the disruption regarding the sense of temporal regularity may be less than that of metric simple or metric complex rhythm condition. Thus, the
perception of the metric complex rhythm condition and/or non-metric rhythm condition may result in oddball-like or mismatched auditory perception, where the temporal expectations of temporal regularity are violated, which has been understood in a similar way to the above violated harmonic sequences.

To minimise such oddball-like or mismatched auditory perception, it will be necessary to modulate rhythmic complexity within the metric rhythm condition. This will allow the understanding of the neural correlates that correspond to the hierarchical cognitive processing of metrical organisation during rhythm perception. The current fMRI study employed beat-based, metric rhythms for all conditions. Recently, perceiving temporal regularity during rhythm perception has been found to activate the basal ganglia (BG) across music and non-music domains (Grahn & Brett, 2007; Grahn & Rowe, 2009; Geiser, Notter, & Gabrieli, 2012). The BG activation during metric rhythm condition may supplement that the idea that the different degrees of rhythmic complexity within metric rhythm condition may be different from the violation of temporal regularity in previous studies, such as metric complex or non-metric rhythm conditions.

In the current thesis, the rhythmic complexity in terms of the frequency of accents to attend to is modulated using additional volume accents, resulting in either enhanced or disrupted metrical organisation of a binary beat pattern (i.e., Beat Accented and Non-beat Accented rhythm conditions; the rhythmic complexity in terms of the amount of information is consistent across all beat-based, metric rhythms). Unaccented rhythm condition is a control condition. The subdivisions of beat in rhythm conditions consist of duplets and triplets of beat (see Methods). Non-isochronous combination of the duplet and triplet of beat increases is designed to increase the ambiguity in beat-based temporal structures, which consequently increase the effect of volume accents during rhythm perception. Thus, although all conditions of Unaccented, Beat Accented,
and Non-beat Accented rhythms are made up of the same rhythmic structures, volume accents are employed to amplify the perception of enhanced and disrupted binary patterns. It was hypothesised, firstly, that all beat-based rhythm conditions would activate the auditory and sensorimotor areas and bilateral IFG. Secondly, it was hypothesised that Non-beat Accented compared to Beat Accented rhythm conditions would increase the activity of left IFG.

4.2. Methods

Participants
Fifteen participants with no history of neurological, psychiatric disorders, or hearing difficulties were recruited. One participant was excluded from the fMRI data analysis because behavioural performance during fMRI scanning was at the chance level (the accuracy was fifty percent), which might suggest that the participant neither could fully understand the task nor could not manage to attend the fMRI session properly. The remaining fourteen participants (mean age 30.28 years; s.d. 9.46; range 19 – 49 years; 6 males; 0 to 5 (mean 1.00, s.d. 1.51 years of musical training; for the details of participant information, Table 4.1.) were all right-handed (N=14, mean L.Q. = + 83.62, s.d. = 16.73, Decile R.6, The Edinburgh Handedness Inventory, Oldfield, 1971). The study was approved by both the Local Research Ethics Committee (West of Scotland Research Ethics Service (WoSRES), REC reference: 12/WS/0229) and the University of Edinburgh Ethics Committee (Edinburgh College or Art). The participants gave fully informed written consent of their willingness to participate, and were compensated with minimum travel expenses with no fixed compensation for their participation.
Table 4.1. Participant demographics

<table>
<thead>
<tr>
<th>No. of participants</th>
<th>Gender (F: female, M: male)</th>
<th>Age</th>
<th>Music education</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>33</td>
<td>None</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>45</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>38</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>F (8), M (6)</td>
<td>Avg 30.28</td>
<td>Avg 1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(s.d. 9.46)</td>
<td>(s.d. 1.51)</td>
</tr>
</tbody>
</table>

Study Design

Participants listened to rhythm pairs during fMRI scan, where the scan measured blood oxygenation level-dependent (BOLD). The rhythm pairs were sequentially heard with a 1s silence gap between each two rhythms. Rhythmic stimuli of Unaccented, Beat Accented, and Non-beat Accented rhythms were presented in random order using a slow event related design. Rhythm pairs were either the same or different. The participants indicated whether the rhythms were the same or different by pressing buttons after listening to a series of rhythm pairs. A very simple task and button response was designed to avoid the limitations of a passive listening task, and was intended not to bring cognitive difficulty. The behavioural responses of the button press were collected, and the accuracy of the task performance was considered to check whether the participants understood and performed the task properly during the MR scan. The
response time was not considered because the participants had not been instructed to perform the task as quickly as they could.

Stimuli and task

All sound stimuli were beat-based rhythms, and rhythmic complexity in accordance with temporal regularity was modulated by volume accents in which Beat Accented and Non-beat Accented rhythms were created. Unaccented rhythms were prepared as a control condition, and had no volume accents (Figure 4.1.). For example, all the rhythm conditions consisted of 4 beats, and the interval between beats was 600ms, an interval that has been used in previous empirical studies (see Figure 4.1., Parncutt, 1994; Drake, Jones, & Baruch, 2000). A beat has the subdivisions of either duple or triple. Duple and triple subdivisions were created by equally subdividing the duration of beat by two or three parts (i.e., the subdivision of duple has the duration of 300ms and the subdivision of triple has the duration of 200ms). 10dB louder volume accents were used to differentiate between Accented and Unaccented sound. Previously, loudness differences between 6 and 10 dB were reasonably distinguished during the auditory fMRI experiment (Chen, Zatorre, & Penhune, 2006).

Due to the additional volume accents on Beat Accented and Non-beat Accented rhythms compared to Unaccented rhythms (Figure 4.1.), the average intensity of sound stimuli (which can be described as a root mean square (RMS)) of both Beat Accented and Non-beat Accented rhythms was greater than Unaccented rhythms. The unaccented note was -16dB, whereas accented note was -6dB, so the RMS of Beat Accented and Non-beat Accented rhythms was -10dB (peak 0dB), and RMS of Unaccented rhythms was -16dB. Audio sample rate was 44,100Hz and audio bit depth was 16 bits with mono
Beat perception in complex rhythms

Audio channel (Sound Forge pro 10, Sony Creative Software, Middleton, WI, USA). Beat Accented rhythms had the volume accents on each first and third beat, and Non-beat Accented rhythms had the volume accents on the second subdivision of each first and third beat, of which positions were off-beats. A marimba sound with the same pitch (Cubase software, Stenberg, Germany) was used for all rhythmic stimuli.

<table>
<thead>
<tr>
<th>Music notation</th>
<th>Sound wave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaccented rhythms</td>
<td><img src="image1" alt="Unaccented Rhythms" /></td>
</tr>
<tr>
<td>Beat Accented rhythms</td>
<td><img src="image2" alt="Beat Accented Rhythms" /></td>
</tr>
<tr>
<td>Non-beat Accented rhythms</td>
<td><img src="image3" alt="Non-beat Accented Rhythms" /></td>
</tr>
</tbody>
</table>

**Figure 4.1.** Music notations and the corresponding sound waves of Unaccented, Beat Accented, and Non-beat Accented rhythm conditions.

A total of 12 rhythm pairs were prepared for the three conditions of Unaccented, Beat Accented, and Non-beat Accented rhythms (Figure 4.2.). The number of trials of the same and different rhythm pairs was the same (i.e., 50 percent of the same rhythm pairs, and another 50 percent of the different rhythm pairs). Different rhythm pairs were created by switching the order of two subdivisions of beats in the middle of rhythms. For example, the duplet and triplet of the second and third beats were switched with each other when different rhythm pairs were created; the subdivisions of the first beat.
remained triplet, and the last beat did not have any subdivision. Thus, a discrimination task to judge the same or different rhythm pairs mainly depends on listening to the subdivisions of the second and third beats of rhythms. The last beat was intended to give the sense of ending during rhythm perception to the participants, so it was intended to design any subdivision of the last note, which was followed by a rest of the same duration of the last beat, i.e., 300ms. All twelve pairs of rhythms were repeated four times during MR scan, and 48 trials were presented in total (i.e., 4 rhythm pairs × 3 conditions of Beat Accented, Unaccented and Non-beat Accented rhythms × 4 repetitions).
CHAPTER 4. Beat perception in complex rhythms

Training Procedure before fMRI scan

Participants practiced a training session immediately prior to MR scan. This training session aimed to enable the participants to be familiar with rhythmic stimuli, the task, and the environment of fMRI experiment. A laptop, headphones and a picture of the MR compatible four-button box were given to participants in an MR preparation room. An

<table>
<thead>
<tr>
<th>Rhythm pairs (S, same; D, different)</th>
<th>Rhythm 1</th>
<th>Rhythm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaccented rhythm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (S)</td>
<td><img src="image1.png" alt="Music notation" /></td>
<td><img src="image2.png" alt="Music notation" /></td>
</tr>
<tr>
<td>2 (D)</td>
<td><img src="image3.png" alt="Music notation" /></td>
<td><img src="image4.png" alt="Music notation" /></td>
</tr>
<tr>
<td>3 (S)</td>
<td><img src="image5.png" alt="Music notation" /></td>
<td><img src="image6.png" alt="Music notation" /></td>
</tr>
<tr>
<td>4 (D)</td>
<td><img src="image7.png" alt="Music notation" /></td>
<td><img src="image8.png" alt="Music notation" /></td>
</tr>
<tr>
<td>Beat Accented rhythm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (S)</td>
<td><img src="image9.png" alt="Music notation" /></td>
<td><img src="image10.png" alt="Music notation" /></td>
</tr>
<tr>
<td>6 (D)</td>
<td><img src="image11.png" alt="Music notation" /></td>
<td><img src="image12.png" alt="Music notation" /></td>
</tr>
<tr>
<td>7 (S)</td>
<td><img src="image13.png" alt="Music notation" /></td>
<td><img src="image14.png" alt="Music notation" /></td>
</tr>
<tr>
<td>8 (D)</td>
<td><img src="image15.png" alt="Music notation" /></td>
<td><img src="image16.png" alt="Music notation" /></td>
</tr>
<tr>
<td>Non-beat Accented rhythm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 (S)</td>
<td><img src="image17.png" alt="Music notation" /></td>
<td><img src="image18.png" alt="Music notation" /></td>
</tr>
<tr>
<td>10 (D)</td>
<td><img src="image19.png" alt="Music notation" /></td>
<td><img src="image20.png" alt="Music notation" /></td>
</tr>
<tr>
<td>11 (S)</td>
<td><img src="image21.png" alt="Music notation" /></td>
<td><img src="image22.png" alt="Music notation" /></td>
</tr>
<tr>
<td>12 (D)</td>
<td><img src="image23.png" alt="Music notation" /></td>
<td><img src="image24.png" alt="Music notation" /></td>
</tr>
</tbody>
</table>

**Figure 4.2.** Music notations of all rhythmic stimuli of Unaccented, Beat Accented, and Non-beat Accented rhythm conditions (a pair of rhythm comprises Rhythm 1 and 2).
experimenter confirmed that the participants fully understood both rhythmic stimuli and task, and then the experimenter gave about 5 minutes to the participants to practice by themselves. They performed a discrimination task of 24 trials, in which rhythm pairs were the same rhythms that were used for the fMRI experiment. After finishing the training session, the participants were guided to a mock MR scanner room, where the experimenter introduced the environment of an MR scan, and the details of fMRI experimental procedures (see Appendix 3. Scripts for the protocol of two fMRI studies).

*fMRI scan*

Radiographers guided participants to the MR room, and assisted them to get ready to perform the experiment in a safe and comfortable posture. The participants were given earplugs to wear inside of MR compatible noise cancelling headphones. The participants were then given a 1-minute MR practice session, which was to familiarise them with the scanner noise and to check the comfort of the participants with fMRI apparatus, such as headphones, goggles, or a button response. The experimenter confirmed that all participants could hear the sound stimuli over the noise of fMRI scan. During an fMRI session for 17 minutes 25 seconds, rhythmic stimuli were presented using a slow event related fMRI design (Figure. 4.3.), which was used to maximise separation of the hemodynamic responses between trials. Each rhythm pair was played within 6 seconds, followed by fixation for 15 seconds. The first and last 15s blocks were fixation, and the presenting order of rhythmic stimuli of all conditions was randomised.
**Data acquisition and MR apparatuses**

MR data was obtained at the Clinical Research Imaging Centre (CRIC), The University of Edinburgh, using a 3T Verio whole-body MRI system (Siemens Medical Systems, Erlangen, Germany). A quadrature birdcage RF coil was used for RF transmission and reception. First, high-resolution 3D anatomical images covering the whole brain were obtained using an MPRAGE sequence (TR 2300 ms; TE 2.98 ms; voxel size 1 mm; FOV 240 by 256 mm; flip angle 9°) encompassing the whole cerebrum parallel to the anterior commissure (AC) – posterior commissure (PC) plane. Functional images were obtained using a blood oxygenation-level dependent (BOLD) contrast by gradient-echo echo-planar T2*-weighted image (EPI) sequence prescribed at the same orientation (TR 3,000 ms; TE 30 ms; flip angle 90°; pixel size 3 mm, FOV 192*192 mm; slice thickness 3.75mm; 36 slices; total volume 345). Conventional Siemens noise cancelling MR compatible headphones with earplugs and the goggles of the Nordic MR visual system.
(www.nordicneurolab.com) were used to convey sound and visual stimuli to participants in the MR scanner. Presentation programming software (www.neurobs.com) was used to synchronise visual and sound stimuli on a desktop computer. The Nordic system was used to synchronise MR sequence and Presentation software of the desktop. An MR compatible four-response button box was used (www.curdes.com).

Data analysis

Behavioural data analysis

The accuracy was calculated from the button response, which was collected from a discrimination task during MR scan. Mean percent of the accuracy for each condition was calculated, and then, a one-way analysis of variance (ANOVA) within-subjects using SPSS 22 (IBM SPSS statistics) was conducted, to examine whether or not there was any significant difference depending on the conditions of Beat Accented, Unaccented, and Non-beat Accented rhythms. As we did not ask participants to perform a task as quickly as possible, reaction time was not considered.

Functional Magnetic Resonance Imaging (fMRI) data processing

Preprocessing of structural and functional images

BrainVoyager QX 2.8.0 (Brain Innovation, Maastricht, The Netherlands) was used for fMRI data analysis. To preprocess structural images, firstly, the skull parts of individual anatomical images (i.e., T1 MPRAGE) were peeled, and the structural images were corrected by intensity inhomogeneity. The images were aligned and transformed with a line of AC (anterior commissure) and PC (posterior commissure), using a cubic spline
spatial interpolation. The aligned AC-PC anatomical images were then normalised to Talairach coordinates (Talairach & Tournoux, 1988), using sinc interpolation using the anatomical template installed with 12 sub-cuboids landmarks in BrainVoyager QX. To preprocess functional images, the first two volumes of functional images were discarded to exclude T1 saturation effects. Motion detection was performed by trilinear interpolation, and motion correction by sinc interpolation, using by 6 rigid motion parameters (3 translation and 3 rotation of x, y, z parameters). Less than 4 mm motion in all x, y, and z directions of all participants was checked. Slice time correction was performed and the data were re-sampled by an ascending interleaved order and cubic spline interpolation. High pass temporal filtering with a Fourier basis set of 2 sines/cosines, including a linear trend, was performed. To co-register structural and functional images, functional images were firstly registered to their own native anatomical images using a header-based initial alignment (IA) for its proximity of global orientation, and a fine-tuning alignment (FA) by Gradient-driven affine transformations (6 - 12 parameters). Individual functional images were then normalised to Talairach coordinates and resampled using trilinear interpolation. The normalised functional images were then spatially smoothed with an isotropic Gaussian filter with FWHM 6 mm.

Statistical analysis of functional data

In the first level of individual analysis, General Linear Models (GLMs) of three conditions were modelled: Unaccented (a control condition), Beat Accented and Non-beat Accented rhythms. As slow event related design of the current experiment used a fixed period of 15s fixation, which is enough to separate a canonical haemodynamic response (with time to response peak 5s and time to undershoot peak 15s) between trials,
CHAPTER 4. Beat perception in complex rhythms

the calculation of a block design of the duration of 6s was used. Motion parameters estimated during the realignment (x, y, z parameters in 3 rotations and 3 translations) were modeled in single subject GLMs. In the second level of group data analysis, a GLM for group data was estimated using a random effect; three conditions modelled in a single subject GLM and motion parameters of individuals were included as residuals. For example, the fixation period was modelled as the constant 0 and each condition of Unaccented, Beat Accented, and Non-beat Accented rhythms was modelled as the constant 1. Serial correlation detection was performed by removing AR (2) and refitting the GLM.

At the second level of group analysis, basic contrast maps were calculated. Firstly, all rhythm conditions relative to fixation were calculated to obtain an overview of the brain activity in the experiment in that BOLD response of BG was checked to confirm that BG was activated by the temporal regularity of Unaccented, Beat Accented, and Non-beat Accented rhythms. Next, this activation map (i.e. all rhythm conditions relative to fixation) was used as a mask for following further contrast maps between conditions in a whole brain analysis. Using this mask with the whole brain activation map (i.e., all rhythm conditions relative to fixation), activation blobs outside of the skull or activations due to eye movements were disregarded for statistical analyses.

Continue from the basic contrast map in the above, additional basic contrast maps in the whole brain analyses were calculated. Firstly, due to volume accents, there was the loudness difference between Accented rhythm conditions of Beat Accented and Non-beat Accented rhythms and Unaccented rhythm condition. To examine loudness effects on neural differences, the contrast maps between Accented rhythms and Unaccented rhythms were calculated (i.e., Beat Accented rhythms > Unaccented rhythms and Non-beat Accented rhythms > Unaccented rhythms). Secondly, in
accordance with the hypothesis of the current experiment, the positions of volume
accents between Beat Accented and Non-beat Accented rhythms would show neural
differences; therefore, direct comparisons between Beat Accented and Non-beat
Accented rhythms were calculated (i.e., Beat Accented > Non-beat Accented rhythms
and Non-beat Accented > Beat Accented rhythms).

Region of Interest (ROI) analyses to examine neural hypotheses on the areas of
BG and IFG were performed. ROIs were basically extracted from the whole brain
activity map of all rhythm conditions versus fixation. The ROIs of IFG were defined as
the common brain regions between a maximum probability map for bilateral IFG (MPM,
BA44/45, Eickhoff et al., 2006) and the activation mask of all rhythm conditions relative
to fixation in a whole brain analysis. Within those common brain regions in both the
activation areas (by the whole brain analysis and by a MPM-based IFG region),
subtraction analyses between Beat Accented and Non-beat Accented rhythm conditions
were performed (i.e., Beat Accented > Non-beat Accented rhythms and Non-beat
Accented > Beat Accented rhythms).

fMRI analyses incorporated a correction for multiple comparisons based on the
false discovery rate (FDR) at the voxel level as described by Genovese et al. (2002).
Percent BOLD signal change was calculated by averaging the time series of activated
voxels for spheres of 10 voxels radius centered on the peak voxel of the relevant clusters
of activation. In the contrast maps between conditions in whole brain analyses (e.g.,
Non-beat Accented rhythms > Unaccented rhythms, and Non-beat Accented rhythms >
Beat Accented rhythms) the uncorrected threshold of \( p < .001 \) at the voxel levels was
used, and then the surviving voxels by a thresholding at a cluster–level threshold
estimator using by Monte-Carlo calculation with 10,000 iterations to correct multiple
comparisons (Forman, Cohen, & Fitzgerald et al., 1995), which is embedded in BV was additionally performed.

In addition, to check whether repetition suppression effects happened between the same and different rhythm pairs in each condition with respect to a predictive coding model, mean percentage signal change in BOLD contrast of each the same and different rhythm pairs were further calculated. Brodmann’s areas (BA) were defined in Talairach and Tournoux (1988) for all ROIs, and the activation of IFG was cytoarchitectonically defined using an MPM.
4.3. Results

*Behavioural data collected in a discrimination task*

The behavioural data collected in a discrimination task during fMRI scan showed that the mean percent of the accuracy was over 80% for all rhythm conditions (Figure 4.4.). A one-way ANOVA across all rhythm conditions showed that there was no significant main effect resulting from volume accents ($F_{2, 26} = 0.507$, $p = .608$; the percent accuracy of Beat Accented rhythms = 82.59%, Unaccented rhythms = 85.27%, and Non-beat Accented rhythms = 81.70%).

![Figure 4.4. Behavioural results in a discrimination task during fMRI scan](image-url)
CHAPTER 4. Beat perception in complex rhythms

*fMRI experimental results*

*Brain regions in rhythm perception*

All rhythm conditions relative to fixation activated the primary and secondary auditory areas, SMA, PMC, BG (putamen and globus pallidus), anterior insula, and IFG in both hemispheres and the left cuneus (Figure. 4.5. \( p \text{(FDR)} < .05; \) Table 4.2.). This brain activity map included all brain areas responding to all rhythm conditions. The map was used to create a comprehensive activation mask for further analyses in the current experiment.

![Brain activation during all rhythm conditions of Unaccented, Beat Accented, Non-beat Accented rhythm conditions.](image)

**Figure 4.5.** Brain activation during all rhythm conditions of Unaccented, Beat Accented, Non-beat Accented rhythm conditions.

From the activation map by all rhythm conditions relative to fixation, the activation of BG (putamen) was examined in terms of BOLD response. As temporal regularity in rhythm perception has been suggested to activate BG, its BOLD response
CHAPTER 4. Beat perception in complex rhythms

would activate in Unaccented and Beat Accented rhythms. The mean percent signal change in the BOLD response in bilateral BG showed that the activation of BG in all rhythms conditions showed no significant difference (Figure. 4.6.).

![Image: Figure 4.6. The mean percent change in BOLD response in the bilateral BG activation during each rhythm condition.](image)

**Accented relative to Unaccented rhythm conditions**

Accented rhythms, such as Beat Accented and Non-beat Accented rhythms, were created by the addition of volume accents on Unaccented rhythms. To investigate whether there was a significant difference in neural activity between modified conditions and a control condition, subtraction analyses between conditions were performed. When the Beat Accented rhythm condition was compared to Unaccented rhythm condition (i.e., Beat Accented rhythms > Unaccented rhythms), no significant activation was observed (e.g., statistical threshold; \( p \) (uncorrected) < .001). When the Non-beat Accented rhythm
CHAPTER 4. Beat perception in complex rhythms

condition was compared to the Unaccented rhythm condition (i.e., Non-beat Accented rhythms > Unaccented rhythms), left superior temporal gyrus (STG, Talairach coordinates, (-57, -15, 3)) was activated (Figure. 4.6. \( p \text{(uncorrected)} < .001 \), cluster-level threshold estimator \( p \text{(alpha)} < .05 \), minimum voxel number-16 voxels). The mean percent change in the BOLD response in the left STG was calculated and it showed a significant difference in BOLD responses between Non-beat Accented and Unaccented rhythms (two tailed paired t-test, \( p < .001 \)), whereas no significant difference in BOLD responses between Beat Accented and Unaccented rhythms was confirmed.

![Figure 4.7](image)

**Figure 4.7.** Left STG activation during Non-beat Accented rhythm condition compared to Unaccented rhythm conditions (bar graph denotes standard errors of the mean (±SEM)).
Beat Accented compared to Non-beat Accented rhythm conditions

To investigate the different neural activities between Beat Accented and Non-beat Accented rhythms, the direct subtraction analyses between the two conditions were calculated. When Beat Accented rhythm condition was compared to Non-beat Accented rhythm condition (i.e., Beat Accented > Non-beat Accented rhythms, \( p \) (uncorrected) < .001), there was no significant activation. In contrast, when Non-beat Accented rhythm condition was compared to Beat Accented rhythm condition, the activations were significantly left lateralised, including left IFG (BA44), left PMC (BA6), SMA, thalamus, left STG, and left occipital lobe (Figure 4.7. \( p \) (uncorrected) < .001, cluster-level threshold estimator \( p \) (alpha) < .05, k-16 voxels). The activations of thalamus and occipital area, however, showed that BOLD responses were below baseline, and the so-called negative BOLD response was not considered in the current experiment. The left IFG (BA44) by this subtraction map reflects that Non-beat Accented rhythms activated the region significantly greater than Beat Accented rhythms.

Figure 4.8. The left lateralised brain activation during Non-beat Accented rhythm condition compared Beat Accented rhythm condition.
CHAPTER 4. Beat perception in complex rhythms

**ROI analysis of IFG**

To investigate the role of IFG in rhythm perception, firstly ROIs were prepared by the common areas between bilateral IFG activation clusters in a whole brain analysis (i.e., all rhythm conditions relative to fixation) and cytoarchitectonically defined MPM-IFG (BA44/45) region. Within those ROIs, subtraction analyses between Beat Accented and Non-beat Accented rhythm conditions were calculated, which were also to specify the activation within IFG region, which would be related to the perception of the volume accent patterns of Beat and Non-beat. Beat Accented rhythms compared to Non-beat Accented rhythms (i.e., Beat Accented > Non-beat Accented rhythms) did not show any significant activation in ROIs, which was consistent with the results from the whole brain analysis. Non-beat Accented rhythms compared to Beat Accented rhythms (Non-beat Accented > Beat Accented rhythms) showed the activation of left IFG (Figure 4.10. Talairach coordinate (-45 11 16)). Cytoarchitectonically defined IFG regions (MPM-IFG, BA44/45) were overlaid on the left IFG activation, and the activation was defined as BA44.

![Non-beat Accented > Beat Accented rhythms](image)

**Figure 4.9.** The left IFG (BA44) activation defined ROI analysis during Non-beat Accented rhythm condition compared to Beat Accented rhythm condition.
The mean percent change in the BOLD response in left IFG was calculated for all rhythm conditions, and it showed that the Non-beat Accented rhythm condition activated the left IFG significantly greater than the Beat Accented rhythm condition (Figure 4.11.). There was no significant difference between Beat Accented and Unaccented rhythm conditions.

**Figure 4.10.** The mean percent change in BOLD response in the left IFG (BA44) activation (bar graph denotes standard errors of the mean (±SEM)).

**BOLD contrast between the same and different rhythm pairs**

To check repetition suppression between same and different rhythm pairs, of which rhythm pairs were designed as fifty and fifty percents of total rhythmic stimuli, BOLD response was calculated separately on each same and different rhythm pair. The left STG showed that in Beat Accented rhythm conditions, the same rhythm pairs were significantly suppressed compared to different rhythm pairs. In the left IFG (BA44)
activation, the repetition suppression by the presentation of the same rhythm pairs was not significant in all conditions.

Figure 4.11. The mean percent change in BOLD response in the left STG and left IFG (BA44) during each rhythm condition of the same and different rhythm pairs.

The results from both whole brain analyses and ROI analysis are summarised in Table 4.2. Brain activations in the table are activation clusters, and the peak voxel of each cluster is indicated. Brodmann’s areas, t-value, and p-value were defined using the peak voxel. The number of voxels of clusters in the Table 4.2. was reported after removing interpolation. Thus, the number of voxels corresponding to the anatomical high resolution of T1 images (1×1×1 mm³) was manually divided by 27 to match with the resolution of functional images (3×3×3 mm³).
Table 4.2. Stereotaxic coordinates of brain regions defined by both whole brain and ROI analyses

<table>
<thead>
<tr>
<th>Brain Regions (Brodmann’s areas)</th>
<th>Talairach coordinates (peak voxel of clusters)</th>
<th>t-value (avg. cluster)</th>
<th>p-value (avg. cluster)</th>
<th>Cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All rhythm conditions (&gt; fixation) in whole brain analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R precentral gyrus (BA43) / Superior Temporal gyrus (BA22)</td>
<td>46</td>
<td>-15</td>
<td>5</td>
<td>5.268</td>
</tr>
<tr>
<td>L Transverse Temporal gyrus (BA42)</td>
<td>-56</td>
<td>-16</td>
<td>11</td>
<td>5.084</td>
</tr>
<tr>
<td>L SMA (BA6)</td>
<td>-5</td>
<td>8</td>
<td>50</td>
<td>4.241</td>
</tr>
<tr>
<td>R IFG (BA9) / R PMC (BA6)</td>
<td>42</td>
<td>7</td>
<td>23</td>
<td>4.149</td>
</tr>
<tr>
<td>L PMC (BA6)</td>
<td>-50</td>
<td>-7</td>
<td>43</td>
<td>3.991</td>
</tr>
<tr>
<td>R insula (BA13)</td>
<td>27</td>
<td>20</td>
<td>4</td>
<td>4.038</td>
</tr>
<tr>
<td>L IFG (BA44/45) / L insula (BA13)</td>
<td>-29</td>
<td>26</td>
<td>5</td>
<td>3.906</td>
</tr>
<tr>
<td>L cuneus (BA19)</td>
<td>-11</td>
<td>-88</td>
<td>29</td>
<td>3.829</td>
</tr>
<tr>
<td>R putamen/globus pallidus</td>
<td>18</td>
<td>8</td>
<td>5</td>
<td>3.672</td>
</tr>
<tr>
<td>L putamen/globus pallidus</td>
<td>-20</td>
<td>4</td>
<td>5</td>
<td>3.670</td>
</tr>
<tr>
<td><strong>Non-beat Accented Rhythms &gt; Unaccented Rhythms in whole brain analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Superior temporal gyrus (STG)</td>
<td>-57</td>
<td>-15</td>
<td>3</td>
<td>4.752</td>
</tr>
<tr>
<td><strong>Non-beat Accented &gt; Beat Accented rhythms in ROI analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L IFG</td>
<td>-45</td>
<td>11</td>
<td>16</td>
<td>4.761</td>
</tr>
</tbody>
</table>

† The original number of voxels by T1 resolution (1×1×1 mm³) was manually divided by 27 to match with the resolution of functional images (3×3×3 mm³)

4.4. Discussion

The current fMRI experiment aimed to assess the activation of IFG in the perception of beat-based, metric rhythms, where rhythmic complexity regarding the different positions to attend to during rhythm perception were modulated by additional volume accents, resulting in Unaccented, Beat Accented, and Non-beat Accented rhythm conditions.
Participants listened to a rhythm pair, and judged whether the rhythms were the same and different. Behavioural results in this discrimination task during fMRI scan showed that the mean percent of the accuracy was over 80 percent for each condition, and there was no significant difference across all conditions, suggesting that the difficulty of task performance was controlled at a similar level across conditions.

fMRI results showed, firstly, that all rhythm conditions relative to fixation activated the auditory and sensorimotor areas and the IFG, including the primary and secondary auditory areas, SMA, PMC, BG (putamen and globus pallidus), anterior insula, IFG in both hemispheres, and the left cuneus. (Figure 4.5.). The results support the previous research findings in rhythm studies that showed activation in auditory and sensorimotor areas during rhythm perception. The mean percentage in the BOLD response in bilateral BG showed that the BG was significantly activated across all beat-based, metric rhythm conditions, while there was no significant difference of BG activation across all conditions (Figure 4.6.), supporting that perceiving temporal regularity involve the activation in the BG (Grahn & Brett, 2007; Grahn & Rowe, 2009; Geiser, Notter, & Gabrieli, 2012). Thus, it is assumed that the brain activation relating to the rhythmic complexity in the current fMRI may reflect higher order cognitive processing of hierarchical metrical organisation, rather than the violation of temporal regularity. However, the activation of BG in the current fMRI study may not be affected by volume accents, suggesting that the enhanced or disrupted binary pattern of metrical organisation during rhythm perception may activate higher-order cortical areas, such as IFG.
CHAPTER 4. Beat perception in complex rhythms

Basic contrast maps in the whole brain analysis

The subtraction analyses between Non-beat Accented and Unaccented rhythm conditions showed the significantly greater activation in the left STG. The activation in auditory areas during Non-beat Accented rhythm condition can be firstly considered with the mechanisms of mismatch negativity (MMN), which was traditionally found to be elicited by oddball sounds in auditory areas, and has been recently suggested to be involved in frontal areas in terms of higher order cognitive processing as well as in memory trace in terms of prediction of forthcoming information (Näätänen, Kujala, & Winkler, 2011). MMN is especially elicited when the predictive events were violated while being independent from attention, thus Non-beat Accented rhythm condition in this study is not related to MMN response in auditory processing. Non-beat Accents in this study may not violate the beat-based rhythmic structures much, and instead increase attention on Non-beat accents in addition to the underlying temporal regularity of the beat-based rhythms.

The subtraction analyses between Beat Accented and Unaccented rhythm conditions did not show any significant activation, suggesting that the enhanced binary pattern using additional volume accents in Beat Accented rhythm condition may not greatly differ from that in a control condition. The significantly greater activation in the left STG in Non-beat Accented compared with the Unaccented rhythm condition, on the other hand, suggests that the left STG has a role of higher cognitive processing to discriminate rhythmic patterns beyond the perception of loudness due to volume accents.

The left auditory areas have been suggested to be activated by rapid processing (Belin, Zilbovicius, & Crozier et al., 1998), in that the activation of the right auditory areas was reduced during rapid processing compared to that of slow processing. The
comparison between the rapid and slow temporal processing in their study, however, may not be the case to explain the increasing activation of the left STG in the current fMRI study because all rhythmic stimuli used in the current fMRI study were the same speed, beat-based rhythms.

Previous studies also showed contradicting results regarding the activation pattern of STG during the perception of temporal patterns. For example, the bilateral aSTG has been found to be activated during the perception of metric rhythms. When metric simple rhythms compared to metric complex or non-metric rhythms, a significantly greater activation in aSTG has been found (Grahn & Brett, 2007). In addition, a lesion study showed that aSTG is important to perceive metric patterns (Liégeois-Chauvel, Peretz, & Babaï et al., 1998). Meanwhile, perceiving temporal patterns in musical rhythm has been suggested to be widespread and bilateral in the brain (Peretz & Zatorre, 2005). In contrast to the significantly greater activation in STG during metric processing, complex temporal patterns have been found to increase the activation of STG. The non-periodic relative to periodic sequences in a volume detection task have been found to activate the STG (Geiser, Notter, & Gabrieli, 2012), and the activations of STG with STS have been shown when sound stimuli are more speech-like (Scott & Wise, 2003), suggesting that STG may have a role in perceiving complex sound sequences that do not need to be regular sound patterns. Although previous studies suggest that STG has a role in perceiving regular, metric temporal patterns, the relationships between the activation and the complexity of sequences are unclear. The increased left STG activation in the Non-beat Accented rhythm condition in the current study is due to additional volume accents on the same complex, but beat-based rhythms, therefore its activation may not be due to irregular rhythmic structures that have been shown in previous studies.
CHAPTER 4. Beat perception in complex rythms

Alternatively, the increased rhythmic complexity in the Non-beat Accented rhythm condition can be explained as the increased frequency to attend to during rhythm listening. According to dynamic attending theory (Jones & Boltz, 1989; Large & Jones, 1999), generated internal beats may increase attention, and the sense of a strong beat in metrical organisation may increase attention. If temporal regularity induces internal beats in the Non-beat Accented rhythm condition as well, which is assumed to be the case (as BG showed the significant activation), the Non-beat Accented rhythm condition induces perceptual accents in both temporal regularity of beat-based rhythm and non-beat accents. Compared to Unaccented and Beat Accented rhythm conditions, therefore, the Non-beat Accented rhythm condition has more frequent accentuation to attend to.

When the Non-beat Accented rhythm condition was compared to the Beat Accented rhythm condition, the left lateralised activation was shown, including IFG, STG, PMC, thalamus, SMA, and parietal lobe (Figure 4. 8.). There was no significant activation when the Beat Accented rhythm condition was compared to the Non-beat Accented rhythm condition. These left lateralised neural networks of PMC, thalamus and parietal lobe in the Non-beat Accented rhythm condition may be modulated by the mechanisms to increase the activation of the left STG and left IFG above, however, the current study delimits the discussion on the left IFG activation, in particular. The discussion of the current study focuses more on the IFG, and the research finding concerning the neural networks remains a subject for future study.

ROI analysis in the left IFG

The ROI analysis using the MPM-based ROI of IFG showed the activation of the left IFG (BA44) during Non-beat Accented compared to Beat Accented rhythm (Figure.
As discussed above, the activation of the left IFG (BA44) in Non-beat Accented rhythms can be interpreted in similar ways to the activation of the left STG in Non-beat Accented rhythm conditions. The activation in left IFG (BA44) during Unaccented rhythm condition suggests that the left IFG (BA44) is involved in metric processing of beat perception. The left IFG (BA45) and left STG have been found to be activated during metric simple rhythms that generate internal beats (Grahn & Brett, 2007), and the activation of left IFG during rhythm perception in the current study supports the previous research finding of the activation in left IFG during beat-based rhythm perception. However, the increasing activation of IFG during the Non-beat Accented rhythm condition is due to disrupting a binary pattern of metrical organisation, or disrupting the higher order cognitive processing in the hierarchical structures of metrical organisation. This suggests that left IFG activation can be increased by not only metric processing, but also disrupting the higher order cognitive processing in hierarchical structures in metrical rhythms, suggesting that two different mechanisms may activate IFG in rhythm perception. The increased activation in IFG can be interpreted as the increased frequency of timing points to attend to (see also the above discussion regarding the left STG activation).

The disruption of a binary pattern of metrical organisation of the current study is not, however, by the same mechanism of an oddball-like auditory perception or the mismatch negative paradigm that has been found to increase IFG activation (Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch, Gunter, & Cramon et al., 2002; Tillman, Janata, & Bharuch, 2003; see also Introduction). The complexity in the current study was the disruption of a binary pattern of metrical organisation, which can be further interpreted as either disrupting hierarchical metric processing or increasing time points to attend.
Repition suppression during Beat Accented rhythm condition

When the mean percent change in BOLD response in the left IFG and left STG was further analysed for the same and different rhythm pairs in each condition, the results showed that there was a significant difference of BOLD response between the same and the different rhythm pairs only when Beat Accented rhythms are heard (Figure. 4.11.). The significant difference in the mean percent change in BOLD response was not found in Unaccented and Non-beat Accented rhythm conditions. This results is rather striking because when the mean percent change in BOLD response in the same and different rhythm pairs were averaged, the significant difference of activation was between the Unaccented and Beat Accented rhythm conditions, and the Non-beat Accented rhythm condition. When the mean percent change in BOLD response was calculated in the same and different rhythm pairs, however, a significant different activation pattern was found between the Beat Accented rhythm condition and the Unaccented and Non-beat Accented rhythm conditions. Although the Non-beat Accented rhythm condition significantly activated the left IFG and left STG compared to other two conditions, the result the above suggests that the mechanism used to listen to the same and different rhythms can be strikingly affected by Beat Accents. This suggests that when metrical organisation, such as a binary pattern, is enhanced by additional accents in the beat-based, metric rhythms, the difference between the presentation of same and different rhythmic stimuli will be amplified in terms of neural activation.

The significantly reduced activation in the left IFG and left STG can be partly interpreted by the repetition suppression phenomenon in fMRI research. Several fMRI studies have found that when the stimuli or tasks are repeated, brain activation is
CHAPTER 4. Beat perception in complex rythms

decreased (Summerfield, Tritschuh, & Monti et al., 2008; Hamilton & Grafton, 2009; Andics, Gál, Vicsi, Rudas, & Vidnyánszky, 2013). Beyond the perceptual repetition, the repeated exposure in learning has been found to decrease cortical activation (Buchel, Coull, & Friston, 1999). The repetition suppression in the left IFG and left STG during the perception of the same rhythm pair in the current study is consistent with these previous studies. Although the same rhythm pairs were heard by listeners in Unaccented and Non-beat Accented rhythm condition in the current study, the repetition suppression did not happen, suggesting that the enhanced metrical organisation during time dependent information processing, such as the perception of beat-based rhythms, may result in repetition suppression.

The repetition suppression of the left IFG (BA44) and left STG during the perception of the same rhythm pairs of Beat Accented rhythms is due to the enhanced top-down processing; in the current study, a binary pattern of metrical organisation. Although the overall magnitude (or % BOLD response) of the same rhythm pair in the Beat Accented rhythm condition was significantly suppressed in the left IFG compared to the different rhythm pair, the significant difference between the same and different rhythm pairs showed lower statistical value in the left IFG (p < .07) compared to left STG (p < .03), suggesting the individual variables to perceive rhythmic structures in higher cortical areas may affect the statistical significance. This can be also interpreted as that higher cortex of left IFG (BA44) can be more affected by top-down processing regarding a binary pattern of metrical organisation.

The current fMRI study showed firstly, that beat-based, metric rhythm in all rhythm conditions activated auditory and sensorimotor areas including IFG, where in particular perceiving temporal regularity was successfully controlled across conditions. This was interpreted by BG activation, which showed significant, but similar activation
across all conditions. Secondly, the Non-beat Accented rhythm condition showed significantly greater activation in both left IFG and left STG, suggesting that the violation of a binary pattern of metrical organisation or the increase of attending time points increases BOLD response. Interestingly, a significantly suppressed activation in the same rhythm pair during the Beat Accented rhythm condition was found, suggesting that top-down processing due to enhanced metrical organisation (i.e., a binary pattern in the current study) may result in the repetition suppression.
Chapter 5. Metre perception in isochronous sequences

5.1. Introduction

This chapter aims to investigate the activation of IFG during the perception of metrical organisation in rhythms. Metrical organisation is a key factor in cognitive processing during rhythm perception, resulting in top-down processing between different hierarchical levels of metrical rhythmic structures. The cognitive processing of metrical organisation, in particular of a binary pattern, has been suggested to be developed early (Hanon & Johnson, 2005; Bergeson & Trehub, 2006), suggesting that this cognitive processing of the binary pattern of metrical organisation in western music culture may develops well in adults. However, neural mechanisms underlying IFG in non-musician adults during their perception of relatively simple metrical organisation, such as 2/4, 3/4, or 4/4, has been poorly understood.

It has been suggested that neural correlates of metre perception are significantly different between musicians and non-musicians. For example, musicians compared to non-musicians showed greater activation in the brain when the strong beat of the binary pattern was weakened in isochronous sequences (Brochard, Abecasis, & Potter et al., 2003). This suggests that the top-down processing in musicians may increase brain activity. Auditory activity in metre perception has been found to be left lateralised in musicians whilst right lateralised in non-musicians (Vuust, Pallesen, & Bailey et al., 2005). Here, the neural responses in auditory cortex were measured during the pre-attentive period before metrical organisation was disrupted, suggesting that the neural activity relates to the increased or entrained musical expectation of metre perception. Where their experimental design was in line with Brochard, Abecasis, and Potter et al.
CHAPTER 5. Metre perception in isochronous sequences

(2003), metre perception in musicians’ groups has shown the left lateralised brain activity (Abecasis, Brochard, & del Río et al., 2009) and, additionally, strong relative to weak beat showed a greater brain activity change.

Researches into the imagination of metre (or metrical organisation) also suggest neural correlates relating to a hierarchical cognitive processing that can be involved in metre perception. For example, when musicians imagined binary and ternary metrical organisation in isochronous sequences, strong beats in the imagined metric patterns evoked additional EEG responses in addition to the neural response corresponding to a steady beat (Nozaradan, Peretz, Missal, & Mouraux, 2011). This additionally evoked EEG response corresponding to the imagination of binary and ternary metric patterns suggests neuronal entrainment, which explains top-down cognitive processing without perceiving actual sensory information. Entrainment in rhythm sequences proposed by Large and Jones (1999) explains that attention to rhythm generates internal oscillation, which enables one to anticipate future events. According to this theory, neuronal oscillation correlates with attention, in which more attention may be induced by strong beats than weak beats, resulting in greater neuronal responses corresponding to the strong beats in metre perception. The entrainment and neuronal oscillations to metrical organisation patterns can be, therefore, dependent on the strong beats in metrical organisation, where the different types of metrical organisation may result in different neural correlates in terms of different time course and intensity of BOLD response in fMRI research.

In this sense, entrainment can also be understood in terms of a pre-attentive period, during which weak beats are perceived between strong beats in metre perception. The perceptual expectation to future events (or in this case strong beats) may increase if the number of weak beats increase. For example, single neuron recordings of rat brains
showed that neural activity in the primary auditory cortex was enhanced when perceptual expectation during entrainment (which can be understood as a pre-attentive period before perceiving sensory information or target stimuli) was enhanced (Jaramillo & Zador, 2011). In human brains, perceptual expectation activated the higher visual cortex even before the visual perception of stimuli activated the primary visual cortex (Chawla, Rees, & Friston, 1999). These two studies suggest that perceptual expectation in entrainment modulates neural activity in either primary sensory or even higher brain areas.

The correlation between the neural activity and the perceptual expectation may not be necessarily correlated. For example, if perceptual expectation is induced due to the repetition of the same stimuli, neural activity for the repetitive sensory information can be suppressed, where numerous fMRI studies showed the repetition suppression of BOLD response (Büchel, Coull, & Friston, 1999; Summerfield, Tritschuh, & Monti et al., 2008). In the research into repetition suppression, the perceptual expectation relating to the repetition of the sensory information (i.e. visual and auditory stimuli) reduces brain activity. Suppression of BOLD response due to the repetition of the stimuli has been found in both primary and higher cortical areas. Even the auditory hair cell at a low level sensory processing showed suppressed response due to the repetition of the same information (Todd, 1994). In this sense, a predictive coding model and free energy principle (Friston, 2008; Friston & Kiebel, 2009; Friston, 2010), proposing that the brain optimises to predict future events and minimises prediction error (perceptual expectation due to the repetition in this context), may well explain this repetition suppression relating to perceptual expectation.

Although metre can be often perceived explicitly in everyday music listening, in experimental circumstance it can be difficult to control the highly subjective and psychological process in metre perception. For example, perceiving isochronous
sequences resulted in subjective metrical grouping (Brochard, Abecasis, & Potter et al., 2003). Metre perception during music listening may, therefore, involve higher order cognitive processing, relating to the hierarchical organisation of rhythmic structures by metrical organisation.

The left IFG has been suggested to be involved in hierarchical cognitive processing (Koechlin & Jubault, 2006; Tettamanti & Weniger, 2006) whilst the right IFG and BG have been suggested to be involved in attention and time perception (Rao, Mayer, & Harrington et al., 2001; Hedden & Gabrieli, 2006, Weissman, Roberts, Wisscher, & Woldorff, 2006). This suggests that bilateral IFG may be involved in metre perception in isochronous sequences.

Thus, the fMRI study in this chapter aimed to investigate the activation pattern of bilateral IFG during the cognitive processing of the different types of metrical organisation in isochronous sequences, of which metrical grouping levels were parametrically organised by every 2 beats (C2), 3 beats (C3), or 4 beats (C4), and the sequence without a grouping level (C1), using pitch accents. Although there are two contradicting mechanisms in the relationship between the magnitude change of brain activation and the perceptual expectation as discussed in the above, metre perception during a steady beat will show the increased pattern of brain activation corresponding to the increased perceptual expectation. For example, the conditions of C2, C3, and C4 have the linearly increased duration of entrainment, and at the same time, have the decreased frequency of strong beat in metrical grouping in a parametric manner. The increased duration of entrainment from C2 to C3 to C4 may increase brain activity due to the increased number of weak beats between the strong beats of metrical organisation patterns, whereas the increased frequency of the repetition of metrical organisation patterns from C4 to C3 to C2 due to the same number of isochronous sequences will
decrease the brain activity. It is hypothesised, therefore, firstly, that beat perception in isochronous sequences will activate auditory sensorimotor areas and bilateral IFG, and, secondly, that parametrically increased metrical grouping levels from C2 to C3 to C4 will increase the IFG activation.

5.2. Methods

Participants

Fifteen participants with no history of neurological, psychiatric disorders, or hearing difficulties were recruited (mean age 29.46 years; s.d. 9.65; range 19 – 49 years; 6 males; 0 to 5 (mean 0.93, s.d. 1.53 years of musical training; for the details of participant information, see also Table 5.1. This participant group is the same participants recruited in the fMRI study of Chapter 4.). All participants were right-handed (N=15, mean L.Q. = + 84.71, s.d. = 16.67, Decile R.6, The Edinburgh Handedness Inventory, Oldfield, 1971). The study was approved by the Local Research Ethics Committee (West of Scotland Research Ethics Service (WoSRES), REC reference: 12/WS/0229) and The University of Edinburgh Ethics Committee (Edinburgh College or Art). The participants gave fully informed written consent of their willingness to participate, and were compensated minimum travel expenses with no fixed compensation for their participation.
Table 5.1. Participant demographics

<table>
<thead>
<tr>
<th>No. of participants</th>
<th>Gender (F: female, M: male)</th>
<th>Age</th>
<th>Music education</th>
<th>Duration (&lt; 5 years)</th>
<th>Starting age (years) / instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>22</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>49</td>
<td></td>
<td>2</td>
<td>8 / piano</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>19</td>
<td></td>
<td>2</td>
<td>16 / guitar</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>33</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>19</td>
<td></td>
<td>2</td>
<td>15 / singing</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>19</td>
<td></td>
<td>5</td>
<td>11 / piano, singing</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>35</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>45</td>
<td></td>
<td>3</td>
<td>11 / piano, violin</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>26</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>35</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>27</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>29</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>38</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>28</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>18</td>
<td></td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>Total</td>
<td>F (9), M (6)</td>
<td></td>
<td>Avg 29.46 (s.d. 9.65)</td>
<td>Avg 0.93 (s.d. 1.53)</td>
<td></td>
</tr>
</tbody>
</table>

Study Design

Participants listened to a series of isochronous sequences during fMRI scan. Rhythmic stimuli comprised of the same number of beat (i.e., 24 beats) in isochronous sequences, which were metrically grouped by 2, 3, or 4 beats. Pitch accents were used to create different types of metrical grouping and isochronous sequences, resulting in the metrical grouping of 2/4 (C2), 3/4 (C3), and 4/4 (C4), and the isochronous sequences without metrical grouping (1/4 or C1). The participants indicated whether the rhythm sequences were C1, C2, C3, or C4 by pressing buttons after listening to the rhythm sequences. C1 did not have a metrical grouping level, and the participants needed to press a button 1 after listening to C1, which is to validate task performance between C1 and other metrical grouping conditions. A very simple task and button response was designed to avoid the limitations of a passive listening task, and was intended not to bring cognitive
difficulty. The accuracy of the task performance was considered to check whether the participants understood and performed the task properly during the MR scan. The response time was not considered because the participants had not been instructed to perform the task as quickly as they could. All conditions of C1, C2, C3, and C4 were repeated eight times during the fMRI scan.

**Stimuli and task**

Rhythmic stimuli comprised of a series of isochronous sequences using woodblock sound (Finale 2007, Makemusic, Inc. Eden Prairie, MN, USA). Each condition of C1, C2, C3, and C4 has 24 beats, and the interval between beats (i.e., inter stimulus interval (ISI)) was 600ms, an interval that has been used in previous empirical studies (Parncutt, 1994; Drake, Jones & Baruch, 2000). Each isochronous sequence of one block lasted for 14400ms. The woodblock sound was chosen to create beat-based rhythmic stimuli rather than duration-based rhythmic stimuli. Rhythmic complexity in accordance with different metrical grouping levels of C2, C3, and C4 were created by pitch accents. High and low pitch accents were counterbalanced for each condition, and strong beats of metrical organisation were accented by either low or high pitch (Figure 5.1.). The averaged intensity of sound stimuli (RMS) of all rhythmic stimuli was normalised to the same level of -16dB (in each condition, high pitch sound was RMS -16dB, peak -0.1dB; low pitch sound was RMS -16dB, peak -0.5dB). Audio sample rate was 44,100Hz, and audio bit depth was 16 bits with mono audio channel (SoundForge pro 10 software, Sony Creative Software, Middleton, WI, USA). The participants indicated whether the rhythm sequences were C1, C2, C3, or C4 by pressing buttons after listening to the rhythm.
The task was intended to sustain listener’s attention during metre perception and, at the same time, to induce minimum cognitive load.

<table>
<thead>
<tr>
<th>Metrical organisation</th>
<th>Listening to beat sequence</th>
<th>Button response</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>• • • • • • • • • • • • • • •</td>
<td>1</td>
</tr>
<tr>
<td>C2</td>
<td>• • • • • • • • • • • • • • •</td>
<td>2</td>
</tr>
<tr>
<td>C3</td>
<td>• • • • • • • • • • • • • • •</td>
<td>3</td>
</tr>
<tr>
<td>C4</td>
<td>• • • • • • • • • • • • • • •</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 5.1. Schematic of rhythmic stimuli.

Training Procedure before fMRI scan

Participants practiced a training session immediately prior to MR scan. This training session aimed to enable the participants to be familiar with rhythmic stimuli, the task, and the environment of fMRI experiment. A laptop, headphones, and a picture of MR compatible four-button box were given to participants in an MR preparation room. An experimenter confirmed that the participants fully understood both rhythmic stimuli and task, and then the participants practiced three trials with a laptop by themselves. After finishing the training session, the participants were guided to a mock MR scanner room, where the experimenter introduced the environment of MR scan, and the details of fMRI experimental procedures (see Appendix 3. Scripts for the protocol of two fMRI studies).
Radiographers guided participants to the MR room, and assisted them to get ready to perform the experiment in a safe and comfortable posture. The participants were given earplugs to wear inside of MR compatible noise cancelling headphones. The participants were then given a 1-minute MR practice session, which was to familiarise them with the scanner noise and to check the comfort of the participants with fMRI apparatus, such as headphones, goggles, and button response. The experimenter confirmed that all participants could hear the sound stimuli over the noise of fMRI scan. During fMRI scan for 16 minutes 21 seconds, rhythmic stimuli were presented using a block fMRI design (Figure 5.2). Each condition lasted for 15 seconds, followed by fixation for 15 seconds. The first and last 15s blocks were fixation, and the presenting order of rhythmic stimuli of all conditions was randomised.

**Figure 5.2.** Schematic of fMRI scan
Data acquisition and MR apparatuses

MR data were obtained at the Clinical Research Imaging Centre (CRIC), The University of Edinburgh, using a 3T Verio whole-body MRI system (Siemens Medical Systems, Erlangen, Germany). A quadrature birdcage RF coil was used for RF transmission and reception. First, high-resolution 3D anatomical images covering the whole brain were obtained using an MPRAGE sequence (TR 2300 ms; TE 2.98 ms; voxel side 1 mm; FOV 240 by 256 mm; flip angle 9°) encompassing the whole cerebrum parallel to the anterior commissure (AC) – posterior commissure (PC) plane. Functional images were obtained using a blood oxygenation level dependent (BOLD) contrast gradient echo echo-planar T2*-weighted images (EPI) sequence prescribed at the same orientation (TR 3,000 ms; TE 30 ms; flip angle 90°; pixel size 3 mm, FOV 192*192 mm; slice thickness 3.75mm; 36 slices; total volume 327). Conventional Siemens noise-cancelling MR compatible headphones with earplugs and the goggles of Nordic MR visual system (www.nordicneurolab.com) were used to convey sound and visual stimuli to participants in the MR scanner. Presentation programming software (www.neurobs.com) was used to synchronise visual and sound stimuli on a desktop computer. The Nordic system was used to synchronise MR sequence and Presentation software of the desktop. An MR compatible four-response button box was used (www.curdes.com).

Data analysis

Behavioural data analysis

The accuracy was calculated from the button response, which was collected from a judgement task during MR scan. The task was to enable participants to attend rhythmic
stimuli without a cognitive load, which was quite similar to passive listening. Thus, the aim of the behavioural data analysis was to check whether the participant listened the rhythmic stimuli properly, and further statistical analysis was not performed. Reaction time was not considered, as we did not ask participants to perform a task as quickly as possible.

Functional Magnetic Resonance Imaging (fMRI) data processing

Preprocessing of structural and functional images

BrainVoyager QX 2.8.0 (Brain Innovation, Maastricht, The Netherlands) was used for fMRI data analysis. To preprocess structural images, firstly, the skull parts of individual anatomical images (i.e., T1 MPRAGE) were peeled, and the structural images were corrected by intensity inhomogeneity. The images were aligned and transformed with a line of AC (anterior commissure) and PC (posterior commissure), using a cubic spline spatial interpolation. The aligned AC-PC anatomical images were then normalised to Talairach coordinates (Talairach & Tournaux, 1988), using sinc interpolation using the anatomical template installed with 12 sub-cuboids landmarks in BrainVoyager QX. To preprocess functional images, the first two volumes of functional images were discarded to exclude T1 saturation effects. Motion detection was performed by trilinear interpolation, and motion correction by sinc interpolation, using by 6 rigid motion parameters (3 translation and 3 rotation of x, y, z parameters). Less than 4 mm motion in all x, y, and z directions of all participants was checked. Slice time correction was performed and the data were re-sampled by an ascending interleaved order and cubic spline interpolation. High pass temporal filtering with a Fourier basis set of 2 sines/cosines, including a linear trend, was performed. To co-register structural and
CHAPTER 5. Metre perception in isochronous sequences

functional images, functional images were firstly registered to their own native anatomical images using a header-based initial alignment (IA) for its proximity of global orientation, and a fine-tuning alignment (FA) by Gradient-driven affine transformations (6 - 12 parameters). Individual functional images were then normalised to Talairach coordinates and re-sampled using trilinear interpolation. Finally the normalised functional images were spatially smoothed with an isotropic Gaussian filter with FWHM 6 mm.

Statistical analysis of fMRI data

In the first level of individual analysis, General Linear Models (GLMs) of four conditions of C1, C2, C3, and C4 were modelled with a conventional canonical hemodynamic response function (i.e., time to response peak is 5s and time to undershoot peak is 15s). Motion parameters estimated during the realignment (x, y, z parameters in 3 rotations and 3 translations) were modelled in single subject GLMs. In the second level of group data analysis, a GLM for group data was estimated using a random effect; the parametrically modelled four conditions of C1, C2, C3, and C4 in single subject GLM and the motion parameters of individuals were included as residuals. Autocorrelation was estimated using an AR (2) model and GLM was refitted to correct serial correlations. The fixation period was modelled as the constant 0, and the conditions from C1 to C2 to C3 to C4 was modelled as the constant 1 to 2 to 3 to 4. Although it has been shown that listeners tend to subjectively group isochronous sequences (Brochard, Abecasis, & Potter et al., 2003), C1 was modelled as the constant of 1 because the participants in the current study had been instructed to indicate C1 as a button press of 1. This was assumed that the participants may not intend to group C1 as they group other conditions of C2, C3, and C4.
CHAPTER 5. Metre perception in isochronous sequences

To identify brain regions that were parametrically increased by the increasing levels of metrical grouping, the GLM in the second level of group data analysis using a parametric modulation from C1 to C2 to C3 to C4 was calculated in the whole brain analysis. Region of Interest (ROI) analysis was performed using the ROI defined by the MPM (Eickhoff, Heim, Zilles, & Amunts et al., 2006) of bilateral IFG. Within the ROIs of IFG, a one-way analysis of variance (ANOVA) within-subjects using SPSS 22 (IBM SPSS statistics) was conducted, to examine whether the mean percent in the BOLD response across the conditions from C2 to C3 to C4 was statistically significant. Finally, within bilateral IFG (BA44/45) using MPM, a GLM using the parametric modulation from C2 to C3 to C4 was calculated to specify the IFG region that responds to the increasing levels of metrical grouping.

All analyses incorporated a correction for multiple comparisons based on the false discovery rate (FDR) at the voxel level as described by Genovese et al. (2002). Percent BOLD signal change was calculated by averaging the time series of activated voxels for spheres of 10 voxels radius centered on the peak voxel of the relevant clusters of activation. Brodmann’s area (BA) was defined in Talairach and Tournoux (1988) for all ROIs, and IFG was additionally defined by cytoarchitectonic Brodmann’s area using an MPM.

5.3. Results

Behavioural data collected in a judgement task

The mean percent of the accuracy of task performance obtained during fMRI scan was over 90% for all conditions as predicted (see also Appendix Figure 5.1.).
fMRI experimental results

Brain regions that parametrically increase from C1 to C2 to C3 to C4 in the whole brain analysis

Brain activations that were parametrically increased during metre perception in isochronous sequences in the whole brain analysis showed the activation of the primary and secondary auditory areas, PMC, anterior insula, SMA, IFG / MFG, BG, cuneus in both hemispheres and left lingual gyrus (Figure 5.3., $p$ (FDR) < .05; Table 5.2.).

Figure 5.3. Brain activations that were parametrically modulated from C1 to C2 to C3 to C4
CHAPTER 5. Metre perception in isochronous sequences

The mean percent change in BOLD responses in the activation regions showed the parametrically increasing patterns from C2 to C3 to C4 (Figure 5.4.). The mean percent change in BOLD response of brain activation in C1 showed a greater activation than C2, in particular in the bilateral IFG, right insula, right PMC and left lingual gyrus.

Figure 5.4. The mean percent change in BOLD response in the activation areas that were parametrically modulated from C1 to C2 to C3 to C4.
ROI analysis using the MPM-based ROI of IFG

ROIs of bilateral IFG (BA44/45), which were cytoarchitectonically defined using an MPM, were used to conduct ROI analysis. Within the ROIs using the MPM-based IFG, GLM using a parametric modulation from C1 to C2 to C3 to C4 were calculated. The results showed the activation of bilateral IFG (Figure 5.5. p (FDR) < .05). Brodmann’s area based on Talairach coordinates was indicated in the Figure, although the bilateral activation regions were cytoarchitectonically defined using the MPM of BA44 and BA45, i.e., the bilateral IFG activation was cytoarchitectonically defined by BA44 and the right IFG was more the dorsal part. ANOVA of the mean percent change in BOLD response in bilateral IFG activation showed that the BOLD response significantly increased from C2 to C3 to C4, i.e., the left IFG (F_{2, 42} = 4.0, P < .02) and right IFG (F_{2, 42} = 6.03, P < .005) (Figure. 5.6.).

![Image of brain scan showing bilateral IFG activation](image)

**Figure 5.5.** The bilateral IFG (BA44/45) activation in ROI analysis.
CHAPTER 5. Metre perception in isochronous sequences

Figure 5.6. The mean percent change in BOLD response in bilateral IFG activation (± standard error of mean)

Table 5.2. Stereotaxic coordinates of brain regions defined by both whole brain and ROI analyses

<table>
<thead>
<tr>
<th>Brain Regions (Brodmann’s areas)</th>
<th>Talairach coordinates (peak voxel of clusters)</th>
<th>t-value (avg. cluster)</th>
<th>p-value (avg. cluster)</th>
<th>Cluster size † (No. of voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Superior temporal gyrus (BA13)</td>
<td>48 -18 6</td>
<td>22.537</td>
<td>&lt; .001</td>
<td>351</td>
</tr>
<tr>
<td>R Middle frontal gyrus (BA9)</td>
<td>49 14 31</td>
<td>8.540</td>
<td>&lt; .001</td>
<td>49</td>
</tr>
<tr>
<td>R Occipital lobe, Cuneus (BA30)</td>
<td>1 -66 7</td>
<td>10.793</td>
<td>&lt; .001</td>
<td>42</td>
</tr>
<tr>
<td>L Medial frontal gyrus (BA6)</td>
<td>-8 -1 55</td>
<td>8.136</td>
<td>&lt; .001</td>
<td>58</td>
</tr>
<tr>
<td>L Occipital lobe, Cuneus (BA18)</td>
<td>-10 -88 18</td>
<td>6.876</td>
<td>&lt; .001</td>
<td>12</td>
</tr>
<tr>
<td>L Occipital lobe, Lingual gyrus (BA19)</td>
<td>-14 -61 -5</td>
<td>7.384</td>
<td>&lt; .001</td>
<td>42</td>
</tr>
<tr>
<td>L insula (BA13)</td>
<td>-35 18 6</td>
<td>8.498</td>
<td>&lt; .001</td>
<td>154</td>
</tr>
<tr>
<td>L Precentral gyrus (BA6)</td>
<td>-32 -7 52</td>
<td>7.931</td>
<td>&lt; .001</td>
<td>15</td>
</tr>
<tr>
<td>L Transverse temporal gyrus (BA41)</td>
<td>-48 -26 11</td>
<td>14.062</td>
<td>&lt; .001</td>
<td>474</td>
</tr>
<tr>
<td>L Inferior frontal gyrus (BA44)</td>
<td>-41 11 13</td>
<td>6.151</td>
<td>&lt; .001</td>
<td>12</td>
</tr>
</tbody>
</table>

Parametrically modulated brain regions from C1 to C2 to C3 to C4 in the ROI of MPM-based IFG

<table>
<thead>
<tr>
<th>Brain Regions (Brodmann’s areas)</th>
<th>Talairach coordinates (peak voxel of clusters)</th>
<th>t-value (avg. cluster)</th>
<th>p-value (avg. cluster)</th>
<th>Cluster size † (No. of voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Inferior frontal gyrus</td>
<td>49 12 29</td>
<td>8.540</td>
<td>&lt; .001</td>
<td>49</td>
</tr>
<tr>
<td>L Inferior frontal gyrus</td>
<td>-41 11 13</td>
<td>6.151</td>
<td>&lt; .001</td>
<td>12</td>
</tr>
</tbody>
</table>

† The original number of voxels by T1 resolution (1×1×1 mm$^3$) was manually divided by 27 to match with the resolution of functional images (3×3×3 mm$^3$).
5.4. Discussion

The fMRI study in this Chapter aimed to investigate the activation of bilateral IFG in metre perception in isochronous sequences, where the metrical grouping was organised by 2, 3, or 4 beats. After listening to the isochronous sequences, participants indicated the perceived metrical grouping level. Behavioural data acquired during fMRI scan showed that the participants had no difficulty to perceive metrical groupings during the fMRI scan.

The results showed that beat perception in isochronous sequences activated the auditory and sensorimotor areas, including the primary and secondary auditory areas, SMA, anterior insula, BG, PMC and the IFG. The left lingual gyrus in visual system was also parametrically increased in accordance with the increase of metrical grouping levels from C2 to C3 to C4. The results in the whole brain analysis suggest, firstly, that the metrical organisation from C2 to C3 to C4 increased the mean percentage signal change in BOLD response of the brain activity. This increased pattern is not limited to some specific brain regions and can be interpreted as a whole brain activity pattern. Predictive coding model and free energy principle (Friston, 2008, 2010), which supports, in particular, the neural mechanisms of the whole brain level, can be understood with the results. The increased BOLD response can be interpreted as suggesting that the increased interval between strong beats (e.g. the duration between strong beats is longer from C2 to C3 to C4) affects the duration of attending to strong beats or the cognitive processing during higher-order metrical grouping levels.

C2 showed significantly less brain activation compared with C3 and C4, and this could be interpreted as several reasons. Firstly, a binary pattern, which has been suggested to be a default metrical grouping (Fraisse, 1984; Drake, 1993; Brochard, Abecasis, & Potter et al., 2003; Bergeson & Trehub, 2006), may result the reduced brain
CHAPTER 5. Metre perception in isochronous sequences

activity in C2. However, Brochard, Abecasis, and Potter et al., (2003) suggested that
cognitive process, rather than automatic default subjective grouping affected even a
default binary metre perception. For example, in their experiment, they have shown that
musicians showed the lager signal change than non-musicians when strong beat was
violated in simple binary pattern in isochronous sequences. The violation in metre
perception was designed by decreasing the volume of the beat without disrupting
temporal regularity. The larger signal difference evoked when strong beats are violated
has been interpreted as being because strong beat receives more attention than weak beat,
and significantly late positive parietal ERPs response is evoked when strong beat is
violated compared to when weak beat is violated, which is in line with the dynamic
attending theory (Jones & Boltz, 1989) that suggests attention increases on strong beats.
Secondly, as mentioned in the introduction, this reduced brain activation in C2 can be
related to the predictive coding model and free energy principle (Friston, 2008, 2010).
For example, the most frequent repetition of C2 compared with C3 and C4 may increase
perceptual expectation, which consequently minimises the prediction errors, which
reduce the free energy of brain activity.

The EEG study showed that the imagination of metre elicited additional neural
oscillation in addition to neural activity corresponding to beat, suggesting attention and
entrainment (Nozaradan, Peretz, Missal, & Mouraux, 2011). As having shown in the
previous EEG study, strong beats in metre produce additional brain activation. For
example, significantly greater activation can be shown in C2 than C3 or C4, as strong
beats in C2 are the most frequently repeated. The results of the current experiment,
however, showed that C2 suppressed the brain activation, suggesting that the frequent
repetition of the binary pattern can be a factor in suppressing BOLD responses in the
brain.
CHAPTER 5. Metre perception in isochronous sequences

The significantly greater activation in C1 compared to C2 suggests that the possibility of higher subjective grouping may contribute to an increase of activation. The greater mean percent change in the BOLD response during C1 condition suggests a subjective metric grouping, although the participants had received an introduction that they would press a button of 1 after they listened to C1. This suggests that the participants may involuntarily grouped C1 as different metric grouping level (after MR scan, while interviewing the participants about post fMRI scan experience, some participants reported that they grouped C1 as the groups of four or eight beats). Subjective grouping has been shown in isochronous sequences (Parncutt, 1994; Temperley, 1963).

Meanwhile, pitch accents in C2, C3, and C4 can result in different intervals or durations between the strong beats of metrical grouping levels, where the interval of C2 has a weak beat, C3 has 2 weak beats and C4 has 3 weak beats (C1 is not discussed in this context). This means that the duration between the strong beats is shortest in C2 and longest in C4, which consequently suggests that the increased duration can be a factor that correlates with the increased percent signal change in BOLD response of activation. In terms of the different durations between different types of metrical organisation, however, it is important to consider the role of weak beats during entrainment or a pre-attentive period pivot on strong beats of metric organisation. For example, when the same duration without weak beats or internally generated beats is given, it does not offer the regular temporal time points to attend to, resulting in, presumably, the different brain activation pattern from C2 to C3 to C4.

The question is whether the increased frequency to attend to weak beats or the increased duration between strong beats regardless of weak beats, is a factor in increasing percent signal change in BOLD response of activation of IFG. Time
perception with and without beats will show the difference in terms of neural activation. This can be partly answered by the result in the experiment of Chapter 5; C1 condition showed greater activation than C2 in IFG activation, suggesting the subjective interpretation affects the brain activity much more. The increased activation from C2 to C3 to C4 does not simply correspond to the increased duration, therefore, and it can instead correspond to the frequency of weak beats between the strong beats.

The interpretation that increased brain activation correlates with an increase in the number of weak beats can contradict with the top-down processing of metre perception, however. For example, during metre perception in isochronous sequences, strong beats suggest a hierarchical structure in the isochronous sequences, which may sustain attention between strong beats rather than counting each weak beat between the strong ones. Thus, if brain activation is increased due to the number or the duration of weak beats, it suggests bottom-up processing rather than top-down processing during metrical organisation in isochronous sequences. This interpretation may need to be resolved in a future study by comparing the groups of participants, i.e., musicians compared with non-musicians.

The increased BG activation in the results suggests that BG is not only related to temporal regularity, but also correlates with the higher cognitive processing of metrical organisation. As temporal regularity was controlled across all conditions, the increase of BG activation is related to the increased attention or increasing levels of metrical grouping. It has been suggested that BG has a function of holding attention during time perception (Ravizza & Ivry, 2001; Nieoullon, 2002), therefore, it is reasonable to interpret that BG activation can be related to sustain the attention between strong beats in different levels of metrical organisation. The research finding of cognitive function of BG in Chapter 5 may not be the same function as that revealed in
To sum up, the current experiment demonstrated that an increase of metrical grouping levels in isochronous sequences increases the activation of bilateral IFG. The significantly lower activation in C2 can be interpreted as two ways. Firstly, the frequency of the repetition increases the perceptual expectation, resulting in the reduced brain activation. Secondly, a less weak beats to attend to between strong beats in the metrical organisation may reduce brain activation. The significantly greater activation in C1 compared to C2 suggests that subjective grouping in C1 can be higher than the binary pattern of metrical organisation.
Chapter 6. General Discussion and Conclusions

Chapter 6 summarises the research findings of IFG activation in the two fMRI studies of Chapter 4 and 5, and discusses the role of IFG in beat and metre perception.

6.1. Evidence of IFG activation in beat and metre perception

The two studies of this thesis investigated the activation of IFG during beat and metre perception (Chapter 4 and 5). In Chapter 4, Beat Accented rhythms enhanced the binary pattern of metrical organisation, whereas Non-beat Accented rhythms disrupted this binary pattern, which was the temporal regularity of beat. The result showed that the Non-beat Accented rhythm condition activated the left IFG (BA44) significantly greater than the Beat Accented rhythm condition. The right IFG was significantly activated during all rhythm conditions (Unaccented, Beat Accented, and Non-beat Accented rhythms), resulting in no significant difference across conditions. In Chapter 5, metre perception in isochronous sequences was measured, of which isochronous sequences were metrically organised by 2 (C2), 3 (C3), or 4 beats (C4), or without grouping (C1). The activation in bilateral IFG was parametrically increased in accordance with the increased metrical grouping levels from C2 to C3 to C4, suggesting that bilateral IFG is involved in higher order cognitive processing of metrical organisation.

The discussion in this section is summarised with three issues: (1) increased BOLD response in Non-beat Accented rhythms, (2) decreased BOLD response in Beat Accented rhythms, and (3) parametrically increased BOLD response from C2 to C3 to C4.
Increased BOLD response in Non-beat Accented rhythm condition

In Chapter 4, Non-beat Accented rhythm condition relative to Beat Accented rhythm condition showed the increased activation in the left IFG (BA44), whereas the right IFG activation showed no significant difference across conditions.

Additional volume accents of off-beat are assumed to weaken the sense of temporal regularity of beat, which is the binary pattern of metrical organisation in complex rhythms. Disrupting this temporal regularity in Non-beat Accented rhythms is, therefore, disrupting the binary pattern, while the temporal regularity of beat-based rhythmic structures was the same. This kind of disruption, detailed in Chapter 4, is different from a mismatched or oddball-like sound paradigm, which violates the temporal expectation because the rhythmic structures are the same as beat-based rhythms. Although whether the sense of pulse was weakened in Non-beat Accented rhythm conditions or not was not directly measured in the study, Non-beat Accented rhythm conditions were based on the same beat-based metrical rhythmic structures, and its induced temporal irregularity was not disrupting the beat-based rhythmic structures.

Firstly, the increased activation of left IFG (BA44) in Non-beat Accented rhythm condition could be due to the increased frequency of perceptual accents to attend to. The difference between Beat Accented and Non-beat Accented rhythm conditions is in the position of additional volume accents. In all rhythmic stimuli of beat-based rhythms, phenomenal volume accents of Beat Accent rhythm conditions overlap with the metric accents due to the temporal regularity of beat in rhythms, whereas Non-beat Accent conditions added phenomenal volume accents during off-beat in addition to metric accents, which are presumably induced during the perception of beat-based rhythms due to the temporal regularity of beat in rhythms. According to Jones and Boltz (1989), attending modes in rhythm processing can be divided into two: ‘future-oriented
attending’ and ‘analytic attending’. When temporal regularity is enhanced by Beat Accents, its attention can be the future-oriented attending mode, and when temporal regularity is disrupted by Non-beat Accents, it can result in analytic attending (temporal regularity in Chapter 4 indicates a binary pattern of metrical organisation). The fact that there was no significant difference in IFG activation between Unaccented and Beat Accented rhythm conditions supports this idea that the attending rhythms (or attending mode) can be different between two of the Unaccented and Beat Accented rhythm conditions and the Non-beat Accented rhythm condition.

Regarding the number of rhythmic elements, the complexity of the two conditions of Beat Accented and Non-beat Accented rhythm conditions are the same. Thus, the neural difference between two conditions were not due to rhythmic structure itself. This supports the interpretation that the two different attending mode used during rhythm perception caused the rhythmic complexity in this study. Non-beat Accented rhythms have more frequent accents to attend to (e.g., attending to both metric accents due to the temporal regularity of each beat and phenomenal accents due to Non-Beat volume Accents), whereas Beat Accented rhythms have less frequent accents to attend to (e.g. attending to metric accents due to the temporal regularity of the beat). Unaccented rhythms, which have ambiguous temporal regularity as a control condition, presumably, have two or four times accents to attend to, depending on the pulse or beat perception that the participants may subjectively hear by. The rhythmic complexity in Chapter 4 is, therefore, related to attending mode, which was modulated by the position of volume accents. Those more frequent accents to attend to consequently can be a factor to increase IFG activation.

Secondly, co-activation of left IFG (BA44) and left STG, which was shown in Non-beat Accented rhythm condition of the current thesis, has been found to be involved
in speech processing in previous studies. For example, Obleser, Meyer, and Friederici (2011) have shown that complexity in auditory comprehension processing varies the percentage of BOLD response in left IFG as well as left STG. The left STG in Wernicke’s’ area and left IFG (BA44) in Broca’s area have long been discussed as language-specific neural networks. Although the current fMRI study did not use linguistic stimuli, the activation of the left IFG (BA44) and STG can be understood as auditory processing in general, supporting that non-language sound stimuli also activate this circuit.

Phonological processing in speech perception also activates STG and IFG of the left hemisphere (Buchsbaum, Hickok, & Humphries, 2001; Zatorre, Evans, Meyer, & Gjedde, 1992; Burton, 2001; Zatorre, Evans, & Meyer, 1994; Hickok & Poeppel, 2007), suggesting the perception of sequence activates STG and IFG in the left hemisphere. The beat-based, metric rhythms in the current thesis are rapid and complex, so can be similar to speech sounds. According to this previous research finding, the activation of left IFG and STG in the current thesis might be due to rapid, complex rhythms. In addition, the research finding of the current thesis supports that these two regions of left IFG and STG are not only simply activated by rapid, complex rhythm sequences but also distinguish the temporal patterns in terms of different positions of accents and, therefore, these two regions are involved in the perception of accents in rapid, complex rhythm sequences.

Different attending modes and different number of frequency to attend between Non-beat Accented and Beat accented rhythm conditions can be further extended to understand higher order cognitive processing. Disrupted perception of a binary pattern in Non-beat Accented rhythm condition may not induce higher order cognitive processing in hierarchical rhythmic structures between the subdivisions of beat level and beat level, however, enhancing perception of a binary pattern in Beat Accented rhythm condition
CHAPTER 6. General Discussion and Conclusions

will result in the higher order cognitive processing, which is a metrical grouping due to the binary pattern. In this sense, the increased activation in the left IFG in Non-beat Accented rhythm condition is understood as the reduced activation in the left IFG in Beat Accented rhythm condition, which can be in line with top-down processing (this is discussed further in the next section).

*Decreased BOLD response in Beat Accented rhythm condition*

Unaccented rhythm condition, which here is a control condition, can be considered as equivalent to metric simple rhythm in previous studies. Although Unaccented rhythm conditions comprise of non-isochronous combinations of duplet and triplet subdivisions of beat, the sound information of on-beat can be still heard, which may assist listeners to generate internal beats. Thus, neural correlates during the perception of Unaccented rhythm conditions may be similar to the neural correlates during the perception of temporal regularity in rhythms, the latter of which have been often studied through the comparison between temporal regularity and irregularity in rhythms. In this sense, greater activation patterns of Non-beat Accented rhythm condition suggest that the percent signal change in BOLD response of the neural activation involved in Unaccented rhythm condition is increased by the increase of complexity in a Non-beat Accented rhythm condition. The result showed, however, that the reverse is in fact the case; there was no significant difference when Unaccented and Beat Accented rhythm conditions were directly compared with.

Although there was no significant difference between Unaccented and Beat Accented rhythm conditions overall, only Beat Accented rhythm condition showed a significant difference in BOLD response between the same and different rhythm pairs.
This probably suggests that Beat Accents assist the generating internal beats and the enhanced beat perception of the binary pattern suppressed activation in the repetition of the same rhythmic stimuli. The repetition suppression in the same rhythm pair in Beat Accented rhythm condition can be understood as cognitive efficiency in relation to metrical grouping.

Generating internal beats has been suggested to activate the left IFG (BA45) (Grahn & Brett, 2007) and the violation of expected harmonic sequences has been also suggested to activate left IFG (Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch, Gunter, & Cramon et al., 2002; Tillman, Janata, & Bharuch, 2003). These two contradicting mechanisms to activate IFG are important. The increased activity in the left IFG has been found when there are salient metric patterns in rhythms (Grahn & Brett, 2007). The increase of right IFG activation in musicians compared to non-musicians in a reproduction task (Chen, Penhune, & Zatorre, 2008a) has been also suggested as related to the musical proficiency of musicians. The mechanisms that increase the activity of the right IFG in musicians are related to better performance or better proficiency. If IFG activation in the former case is related to cognitive proficiency, the latter can be related to the proficiency of expertise, both of which activate the IFG.

The decreased IFG activation in Beat Accented rhythm condition is, therefore, not consistent with the mechanism to activate IFG in the previous two fMRI studies. Rather, the result is consistent with the other previous studies into cognitive complexity. For example, the complexity in syntactic processing of language activated the left IFG (Caplan, Alpert, Waters, & Olivieri., 2000; Embick, Marantz, & Miyashita et al., 2000; Obleser, Meyer, & Friederici, 2011; Rogalsky, Matchin, & Hickok, 2008). When the expected harmonic sequences were violated, the greater activation of IFG has been shown in bilateral IFG (Mass, Koelsch, Gunter, & Friederici, 2001; Koelsch, Gunter, &
von Cramon et al., 2002; Tillmann, Janata, & Bharuch, 2003). A metric complex rhythm condition may be similar to this case, whereas a metric simple rhythm condition which does not induce any disruption. Comparison between metric simple and complex rhythm conditions may suggest an expected familiar experience relative to unexpected disrupted experience in temporal processing. Thus, Beat Accented rhythm compared to Non-beat Accented rhythm significantly reduced the activation of IFG.

Although modulated brain regions, such as left IFG and left STG, have been studied in this context, right IFG activation has not been fully understood in the fMRI study detailed in Chapter 4. The right IFG has been suggested to be activated during time perception (Rao, Mayer, & Harrington et al., 2001), inhibition (Aron, Fletcher, & Bullmore et al., 2003; Chikazoe, Konishi, & Asari et al., 2007; Hampshire, Chamberlain, & Monti et al., 2010), and attention (Hedden & Gabrieli, 2006, Weissman, Roberts, Wisscher, & Woldorff, 2006), therefore, it was important to define whether the activation of the right IFG is also related to higher cognitive processes during rhythm perception (this issues is further discussed in the next section).

The decreased activation of both left IFG (BA44) and left STG in Beat Accented rhythms compared to Non-beat Accented rhythms is, however, mainly due to the suppressed activation of the same rhythm pairs of Beat Accented rhythms. Interestingly, the mean percent change in BOLD response between the same and different rhythm pairs showed that the same rhythm pairs significantly decreased the percent signal change in BOLD response of activation compared to the different rhythm pairs in Beat Accented rhythms. Although the left IFG did not show that much significance, the patterns of suppressed BOLD response were similar. Although the same rhythms were repeated in both Unaccented and Non-beat Accented rhythm conditions, the percent signal change in BOLD response of left IFG in same rhythm pairs was not
significantly decreased compared to the different rhythm pairs. In short, when a metrical organisation of a binary pattern becomes ambiguous or disrupted, the suppression of BOLD response is not found in rapid, complex rhythm perception.

Regarding sensory processing, it has been suggested that higher sensory areas will be affected by the prediction (i.e., predictive coding model, Friston, 2008, 2010), in that minimised predictive error or enhanced prediction. Beat-based rhythms may reduce free energy, which can be understood as the percent signal change in BOLD response of brain activation correlated with effort. Within this theoretical framework, enhanced perception of a binary pattern due to Beat Accents may enhance top down processing of metrical organisation, which consequently suppresses the brain activity the same rhythm pairs of Beat Accented rhythms.

The evidence of suppression in the same rhythm pair of Beat Accented rhythms is consistent in the repetition suppression in perceptual expectation (Summerfield, Tritschuh, & Monti et al. 2008) and in the learning process (Büchel, Coull, & Friston, 1999). Rhythm perception in the current thesis requires the duration of seconds to perceive each event; therefore, the mechanism of repetition suppression in the rapid, complex rhythm perception can be different from that of other modalities. For example, although the same rhythms were presented in the other two conditions of Unaccented and Non-beat Accented rhythms, there was no repetition suppression between the same and different rhythm pairs, suggesting a particular role of Beat Accented rhythms in enhancing the metrical organisation of a binary pattern.
CHAPTER 6. General Discussion and Conclusions

Parametrically increased BOLD response from C2 to C3 to C4

The results in Chapter 5 showed that the increase of metrical grouping levels from C2 to C3 to C4 increase the activation of bilateral IFG. In Chapter 4, the role of right IFG has not been further understood, but in Chapter 5, the increased activation pattern of right IFG suggests that the right IFG can be related to both higher order cognitive processing and holding attention. BG activation during metre perception in the current thesis showed that the BG corresponds to the different types of metrical organisation. In a similar way, the right IFG has been found to be activated in attention, timing, inhibition, or working memory. In the current thesis the right IFG was also modulated by the metrical grouping levels in rhythm, suggesting its role in higher order cognitive processing relating to hierarchical rhythmic structures.

Regarding repetition suppression, the significantly less activation in C2 than other conditions suggests that the most frequently repeated temporal pattern creates reduced the brain activation. Within the same number of isochronous sequences, the frequency of repetition and the duration or interval of temporal patterns of C2, C3, and C4 are negatively correlated. The shorter the metrical grouping patterns are, the more they repeat. In either case, the brain activation in bilateral IFG has been found to be significantly decreased in C2 as described in Chapter 5.

In Chapter 4 and 5 it was shown that bilateral IFG is involved in beat and metre perception. In particular, metrical organisation modulated the activation of IFG. The left IFG (BA44) activation was significantly reduced when metrical organisation was enhanced in Beat Accented rhythm condition (Chapter 4) and C2, with a binary pattern (Chapter 5). The right IFG (BA9/45) has been found to be activated during both beat and metre perception, but the activation was not modulated during the rapid, complex beat-
based rhythm perception. The increase of metrical grouping levels in isochronous sequences significantly increased the activation of right IFG (BA9).

6.2. The role of IFG in beat and metre perception

Hierarchical temporal processing in relation to various functions in language processing has previously been shown in Broca’s area, in particular. Syntactic complexity in sentence processing has been found to activate the left IFG (Caplan, Alpert, Waters, & Olivieri, 2000; Obleser, Meyer, and Friederici, 2011; Rogalsky, Matchin, & Hickok, 2008), suggesting that the left IFG has a role in grammatical processing, where grammar is organised in a hierarchy. Functions in Broca’s area have been suggested as being dissociated as well, in that syntactic tasks activate the left IFG (BA44) and semantic tasks activate the left IFG (BA45) (Friederici, 2012). The region has been found to be important during speech perception (Wilson, Saygin, Sereno, & Iacoboni, 2004). The left IFG (BA44) has also been found to be involved in phonological processing (Buchsbaum, Hickok, & Humphries, 2001; Burton, 2001; Zatorre, Evans, Meyer, & Gjedde, 1992; Zatorre, Evans, & Meyer, 1994) and motoric functions (Heim, Eickhoff, & Amunts, 2008).

The IFG activation in the current thesis supports its role in higher order cognitive processing in music, however, the left and right IFG functions may need to be further dissociated depending on the hierarchical levels of the metrical organisation. For example, in Chapter 4, the left IFG (BA44) was been found to be activated in the metrical organisation between the subdivisions of beat and beat levels, whereas the right IFG (BA9) was not different across conditions, suggesting that the right IFG might correspond to temporal regularity, which is underlying the rhythm conditions used.
CHAPTER 6. General Discussion and Conclusions

In Chapter 5, the left IFG (BA44) and the right IFG was found to be increased by the increased metrical grouping levels from C2 to C3 to C4. The metrical grouping levels indicate the modulation in terms of the repetitive frequency of the intervals between strong beats. Repetition often happens in temporal structures in music, where the same or different musical phrases repeat accompanied by metrical grouping. Metrical organisation underlying music can be a factor in the decreased percent signal change in BOLD response of higher cortical areas, such as IFG, as shown in the current study. If the percent signal change in BOLD response of brain activation is reduced by the repetition of the same temporal pattern, such as the metric grouping underlying the rhythm sequences, the accumulative brain activation might be significantly less in total than the information that does not have underlying repetitive temporal patterns.

If suppression in the brain results in higher cognitive processing due to expectation or prediction while the same information repeats, there is also a possibility that the higher order cognitive processing during the repetition of the same information also can enhance brain activation. For example, the perceptual expectation increased by metrical organisation in rhythms may increase attention, which sequentially increases neural activity during the entraining to future events.

Thus, two mechanisms can be considered to activate left IFG (BA44) in metrical organisation. Firstly, the suppression of the activation in the left IFG may happen when metrical organisation becomes salient, which suggests the reduced complexity or minimised prediction error due to top-down processing involved. In this case, the increased activation of the left IFG is due to the increased complexity or increased prediction error. Secondly, the enhanced or increased activation in the IFG can happen with the increase of attention, which is not necessarily related to an increase in complexity, such as emotional appreciation in music or the proficiency of musicians in a
music relevant task. Also, increased prediction in terms of cognitive processing in hierarchical rhythmic structures can activate the IFG. The IFG (BA45) activation in beat perception (Grahn & Brett, 2007) may be an illustration of this.

The former binary pattern in Study 1 may contribute the top-down processing of metrical organisation more than bottom-up processing of rhythmic grouping in complex rhythm perception. The latter binary pattern in Study 2 may increase perceptual expectation due to the repetition of the same rhythmic patterns during metre perception in isochronous sequences. Two different mechanisms are considered here: the former, minimising predictive error; the latter, optimising anticipation rather than minimising predictive error because in this case, there was no need to detect error, as this often happening in normal music listening.

The activation difference between Beat Accented and Non-beat Accented rhythms can be considered in terms of metric grouping versus rhythmic grouping. It has been suggested that the perception of pattern in relation to hierarchical processing was left lateralised in speech studies (Scott and Wise, 2003). With respect to hierarchical processing of temporal structure, temporal regularity in rhythms is a basis on which to build them up to higher levels of temporal structure, such as metre. Unaccented and Beat accented rhythms can be easily grouped from single beats into metre of 2/4 or 4/4, however, Non-beat Accented rhythms may not be metrically grouped into the higher levels due to Non-beat Accents. Unaccented, Beat Accented, and Non-beat Accented rhythms have subdivisions of beats, and those serials of subdivisions can be grouped differently depending on accent types, although all rhythm was based on beats, and this aspect of rhythmic stimuli can be possibly linked to the activation of the left IFG (BA44).

Regarding top-down and bottom-up processing, the study results in Chapter 4 can be understood as an interactive cognitive processing model depending on the accents,
whereas Chapter 5 result are effectively all top-down processing. The significantly lower level of activation in Beat Accented rhythm (Study1) and Metrical grouping level of 2 beats (Study 2) suggests that neural mechanisms depend on attention and prediction. As discussed in Chapter 2, the two attending modes of future-oriented and analytic attending mode can place differing demands on rhythm perception, each of attending mode can also be understood relating to top-down and bottom-up processing. For example, in Chapter 4, the perception of complex rhythms can place more bottom-up processing in order to infer beat, whilst in Chapter 5, the cognitive processing of metrical organisation in isochronous sequences can employ more top-down processing during the hierarchically organised metrical processing. In the study of beat perception, the complexity was defined with respect to the types of accent, which differentiate the location of attention. In the study of metre perception, the complexity was defined by the duration of attention, which was parametrically increased by metrical grouping levels.

Significantly reduced brain activity in both Beat Accented rhythms in Chapter 4 and a binary metric pattern in Chapter 5 needs to be further investigated whether it can suggest facilitation of cognitive and motor functions. For example, in the current thesis, neural activities of these two rhythmic structures may suggest a default neural network when an additional task was not added. When additional cognitive tasks were put on those rhythmic structures, it might produce better performance, but whether or not these better performances would be related to greater activity should be further studied. Thus, other rhythmic cognitive tasks, in addition to replicating the perception of solely rhythmic structures, should be measured in both healthy non-musician volunteers and other participant groups with an interest.
6.3. Limitations of the current thesis

Although the current study demonstrated that the left IFG was significantly involved in metrical origination, the role of right IFG in metrical organisation was not fully understood. The right IFG in beat perception within complex rhythms presented in the Chapter 4 showed significant activation across all rhythm conditions, but there was no significant difference between conditions of Unaccented, Beat-Accented and Non-beat Accented rhythms. This suggests that the right IFG activation may not be involved in local accents in rapid, complex rhythms. The parametric increase of the right IFG in metre perception within isochronous sequences in Chapter 5 suggests that the right IFG is not only responding to temporal regularity, but is also modulated by higher-level metrical organisation. In this case, however, the increase and decrease activation pattern in the brain is found at the whole brain level, rather than at the localised brain activation, suggesting that the right IFG activation may need to be understood regarding in the whole brain networks.

The movements or the preparation of the movement related to a button press in the current thesis may not completely be free from the potential motor function of left IFG (BA44), of which motor functions have been suggested in several previous studies. Although passive listening without cognitive tasks can lead unexpected results due to uncontrolled free thought or feeling in music, it may have a benefit to control motor function. It may well be challenging to come up with a way of dissociating motor and cognitive functions in the brain regions because those two processes often happen simultaneously without actual movement, as previous studies have shown. Converging evidence has shown left IFG (BA44) activation in both motor and cognitive functions in various cognitive processing tasks. Although research in dissociating motor and
cognitive functions in left IFG is beyond the scope of the current thesis, this may need to be considered in the future.

The fMRI designs of both block and slow-event related design used a fixed inter-stimulus-interval (ISI) between stimuli blocks or events. Not using jittered ISI may increase the participants’ anticipation or guessing about the target stimuli. However, as the following fixation period is enough duration not to overlap haemodynamic response, this anticipatory effect of individual participant might not be significant to calculate the BOLD response of each block and event. The same period of anticipation that was given for all rhythm conditions may cancel out the anticipation effect at a same level. The events in both designs were randomised in the session, but the order of events was not randomised across participants. This is partly supplemented by the simple designs implemented in this thesis. For example, order effects, which were mostly observed during learning tasks, can be critical in terms of the randomisation across participants, although the participants are in a group. As auditory stimuli employed in this thesis were very easy and simple, the effect by learning may not happen. In fMRI experiment particularly the fatigue is an important factor, which needs randomisation across participants. In this thesis, 12 rhythmic stimuli were repeated four times in Study 1, and 4 rhythmic stimuli were repeated eight times in Study 2. When the rhythmic stimuli were repeated, each 12 trials in Study and each 4 trials in Study 2 were randomised within a cycle, thus there was no case that a certain trials were posited in the beginning or ending of the session.

Finally, although two fMRI studies of the current experiment examined beat and metre perception based on the hierarchical relationship between two levels in theories, hierarchy in the cognitive processing between beat and metre has not been directly compared. Theoretically, these hierarchical relations and beat perception within complex
rhythms suggest some of hierarchical processing, and C2 and C4 in metre perception within isochronous sequences included the hierarchical processing in rhythmic structures, partly. Higher order cognitive processing regarding the hierarchical relationship between beat and metre needs, however, to be directly examined with respect to whether or not the left IFG (BA44) activation is modulated in the same way as the results of the current thesis.

6.4. Implications and future directions

The current thesis demonstrated that bilateral IFG was involved in resolving rhythmic complexity in both beat and metre perception, where the left IFG was significantly involved in discriminating accented complex rhythms. In addition, Beat Accented rhythm condition showed significantly suppressed BOLD response compared with Unaccented and Non-beat Accented rhythm condition, when the same rhythm pairs were compared. These results support theoretical frameworks regarding predictions and free energy principle in the brain in term of predictive information, where the neural mechanisms of left IFG imply that perceiving temporal regularity during auditory perception may place positive and facilitating effect on other cognitive and motor domains, although the positive effect has not been presented in behavioural data in this study. For example, the patients with Parkinson’s disease (PD) compared with healthy volunteers had less benefit during beat perception (Grahn & Brett, 2009). Future research into whether or not the perception of Beat Accented and Non-beat Accented rhythm conditions can differently be impaired in the patients with PD can shed light into further understanding about the cortico-subcortical neural mechanisms of the patients with PD.
CHAPTER 6. General Discussion and Conclusions

IFG has been suggested to have both motor and cognitive functions in numerous previous studies, suggesting dissociating two functions in this region is important for not only rhythm processing but also other cognitive processing. In this thesis, motor preparation during a task performance in the fMRI studies could not be completely free from potential brain activation related to motor movement of button press, suggesting potential compounds of both higher cognition and motor-relevant functions. Both perception (i.e., passive listening to rhythmic sequences) and production (i.e., finger movements) has not been able to be solely clarified due to the factor of button press or the preparation for a button press (i.e., it was assumed that participants prepared for which button they would press during listening passages). Although subtraction analysis is supposed to cancel out motor factors, it will be still important to keep developing the experimental design to dissociate two cognition and motor functions in IFG. Although a task including a button press has been used in numerous studies, for example, activation of higher cortical areas has been shown to be involved in higher order cognitive processing even without tasks (Hasson, Nir, & Levy et al., 2004).

The activation in visual areas during rhythm perception needs also to be better understood. Although there was no visual stimuli presented, the activation of visual areas (e.g., lingual gyrus (BA19) during the auditory perception of rhythms requires further research, in particular, whether or not rhythm perception enhances visual processing. For example, the activity of cunues (BA18) in Beat Accented rhythms in Chapter 4 may suggest a potential correlation between enhanced temporal regularity and visual areas during auditory perception.

Finally, the experimental results in the current thesis offer evidence regarding the cognitive processing of metrical organisation in healthy non-musicians. Based on the
CHAPTER 6. General Discussion and Conclusions

research findings of this thesis, comparison studies with musicians or other groups, such as age, culture, or specific patients group, can be compared.

6.5. Conclusions

This thesis investigated the role of IFG during beat and metre perception. Converging research findings suggest that the most predictable rhythmic patterns - Beat Accented rhythms (Chapter 4) and the binary pattern of metrical organisation (Chapter 5) - significantly decrease activation in the left IFG (BA44), suggesting that the left IFG (BA44) may have a role of higher order cognitive processing in rhythm perception. In particular, a binary pattern of strong and weak beats in metric rhythms has been found to have the most significantly suppressed the activation in the left IFG (BA44). This suggests that the significantly suppressed brain activity during the perception of binary beat pattern in rhythm sequences may relate to the efficiency in information processing via optimising prediction (Friston, 2010). Predictability is one of the key mechanisms of rhythm processing (Vuust & Witek, 2014), and of further music processing such as expectation and emotional response (Salimpoor, Zald, & Zatorre et al., 2015).

Positive effects relating to various rhythmic behaviours, in particular, during higher order cognitive processing in metrical organisation in rhythms may result in the significantly suppressed brain activation. This, however, should be further investigated. How the suppressed brain activity due to the benefits of cognitive processing in rhythms can be linked to or transferred to motor functions should be also further studied. We better tap to salient beat patterns, and whether better perception improves motor movements, or vice versa remains future studies as well. In these contexts, rhythm perception departing from its neural correlates may be a useful tool to develop the
cognitive processing of predictability, which is the basis of human information processing.
References


References


References


References


Dreher, J-C., & Grafman, J. (2002). The roles of the cerebellum and basal ganglia in timing and error prediction. *European Journal of Neuroscience, 16*, 1609-1619.


References


References
References


References


175

References


initialisation, synchronisation, and continuation phases of paced finger tapping. 
Neuropsychologia, 42, 1301-1312.


References


References


Appendices

Appendix 1. Behavioural results in Chapter 5

Behavioural results in Chapter 5. The results showed that the accuracy was above 90% for all conditions, suggesting that participants correctly understood the task (Although all participants correctly responded to non-grouping isochronous beat sequences as 1, several participants verbally reported after fMRI scanning that they subjectively interpreted the non-grouping isochronous beats as different metrical grouping levels of 2, 3, or 4 beats)
Appendix 2. Brain activations that were parametrically modulated from to C2 to C3 to C4 in Chapter 5

<table>
<thead>
<tr>
<th>Brain regions</th>
<th>Talairach coordinates</th>
<th>T value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x  y  z</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right insula (BA13)</td>
<td>46 -15 5</td>
<td>10.907</td>
<td>0.001</td>
</tr>
<tr>
<td>Right Middle frontal gyrus (BA9)</td>
<td>48 15 30</td>
<td>5.710</td>
<td>0.001</td>
</tr>
<tr>
<td>Right Culmen</td>
<td>34 -55 -25</td>
<td>6.838</td>
<td>0.001</td>
</tr>
<tr>
<td>Right Superior parietal lobule (BA7)</td>
<td>31 -49 40</td>
<td>7.533</td>
<td>0.001</td>
</tr>
<tr>
<td>Right lingual gyrus (BA19)</td>
<td>9 -66 0</td>
<td>9.674</td>
<td>0.001</td>
</tr>
<tr>
<td>Left Medial frontal gyrus (BA6)</td>
<td>-8 -1 55</td>
<td>6.963</td>
<td>0.001</td>
</tr>
<tr>
<td>Left Cuneus (BA18)</td>
<td>-10 -88 18</td>
<td>6.094</td>
<td>0.001</td>
</tr>
<tr>
<td>Left Culmen</td>
<td>-9 -61 -2</td>
<td>7.295</td>
<td>0.001</td>
</tr>
<tr>
<td>Left insula (BA13)</td>
<td>-35 21 3</td>
<td>7.225</td>
<td>0.001</td>
</tr>
<tr>
<td>Left Precentral gyrus (BA6)</td>
<td>-32 -7 52</td>
<td>6.902</td>
<td>0.001</td>
</tr>
<tr>
<td>Left Suprimarginal gyrus (BA40)</td>
<td>-39 -43 38</td>
<td>6.999</td>
<td>0.001</td>
</tr>
<tr>
<td>Region</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Left Transverse temporal gyrus (BA41)</td>
<td>-47</td>
<td>-24</td>
<td>12</td>
</tr>
<tr>
<td>Left insula (BA13)</td>
<td>-41</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Left precentral gyrus (BA4)</td>
<td>-48</td>
<td>-7</td>
<td>43</td>
</tr>
</tbody>
</table>
Appendix 3. Scripts for the protocol of two fMRI studies

Table 1. Experimental procedure

<table>
<thead>
<tr>
<th>Experimental procedure</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-scanning session</strong> (30 minutes)</td>
<td></td>
</tr>
<tr>
<td>• Complete two signed forms - MRI consent form - MR screening form - Edinburgh handedness inventory (Participant Information Sheet will be prepared)</td>
<td>10 minutes</td>
</tr>
<tr>
<td>• Learn how to perform the experiments with Practice trials</td>
<td>15 minutes</td>
</tr>
<tr>
<td>• Learn an fMRI scanning experience environment (a colour picture is prepared, and a mock scanner can be visited)</td>
<td>5 minutes</td>
</tr>
<tr>
<td>(a short break)</td>
<td>5 minutes</td>
</tr>
<tr>
<td>(Changing clothes)</td>
<td>5 minutes</td>
</tr>
<tr>
<td><strong>Scanning</strong> (50 minutes)</td>
<td></td>
</tr>
<tr>
<td>• Scanning</td>
<td></td>
</tr>
<tr>
<td>- Scanning brain structure: about 5 minutes (Practice session is to check comfort, button press, &amp; sound stimuli before each session)</td>
<td>50 minutes</td>
</tr>
<tr>
<td>- Practice Session 1: 1 minute</td>
<td></td>
</tr>
<tr>
<td>- Session 1: about 17 minutes</td>
<td></td>
</tr>
<tr>
<td>- Practice Session 2: 1 minute</td>
<td></td>
</tr>
<tr>
<td>- Session 2: about 16 minutes</td>
<td></td>
</tr>
<tr>
<td>(Changing clothes)</td>
<td>5 minutes</td>
</tr>
<tr>
<td><strong>Post-scanning session</strong> (10 minutes)</td>
<td></td>
</tr>
<tr>
<td>• Complete two forms</td>
<td></td>
</tr>
<tr>
<td>- Questionnaires for fMRI study and personal information</td>
<td>10 minutes</td>
</tr>
<tr>
<td>- Expense reimbursement</td>
<td></td>
</tr>
<tr>
<td><strong>Total duration</strong></td>
<td>115 minutes (approx. 2 hours)</td>
</tr>
</tbody>
</table>

(Starting time: am / pm)
(Greeting)

“Hello, nice to meet you. My name is Sujin Hong. Thank you for your participation in this fMRI listening experiment.”

“We will move on to a preparation room now.”

“I’m going to read the instructions out using this sheet – if you have any questions don’t hesitate to ask.”

“To begin with you will have a computer and headphones, and around 15 minutes to practice two sessions of the experiment.”

“After a short break we will move on to the actual scanning session, which will take about 50 minutes.”

“Finally, once you have finished your scan we will have 10 minutes to discuss how it went.”

“Please ask me as we go along if you have any questions or concerns. Do you want to ask me anything before we move on?”

(Getting forms)

“There are short forms for you to fill out before we begin the experiment.”

(MR consent form, Screening form, Edinburgh handedness inventory; if the participant has a question with regard to a Screening form, the question can be further answered by radiographers.)

“I have already sent you the Participant Information Sheet (PIS). If you did not read the PIS beforehand, I have one for you to read now.”

(Participant Information Sheet read)

“Thank you. Now I will explain what I want you to do next.”

(Training session)

“There are two sessions in this experiment. You will perform a task during the sessions. Before you begin two sessions, there will be 5 minutes more of scanning to take an image of your brain structure. Total scanning time will be about 50 minutes.”
(Training session 1)

“In session 1, you are going to listen to beat sequence of notes, which lasts a few seconds, and judge if they are grouped by 1, 2, 3, or 4 beats. After you listen to the beat sequences, you press a button to indicate your response. As you can see (button picture) here, press button 1 to indicate beat sequence grouped by 1 beat, press button 2 to indicate beat sequence grouped by 2 beats, button 3 to indicate beat sequence grouped by 3 beats, and button 4 to indicate beat sequence grouped by 4 beats (SH showed how to press button).”

- You WILL NOT MOVE your fingers to count 1, 2, 3 or 4 beats.
- Also, you will use only your RIGHT fingers to press button.
- If you make a mistake, don’t worry about it. Just wait for the next rhythms.

“I am going to play each beat sequence. Please listen to a beat sequence grouped by 1 beat. There are two types of this beat sequence (An experimenter play both low and high pitch starting beat sequences, and check if a participant understand sound stimuli).”

“Please listen to a beat sequence grouped by 2 beats. There are two types of this beat sequence (An experimenter play both low and high pitch starting beat sequences, and check if a participant understand sound stimuli).”

(Most participants will understand sound stimuli in this stage, but an experimenter keeps playing remained two conditions)

“Please listen to a beat sequence grouped by 3 beats. There are two types of this beat sequence (An experimenter play both low and high pitch starting beat sequences, and check if a participant understand sound stimuli).”

“Please listen to a beat sequence grouped by 4 beats. There are two types of this beat sequence (An experimenter play both low and high pitch starting beat sequences, and check if a participant understand sound stimuli).”

“Now, I will play sound samples, and would you please practice buttons to indicate beat sequence grouped by 1, 2, 3 or 4?” (an experimenter randomly plays sound samples without a fixed particular order, and confirmed that the participant understands sound stimuli completely)
“Great, well done. Now, We will practice exactly what you will do, in here with the computer. I will give you 3 trials.”

“In the beginning of each session, there will be a short silence and you will see a blank screen with a white cross. You will see the white cross during the experiment too. Just let your eyes rest on the cross.”

“Do you have any question for the session 1?”

(Training session 2)

In session 2, you will listen to a pair of rhythms and judge if they are same or different. After you listen to the rhythms, you press a button to indicate your response. As you can see here (a picture of MR compatible button box, which is an exactly same one for fMRI session), you will press button 1 for the same, button 2 for the different case. The order of button press can be switched if you prefer a different option.”

(Training session finished)
“Great, thank you.”
“Thank you. Let’s go through to the mock scanner now, and I will explain how the main part of the experiment will work.”

“Great, thank you.”
“Thank you. Let’s go through to the mock scanner now, and I will explain how the main part of the experiment will work.”

(Go to mock scanner; the script of Mock Scanner in this page has been originally written by Kirsteen David Kelly, and this is used for standard fMRI protocol for CRIC-IMHSD neuroimaging studies to introduce MR scan to participants. The current thesis used this script by modifying the procedures of experimental session to introduce MR scan procedure for participants.)

“This is a simulation of the real scanning experience; the big difference is that in the real scanner it will be very noisy. The scanner makes a repetitive and quite loud noise during the actual experiment. Going into the scanner for the first time can seem a bit daunting, but it does not take long to get used to it. The noise should not be uncomfortable but it is very loud. You will get noise cancelling headphones to wear. You might find that it actually becomes quite soothing – some people fall asleep, but we are going to ask you to concentrate for the tasks during the scan. Please make sure that you wear earplugs in loosely though, not too much tightly, and also make sure that headphones cover your ear completely so that you can listen to sound better.”

“We can talk to you through the headphones and we can hear you talking to us before and after scanning. During the scanning, this is not available due to the MR noise. There will be an emergency buzzer that you can press ’AT ANY TIME’ if you wish to stop the experiment. You will also be given goggles to wear, which will present written instructions on a screen. You will also have a button box that you already practiced with.”

“It is important to collect image data while you are still in the scanner. You need to stay absolutely still once you are in position, so it is important to get comfortable before we start. It is absolutely crucial that you do not move during scanning – especially not your head, shoulders, elbows or wrists. When scanning finished, you can relax, and once each scan is complete we will check that you are ok to continue”
“Once you are inside the scanner you should make sure that your hands are comfortable to press the buttons. The radiographer can provide you with foam blocks for your elbows or wrists if you like.” (written by Kirsteen David Kelly)

“Before the beginning of each session, I will check your headphones and button response. Also, before the beginning of each session, I will give you a short MR noise practice for about 1 minute, which will make you get used to the noise, button box, goggles and headphones.”

“This experiment will start a scan session to get your brain structure for 5 minutes first. Then, we will have practice session 1 for 1 minute, and Session 1 will last about 16 minutes. After a short break, we will have practice session 2 for 1 minute, and session 2 will last about 17 minutes. I will remind you the instruction before we start each session.”

(5 minutes break; SH will pass the signed form to radiographers)

“Ok, now we will have a short break, about 5 minutes. We will meet here in the preparation room again.”

(MR Scanning; radiographers will help participants to lie on the bed of MR scanner, to wear MR kits (e.g. headphones, goggles, and button box) properly.)

“Are you comfortable and ready to begin?”

Session 1
(a PPT file: ‘Please relax (slide1)’)
(a PPT file: ‘Headphones check (slide2)’)
“I will check your headphones first. Can you hear me?
“I will play sound samples for you.”
“Can you hear the rhythms?”
(a PPT file: ‘Button check (slide3)’)
“Now, I will check your button response. Please press button to indicate beat sequence grouped by 1, 2, 3 or 4 beats serially - Great.”
“Now, I will talk you through the instruction again. You will listen to beat sequences, and will judge if they are grouped by 2, 3, or 4. You press button 1 when a beat
sequences is not grouped, press button 2 when the beat sequence is grouped by 2 beats, press button 3 for beat sequences grouped by 3 beats, and press button 4 for beat sequences grouped 4 beats. If you make a mistake, don’t worry about it. Just wait for the next rhythms”

“We are going to carry out a short practice session 1 for about 1 minute.”

(Practice session finished)

“Is everything ok?”

“Ok, we will start the first session for 16 minutes. Please be ready to start.”

(16 Minutes scanning finished)

“Thank you. We have now finished Session 1, well done. Are you comfortable in there? Please relax for a few moments and then we will start the next session.”

(Preparation for about 30 seconds)

Session 2

“Ok, are you comfortable and ready to begin? Now we will start Session 2 shortly.”

(a PPT file: ‘Please relax (slide1)’)

(a PPT file: ‘Headphones check (slide2)’)

“I will check your headphones first. I will play sound samples for you.”

“Can you hear the rhythms?”

(a PPT file: ‘Button check (slide3)’)

“Now, I will check your button response. Please press button to indicate the same case. Please press button to indicate the different case - Great.”

“I will talk you through the instructions again. You will listen to a pair of rhythms, and will judge if they are the same or different. You will press button 1 when the rhythms are the same, and press button 2 when the rhythms are different. If you make a mistake, don’t worry about it. Just wait for the next rhythms.”

“We are going to carry out a short practice session for about 1 minute.”

(Practice session finished)

“Is everything ok?”

“Ok, we will now start the Second session which last for17 minutes. Please be ready to start.”

(17 Minutes scanning finished)
“Thank you. We finished all sessions, and we will come through now and get you out.”

(Post scanning session)

“Were you comfortable during the scanning? How did you feel?”
“I have two forms that you will fill out. Feel free to ask any questions.”
(Two forms of fMRI questionnaire and Expense reimbursement are completed)
Additional questions are asked to participants (verbally)
1. “In Session 1, how did you perform a task? Did you count beat sequences verbally? Or did you use your fingers or toes to count?”
2. “When you count beat sequence of 1 beat, have you counted it in a different way such as 2, 3, or 4 beats?”
3. “Can you say which beat sequence was the most comfortable of preferable?
4. “In Session 2, have you noticed accents in some rhythms during your listening in MR scanner?”

“Thank you again for your participation in this fMRI experiment. When the study has been completed we will send you a short report of our findings. Bye now.”
Appendix 4. Edinburgh Handedness Inventory

Edinburgh Handedness Inventory (Oldfield, 1971)

Please indicate your preferences in the use of hands in the following activities by putting ✓ in the appropriate column. Where the preference is so strong that you would never try to use the other hand unless absolutely forced to, put ✓✓

If you are really indifferent, put ✓ in both columns.

Some of the activities require both hands. In these cases the part of the task, or object, for which hand preferences is wanted is indicated in brackets.

Please try to answer all the questions, and only leave a blank if you have no experience at all of the object or task.

<table>
<thead>
<tr>
<th></th>
<th>LEFT</th>
<th>RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Writing</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Drawing</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Throwing</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Scissors</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Toothbrush</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Knife (without fork)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Spoon</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Broom (upper hand)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Striking match (match)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Opening box (lid)</td>
<td></td>
</tr>
</tbody>
</table>

i Which foot do you prefer to kick with?

ii Which eye do you use when using only one eye?
Appendix 5. Questionnaire for fMRI study and personal information

QUESTIONNAIRE for fMRI study and personal information

1. If you have any opinions after your fMRI sessions, please describe them briefly. 
________________________________________________________________

2. What year were you born? ____________________________

3. Please circle your gender. ( Female  male )

4. Have you had musical training (e.g. instrument, voice, etc.)? Y / N
   4-1. If yes, when did you start your musical training? Age _______
   4-2. How long did you continue your musical training? ________ years
   4-3. Can you describe your musical training? (e.g. 1 year of flute training, 2 years of piano training, etc.)
       ______________ ________________________________________

5. If English is not your first language, what is your first language? 
________________________________________________________________

5-1. If English is your second language, how long have you learned English? ________ years

6. Have you ever experienced learning language using music, regardless of (native or foreign) language? Y / N
   6-1. If you have, did using music assist you in learning the language? Y / N
   6-2. If you answered yes to 6-1, could you please describe why it was helpful? 
       ___________________________________________________________