Perceptual functions of auditory neural oscillation entrainment

Perceptual functions of auditory neural oscillation entrainment

By Andrew Chang, Bachelor of Science

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University
Doctor of Philosophy (2019)
Hamilton, Ontario (Psychology, Neuroscience & Behaviour)

TITLE: Perceptual functions of auditory neural oscillation entrainment
AUTHOR: Andrew Chang (McMaster University)
SUPERVISOR: Dr. Laurel J. Trainor
NUMBER OF PAGES: xviii, 134

Lay Abstract

Perceiving speech and musical sounds in real time is challenging, because they occur in rapid succession and each sound masks the previous one. Rhythmic timing regularities (e.g., musical beats, speech syllable onsets) may greatly aid in overcoming this challenge, because timing regularity enables the brain to make temporal predictions and, thereby, anticipatorily prepare for perceiving upcoming sounds. This thesis investigated the perceptual and neural mechanisms for tracking auditory rhythm and enhancing perception. Perceptually, rhythmic regularity in streams of tones facilitates pitch perception. Neurally, multiple neural oscillatory activities (high-frequency power, low-frequency phase, and their coupling) track auditory inputs, and they are associated with distinct perceptual mechanisms (enhancing sensitivity or decreasing reaction time), and these mechanisms are coordinated to proactively track rhythmic regularity and enhance audition. The findings start the discussion of answering how the human brain is able to process and understand the information in rapid speech and musical streams.

Abstract

Humans must process fleeting auditory information in real time, such as speech and music. The amplitude modulation of the acoustic waveforms of speech and music is rhythmically organized in time, following, for example, the beats of music or the syllables of speech, and this property enables temporal prediction and proactive perceptual optimization. At the neural level, external rhythmic sensory input entrains internal neural oscillatory activities, including low-frequency (e.g., delta, 1-4 Hz) phase, high-frequency (e.g., beta, 15-25 Hz) power, and their phase-amplitude coupling. These neural entrainment activities represent internal temporal prediction and proactive perceptual optimization. The present thesis investigated two critical but previously unsolved questions. First, do these multiple entrainment mechanisms for tracking auditory rhythm have distinct but coordinated perceptual functions? Second, does regularity in the temporal (when) domain associate with prediction and perception in the orthogonal spectral (what) domain of audition? This thesis addressed these topics by combining electroencephalography (EEG), psychophysics, and statistical modeling approaches. Chapter II shows that beta power entrainment reflects both rhythmic temporal prediction (when events are expected) and violation of spectral information prediction (what events are expected). Chapter III further demonstrates that degree of beta power entrainment prior to a pitch change reflects how well an upcoming pitch change will be predicted. Chapter IV reveals that rhythmic organization of sensory input proactively facilitates pitch perception. Trial-by-trial behavioural-neural associations suggested that delta phase entrainment reflects temporal expectation, beta power entrainment reflects temporal attention, and their phase-amplitude coupling reflects the alignment of these two perceptual mechanisms and is associated with auditory-motor communication. Together, this thesis advanced our understanding of how neural entrainment mechanisms relate to perceptual functions for tracking auditory events in time, which are essential for perceiving speech and music.

Acknowledgements

I am grateful to complete my Ph.D. in a very supportive and kind laboratory filled with joy. There are many people whom I would like to thank. First of all, thank you very much to my supervisor Laurel J. Trainor from the very bottom of my heart. Thanks to her great support on my research, professional development, and networking (I must have spent over millions of dollars attending conferences). I was very lucky to freely explore the research topics I am interested in while being able to have her guidance on how to implement them. She built a perfect lab (actually, "labs") composed with skilled and nice colleagues that I could never dream of. Beyond professional aspects, I feel that Laurel is my friend. She never questioned me for attending rehearsals for Symphony on the Bay instead of working late in the lab. Doing my Ph.D. with Laurel was one of the best decisions I have ever made, even thought I was merely tempted by the Niagara Falls and winery tour she took me on when she tried to recruit me.

Thanks to my progress committee members, Dan Bosnyak and Ian Bruce, my comprehensive exam committee members, Bruce Milliken and Larry Roberts, and my external examiner, Sidney Segalowitz. Their opinions and input have always made my research better. They were always there when I needed help, and have always guided me through technical and scientific challenges.

Thanks to my lab-mates, Haley Kragness, Laura Cirelli, Chris Slugocki, Kate Einarson, Rayna Friendly, Sima Hoseingholizade, Sarah Lade, Hector Orozco Perez, David Prete, Chantal Carrillo, Dobri Dotov, Dan Cameron, Erica Flaten, and Debanjan Borthakur. They made doing science much more fun. Thanks to the lab staff members, Dave Thompson, Dan Bosnyak, Elaine Whiskin, Susan Marsh-Rollo, Carl Karichian, and Steven Livingstone. Research would have been much more challenging without their assistance. Thank you very much to my undergraduate assistants, Alexandra Rice, Michael Wan, Jessica Empringham, Michael Ku, Keeyeon Mark Hwang, Jennifer Chan, Katie Clayworth, Emily Kaunismaa, Brittany Ung, Elger Baraku, Jasmine Zhang, Jessica Otoo-Appiah, Taylor Barton, and Tessa Dickison. The data collection process would have been much much much more painful without them. Their hard work and their consent of slavery are deeply appreciated.

Specifically, I am truly indebted to Dave Thompson for making all my experiments possible, no matter how technically complicated they were. Thanks to Elaine and Susan for always being so enthusiastic. Having lunch with them was my daily retreat. Also, I enjoyed playing squash with Chris, even though failing to defeat him before completing my Ph.D. might be my biggest regret forever. Last but not least, I appreciate Haley and Laura for dragging me out of my cave to hang out with people and explore the city. They have helped me appreciate the North American bad jokes.

Thanks to the music cognition gang of my cohort, Blair Ellis, Lorraine Chuen, and Haley (again). It was truly fun to hang out with these guys, no matter if we were discussing science, music, or any crazy topic, although I still cannot appreciate the weird indie movie we have watched together (picked by Blair, btw). Thanks to all my friends in the program, Brandon Paul, Fiona Manning, Blake Anderson, Jessica Cali, Mike Galang (GO Raptors!), Mike Slugocki (don't support the Warriors), Ali Hashemi, Kiret Dhindsa, Lux Li, Ye Yuan, Anna Siminoski, Aimee Battcock, Sharmila Sreetharan, and many others. They all have made my graduate life special.

Thanks to all the Taiwanese friends I have had in Hamilton and Toronto, Fred Liao, 李曜全, 賴宣安, Leo Hsu, Allison Yeh, 陳奕全, Darren Wang, 葉婷婷, Johnson Chen, Mei-Ju Shih, Mei-Cheng Shih, 張榮勛, Lucia Huang, Yu-Sian Li, 唐婕, 蕭斌, Mandy Chu, Ming-Feng Chiang, Amy Liao, and many others. Thanks for their hospitality when I first arrived Hamilton. They all made me feel at home, except once we got lost in the woods in Dundas Peak.

Thanks to my parents, I am extremely lucky to be in a family that fully supports my career path and has a background of working in academia. They supported me to do my Ph.D. 12,104 km away from them. They have offered me great and honest guidance of academic survival skills, not to mention all the free statistical consulting sessions offered by my dad. I cannot enjoy doing research without their unconditional support and love, especially that they fed me quite a lot every time I went back home. Thanks to my parents-in-law, again, I am extremely lucky to have their unconditional trust and support, as they never questioned why the husband of their lovely daughter is still a poor graduate student at the age of 29.

Most importantly, I cannot express how grateful I am for the love of my wife Wei Vivian Tsou. She supported my goal to pursue my Ph.D. abroad,

even when we just met in Taiwan and I planned to leave in less than a year. She defied the challenges of our 40-month long-distance relationship between Hamilton and Taipei as well as Hamilton and New York. She relocated to Hamilton to start our family together after we were married. She was happy and sad for all the accomplishments and obstacles I have had, and she cheered me up whenever I was too anxious to make any moves. Her unconditional love and endless support made me a better person and that I can better enjoy this career path, which should also be granted with a Ph.D. (Push husband to Doctorate). This thesis could not be completed without her, not only because she is my proofreader, but she is also my perfect companion on this journey.

Table of Contents

List of Figures

List of Tables

Declaration of Academic Achievement

I, Andrew Chang, declare that this thesis titled, "Perceptual functions of auditory neural oscillation entrainment" and the work presented in it are my own. This thesis consists of five chapters in total and is presented in the format of a sandwich thesis. The thesis consists of a general introduction, three empirical chapters, and a general discussion. All three empirical chapters are published in peer-reviewed scientific journals.

I am the primary author of all five chapters. I conceptualized and designed each experiment in consultation with Laurel J. Trainor and Dan J. Bosnyak, who co-authored the three empirical chapters. For each study, I was the primary individual responsible for creating stimuli, collecting data, supervising data collection by undergraduate students, analyzing the data, and preparing the manuscripts. These studies were implemented with technological assistance of research staff Dave Thompson.

The LaTeX typeset of this thesis follows the McMaster Thesis Example by Benjamin Furman, license: CC BY-NC-SA 3.0. (`https://github.com/benjaminfurman/McMaster_Thesis_Template`)

This thesis includes three published research articles with permissions from the Frontiers and Elsevier Ltd.:

- Chang, A., Bosnyak, D.J., Trainor, L.J. (2016). Unpredicted pitch modulates beta oscillatory power during rhythmic entrainment to a tone sequence. *Frontiers in Psychology*, 7:327. doi: 10.3389/fpsyg.2016.00327
- Chang, A., Bosnyak, D.J., Trainor, L.J. (2018). Beta oscillatory power modulation reflects the predictability of pitch change. *Cortex*, 106, 248-260. doi: 10.1016/j.cortex.2018.06.008
- Chang, A., Bosnyak, D.J., Trainor, L.J. (2019). Rhythmicity facilitates pitch discrimination: Differential roles of low and high frequency neural oscillations. *NeuroImage*, 198, 31-43. doi: 10.1016/j.neuroimage.2019.05.007

Chapter I

General introduction

Humans constantly process fleeting auditory information, such as speech and music, in real time in everyday life. The dynamic nature of auditory signals makes it challenging to perceptually capture and process the incoming sensory information. However, auditory information in communication signals is commonly temporally structured in regular patterns, rather than randomly organized. In this thesis, rhythm refers to the temporal amplitude modulation (AM) pattern of the acoustic waveform. In isochronous rhythmic patterns, events tend to occur at regular time intervals, and this cyclical feature can be described by frequency (number of events/cycles per second) and phase (the temporal position within a cycle). Isochronous rhythmic temporal structure in acoustic waveforms is very useful for audition. This regularity greatly simplifies the burden on perceptual processing, because it enables temporal predictions toward the future, based on the past isochronous rhythmic temporal regularity, which the enables proactive temporal optimization of perceptual processing for upcoming events that are at predicted temporal positions (Haegens & Zion Golumbic 2018; Nobre & van Ede 2018).

Following the perceptual advantage for regularity, AM in speech and music acoustic waveforms, across languages and music genres, is commonly rhythmically structured (Ding et al., 2017). The important acoustic information is often highlighted in time, as it is usually positioned at regularly spaced isochronous time points. In the case of music, AM rhythm (the onset to onset times between adjacent musical notes) usually has varying lengths rather than always being isochronous. Nevertheless, these varying onset-to-onset times are usually related by small integer ratios, producing a temporal organization under an invariant isochronous unit of time (i.e., beat). Musical rhythmic

structures are often hierarchical, with beats occurring at different nested tempos (Kotz et al., 2018; Patel, 2010). In the case of speech, the inter-onset intervals of syllables are quasi-isochronously organized. As in music, there are nested hierarchical beat levels, corresponding to phoneme, syllable, word, and phrase structures (Keitel et al., 2018).

Beyond the perceptual domain, isochronous rhythmic regularity also plays a critical role in sensorimotor coordination (Repp, 2005), interpersonal coordination (Savage et al., 2015), and even prosocial behaviours (Cirelli et al., 2018). Indeed, the temporal optimization of perceptual processing is hypothesized to have an evolutionary motor origin (Kotz et al., 2018). The capacity for producing periodic or isochronous motion is nearly universal among animals, and thus perceiving these motion-generated periodic perceptual events (e.g., sound) could be critical in evolution. In the case of vocalization, the motions involved in facial displays (e.g., lip-smacking) of certain non-human primates appear rhythmic and are repeated in the same frequency range as syllables in human speech and beats in music, raising the possibility that such facial displays might be an evolutionary precursor of speech production (Ghazanfar et al., 2012). Although the current thesis does not aim to investigate the evolutionary origin of rhythmic auditory perception, the motor origin hypothesis suggests that the rhythmic organization of human audition is closely tied to benefits of perceiving periodic fluctuations in acoustic information.

Note that the use of term "rhythm" is not entirely consistent across research fields. The field of music cognition uses "rhythm" to describe the temporal pattern of AM of the acoustic waveform, which usually consists of varying onset-to-onset times between adjacent tones in music, and "beat" refers to the invariant isochronous unit of time underlying the AM rhythm (Kotz et al., 2018; Patel, 2010). Under this definition, beat is the level at which temporal prediction occurs. However, the field of temporal prediction usually uses the term "rhythm" to refer to an isochronous AM temporal pattern (Nobre & van Ede 2018), as most of the studies in this field only use isochronous AM stimuli as sensory inputs. In this case, "rhythm" and "beat" are equivalent terms, and thus temporal prediction occurs at this level. For convenience, following the field of temporal prediction, the current thesis uses the term "rhythm" to describe isochronous rhythmic patterns. Rhythm in the sense used by music cognition will be specifically referred to as "musical rhythm".

I.1 Perceptual mechanisms of auditory rhythm tracking

Dynamic attending theory, one of the most popular and successful perceptual theories related to tracking auditory rhythms, proposes that internal temporal attention is rhythmically allocated at regular temporal positions according to rhythmic regularity (i.e., places of expected maximal amplitude) in an external sensory input (Jones, 2010). Specifically, the rhythmic attending mode extracts the statistical temporal regularity of the acoustic waveform (e.g., the beats of music, the syllables of speech), and it anticipatorily directs attention to certain temporal positions (e.g., beat or syllable onsets) in the sensory input, based on the extrapolated temporal regularity (Jones & Boltz, 1989). This mechanism is described by an oscillatory model: an internal oscillator, which represents periodic and isochronous temporal expectation, adapts its period to match the external temporal regularity of the acoustic input, and an attentional pulse, which is centred at a specific phase of the oscillator (Large & Jones, 1999). During rhythmic attending, the oscillatory temporal expectation "rolls" along the temporal dimension and allocates the attentional pulse rhythmically.

Internal rhythmic expectation can be automatically, involuntarily, and exogenously (bottom-up) driven by external stimuli, with its precision depending on the temporal coherence (i.e., the degree of statistical regularity) of external events (Large & Jones, 1999; Jones & Boltz, 1989; Jones, 2010). Empirical evidence shows that the rhythmic pattern of a stimulus sequence automatically orients attention towards expected time points, independent from voluntary control (Rohenkohl et al., 2011). As the same time, rhythmic attending can be deliberately or voluntarily (endogenously, top-down) controlled to a certain extent. For example, in music, there might be hierarchically organized nested levels of temporal regularity, such as an 8 Hz, a 4 Hz and a 2 Hz beat rate. Studies have shown that participants can endogenously orient their attention toward a particular beat level of the temporal hierarchical structure or acoustic stream (Jones & Boltz, 1989). It is also possible for complex music to be ambiguous in that it can be, for example, perceived as either in groups of 2 as in a march or in groups of 3 as in a waltz. In these cases, endogenous processes are involved in beat perception (Fujioka et al., 2015). Also, while

voluntary temporal orienting of attention can be impaired by focal brain damage or transcranial magnetic stimulation (TMS), the rhythmic cues in such cases, may still be able to drive automatic involuntary aspects of temporal orientation of attention (Correa et al., 2014; Triviño et al., 2011).

Empirical evidence shows that rhythmic regularity proactively optimizes perceptual processing for events that occur at expected time positions (Large & Jones, 1999; McAuley & Jones, 2003), and the gradient of perceptual enhancement is sharpened by the increased coherence of external temporal regularity or by simpler rhythmic structure (Jones et al., 2002; Klein & Jones, 1996). Also, near-threshold deviant stimuli presented at a time point matching the preceding rhythmic context will be better perceived, compared to those presented at non-isochronous times, or in a temporally random context (Haegens & Zion Golumbic 2018; Nobre & van Ede 2018). This rhythmic facilitation effect has been observed and replicated in multiple domains. Performance accuracy for comparing the duration of two empty intervals is facilitated if it is preceded by an isochronous tone sequence (see Henry and Herrmann 2014 for a review). Also, participants' detection thresholds for tones embedded in noise are lower when the tones are presented in an isochronous sequence rather than a random sequence, suggesting that rhythmic regularity facilitates signal to noise detection sensitivity (ten Oever et al., 2017). The same effect is also observed in visual perception. The grating orientation of a visual target is better perceived when it is presented embedded in a visual isochronous rhythmic sequence than in a random visual sequence (Cravo et al., 2013; Rohenkohl et al., 2012).

I.2 Dynamic neural mechanisms of auditory rhythm tracking

One of the important goals of neuroscientific studies on tracking auditory rhythm is to understand how neural activities implement the necessary computations in real time. Given the dynamic nature of tracking auditory rhythms, this goal is usually investigated using high temporal-resolution neuroimaging approaches, such as invasive local field potential (LFP) recordings in non-human animals (e.g., monkeys), and non-invasive electroencephalography (EEG) and magnetoencephalography (MEG) recordings in humans.

I.2.1 Definition of neural entrainment

Entrainment, a term originating from complex system theory, describes that two or more independent oscillators, which share the same physical context but with independent oscillation frequencies, mutually influence each other's oscillations. The coupling force of entrainment will adjust each oscillator's frequency and phase until they are the same. In neuroscience, neural entrainment specifically describes how the external physical oscillation of a sensory stimulus (e.g., a sound, light) influences the internal neural oscillations in the brain of an individual until they have the same frequency. Note that the scope of the current thesis only includes the sensory aspect of neural entrainment, only considering the unidirectional influence of physical stimuli on an individual's neural oscillations. The reverse direction (an individual influencing the physical world) and bidirectional interactions are beyond the scope of the current thesis.

I.2.2 Neural entrainment for tracking auditory rhythm

Neural entrainment is thought to play a critical role in tracking auditory rhythm, and both the phase of the low-frequency and the power envelope of the high-frequency oscillations entrain to the rhythmic regularity of the amplitude-modulated sensory input (Figure 1).

LFP recordings from the sensory cortex (visual or auditory) of monkeys showed that AM of a sensory input stream entrains the phase of low-frequency neural oscillations (Calderone et al. 2014; Lakatos et al. 2008, 2013, 2016; Schroeder & Lakatos 2009). Specifically, the excitation phase of the neural oscillation will match the onset timings of the upcoming sensory input (Figure 1), and thus anticipatorily facilitate perceptual processing. Such neural activity can also be observed in human EEG and MEG studies, in which external isochronous or quasi-isochronous rhythmic sensory input at 1-4 Hz AM rates entrains the phase of delta (1-4 Hz) oscillations. Functionally, similar to the findings on LFP recordings in animals, the human neuroimaging studies showed that the phase angle of neural the oscillations relative to the timing of sensory input modulates the quality of perceptual processing, including detection of timing or intensity deviations, speech perception (e.g., Arnal et al., 2015; Bauer et al., 2018; Henry & Obleser, 2012; Henry et al., 2014; Herrmann

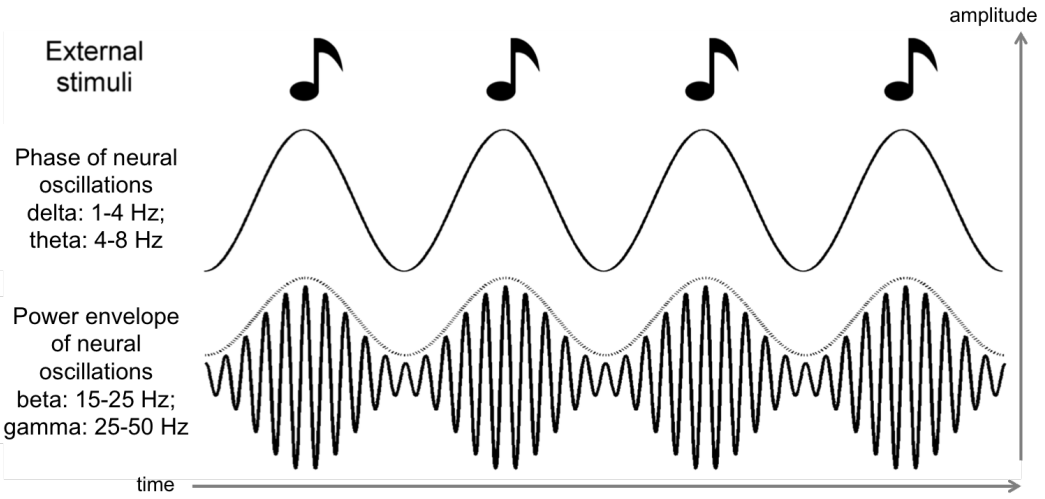

FIGURE I.1: Figure 1. Schematic illustrations of rhythmic AM of the sound entraining neural oscillations

et al., 2016; Stefanics et al., 2010; ten Oever et al., 2017), and even visual orientation discrimination (e.g., Cravo et al., 2013).

The power envelope of high-frequency oscillations can also be entrained by rhythmic AM input (Figure 1), but the perceptual functions of this entrainment have been investigated much less compared to low frequency phase entrainment. EEG and MEG recordings in human participants showed that beta (15-25 Hz) power decreases following each tone onset in an isochronous stream, and then increases with the appropriate slope to anticipate the predicted onset time of each upcoming tone, as a function of the tempo of the tone sequence, suggesting that beta power entrainment reflects isochronous rhythmic temporal prediction (Cirelli et al., 2014; Fujioka et al., 2012, 2015). This power entrainment activity can be disrupted by non-isochronous sequences (Fujioka et al., 2009, 2012), and modulated by hierarchical timing structures (e.g., waltz, march) (Snyder and Large, 2005; Iversen et al. 2009; Fujioka et al. 2015). However, prior to the research in this thesis, the associations between high-frequency power entrainment and behavioural and perceptual consequences were largely unknown.

Low and high frequency neural entrainment activities work in concert for tracking auditory rhythms. Phase-amplitude cross-frequency coupling is thought to result when the excitatory or inhibitory phase of low-frequency

oscillations modulates the power fluctuation in high-frequency oscillations (Canolty & Knight, 2010; Hyafil et al., 2015b), and this coupling is optimized for rhythmic input (Lakatos et al., 2005; Schroeder & Lakatos, 2009). Functionally, studies show that the degree of delta-beta phase-amplitude coupling associates with the accuracy of detecting an auditory temporal delay following an isochronous rhythmic sequence (Arnal et al., 2015); as well, better coupled entrainment is associated with better speech comprehension (Keitel et al., 2018).

Speech AM rates are typically above 4 Hz (Ding et al., 2017) and, as a result, theta (4-8 Hz) phase and the gamma (25-50 Hz) power entrainment, rather than delta phase and beta power, are involved in tracking speech amplitude envelopes, and are associated with speech comprehension (e.g., Doelling et al., 2014; Gross et al., 2013; Kösem et al., 2018; Pefkou et al., 2017). Specifically, gamma power is modulated at the theta rate (i.e., theta-gamma phase-amplitude coupling), and it has been hypothesized that theta oscillations reflect chunking of the continuous speech waveform into meaningful segments, whereas gamma oscillations reflect encoding of speech information (Zion Golumbic et al., 2013; Hyafil et al., 2015a; Giraud & Poeppel, 2012; Peelle & Davis, 2012). Note that the current thesis did not investigate speech perception, and the AM rates of the stimuli used in the studies were around 2 Hz; thus, the current thesis did not investigate theta and gamma entrainments.

These neural entrainment activities also reflect communication among brain regions. Delta phase entrainment activities are mainly generated from auditory cortex (Henry et al. 2014; Stefanics et al. 2010; ten Oever et al. 2017). However, intracranial electrophysiology recordings in humans showed that the posterior parietal, inferior motor, inferior frontal and superior midline frontal cortex also generate delta phase entrainment activities, reflecting communication among these regions as well as between them and sensory areas (Besle et al., 2011). Regarding beta power entrainment, auditory cortex is also the most dominant source, and its cortico-cortical coherence reflects communication between auditory cortex and sensorimotor cortex, inferior-frontal gyrus, supplementary motor area, and cerebellum (Fujioka et al., 2012). A recent MEG study further showed that bottom-up communication from the auditory cortex to the sensorimotor cortex is reflected by delta phase entrainment at the AM (musical rhythm) rates of the sensory input, and top-down communication from the sensorimotor cortex to the auditory cortex is reflected by delta

phase entrainment and beta power entrainment at the beat rate, which is the periodic temporal structure underlying the acoustic AM envelope (Morillon & Baillet, 2017). Thus, together, low and high frequency entrainment activities reflect different directions of cortical communication, and they coordinate bottom-up and top-down signals to track auditory rhythms.

I.3 Aims of this thesis

Why do humans have multiple entrainment mechanisms for tracking auditory rhythm, including low frequency phase, high frequency power, and their coupling? Do these mechanisms have different perceptual functions? Most previous studies only focused on either low or high frequency entrainment, so they could not address this question. Furthermore, the perceptual correlates of high frequency power entrainment have been investigated much less compared to those of low frequency phase entrainment. Therefore, the first aim of the present thesis is to investigate the different perceptual functions of low and high frequency entrainment activities, with an emphasis on the function of high frequency power entrainment.

It is also unclear whether neural oscillations specifically reflect prediction for what is expected to occur as well as when it will occur. Accordingly, the second aim of this thesis is to investigate whether high frequency entrainment reflects prediction for pitch. In general, sounds can vary in time, intensity, and spectral dimensions. Temporal variation includes duration, sound onset spacing, and rhythm; intensity variation constitutes patterns of sound pressure change resulting in perceived loudness changes; and spectral variation involves changes in frequency content resulting, for example, in perceived pitch or timbre changes. However, most previous studies on neural entrainment and rhythmic tracking have ignored spectral factors. It is not obvious a priori that rhythmic regularity would enhance pitch perception similarly to how it enhances time and intensity perception (Haegens & Zion Golumbic 2018), because pitch perception can be dissociated from time and intensity perception. For example, people with amusia or tone deafness typically have auditory perceptual deficits in the spectral domain but not in the temporal or intensity domains (Peretz 2016; Zendel et al. 2015). This topic is important for understanding audition, because spectral features are critical for identifying

auditory objects and understanding speech and music, which are part of the what domain of auditory perception.

Note that spectral prediction is a type of sensory prediction based on statistical regularities in the spectral domain of the audition. For example, an ascending or descending sequence of musical tones sets up the expectation that the next tone will continue in the same direction. Similar to temporal prediction, spectral regularities enable the perceptual system to predict upcoming auditory spectral information. Because spectral information informs as to object identity, it is regarded as part of the what domain of sensory prediction (Arnal & Giraud, 2012). Although what and when are physically orthogonal in audition, the information from these two dimensions is usually associated. For example, the inter-note time interval (when) tends to be longer between two phrases than within a phrase (what) in music (Kragness et al., 2016), suggesting that these two domains are statistically dependent in real-world stimuli.

Chapter II investigates whether beta power entrainment activity is disrupted by an unexpected pitch in an isochronous sequence (i.e., no temporal deviations) of tones of one pitch. If beta power entrainment activity only reflects temporal prediction, regardless of the spectral content, beta power should not be affected by an unexpected pitch if it is presented at the expected time of the next auditory event. In contrast, if an unexpected pitch affects beta power entrainment activity, then beta power entrainment reflects spectral as well as temporal prediction. The results showed that beta power does reflect both temporal expectations and violations of spectral expectations.

Chapter III investigates whether beta power entrainment activity is associated with pitch prediction by examining changes in beta entrainment just prior to a predicted pitch change. The results of Chapter II cannot satisfactorily show whether beta power entrainment activity reflects spectral prediction because the neural modulation was measured following the pitch change. Therefore, in the study of Chapter III, isochronous auditory tone sequences with infrequent pitch changes were created in which the pitch changes were either at predictable positions (every 5th tone) or at unpredictable (random) positions (but with the same 20% rate). If beta power entrainment reflects prediction for pitch, predictable pitch changes should be preceded by changes in beta power modulation that reflect the prediction. This is what was found,

providing strong evidence that beta power entrainment reflects both temporal and spectral prediction.

Chapter IV studies the different roles of delta phase, beta power, and their coupling entrainment activities on perception. Critically, the studies of Chapters II and III could not directly reveal the perceptual functions of these oscillations because behavioural outcomes were not measured. In Chapter IV, participants were required to perform a pitch discrimination task with target tones embedded in either rhythmic (isochronous) or arrhythmic (non-isochronous) tone sequences. In order to investigate how delta phase, beta power and their coupling proactively facilitate perceptual performance, each of these were measured prior to target tones and correlated with behavioural perceptual performance (pitch discrimination sensitivity and reaction time) on a trial-by-trial basis. The results suggested different roles, with delta relating to temporal regularly, and beta also to pitch prediction and attention.

Chapter II

Unpredicted pitch modulates beta oscillatory power during rhythmic entrainment to a tone sequence

Chang, A., Bosnyak, D.J., Trainor, L.J. (2016). Unpredicted pitch modulates beta oscillatory power during rhythmic entrainment to a tone sequence. *Frontiers in Psychology*, 7:327. doi: 10.3389/fpsyg.2016.00327

II.1 Preface

Fluctuations in power of beta band (15–25 Hz) oscillations in auditory cortex are involved in predictive timing during rhythmic entrainment, but whether such fluctuations are affected by prediction in the spectral (frequency/pitch) domain was unclear from previous research. In Chapter II, university undergraduates were recruited to passively listen to isochronous auditory tone sequence with unpredictable infrequent pitch changes while EEG was recorded.

The induced beta power was affected by unpredictable infrequent pitch changes, and this response was larger when the pitch change was less predictable. This study showed, therefore, that beta power rhythmic entrainment activity not only reflects temporal prediction, but can also be modulated by unpredicted spectral (frequency/pitch) information.

II.2 Abstract

Extracting temporal regularities in external stimuli in order to predict upcoming events is an essential aspect of perception. Fluctuations in induced power of beta band (15–25 Hz) oscillations in auditory cortex are involved in predictive timing during rhythmic entrainment, but whether such fluctuations are affected by prediction in the spectral (frequency/pitch) domain remains unclear. We tested whether unpredicted (i.e., unexpected) pitches in a rhythmic tone sequence modulate beta band activity by recording EEG while participants passively listened to isochronous auditory oddball sequences with occasional unpredicted deviant pitches at two different presentation rates. The results showed that the power in low-beta (15–20 Hz) was larger around 200–300 ms following deviant tones compared to standard tones, and this effect was larger when the deviant tones were less predicted. Our results suggest that the induced beta power activities in auditory cortex are consistent with a role in sensory prediction of both "when" (timing) upcoming sounds will occur as well as the prediction precision error of "what" (spectral content in this case). We suggest, further, that both timing and content predictions may co-modulate beta oscillations via attention. These findings extend earlier work on neural oscillations by investigating the functional significance of beta oscillations for sensory prediction. The findings help elucidate the functional significance of beta oscillations in perception.

II.3 Introduction

Perceptual systems extract regularities from the stream of continuous sensory input, and form internal representations for predicting future events. Predictive timing is the sensory prediction (or expectation) of when an event will

occur (Nobre et al., 2007; Schroeder and Lakatos, 2009). Such predictions are hypothesized to be essential for many human behaviors, including understanding speech and music (Ding et al., 2015; Doelling and Poeppel, 2015), and synchronizing movements (Jenkinson and Brown, 2011; Fujioka et al., 2012, 2015; Kilavik et al., 2013). Predictive timing can be studied at a basic level in that an isochronous stream of metronome clicks sets up a strong prediction for when the next click will occur.

Entrainment is the process of internal neural oscillations becoming synchronized with temporal regularities in an external auditory rhythmic input stream, and it provides a mechanism for predicting future events in time (Jones, 2010). Such entrainment appears to be accomplished in the brain by neural oscillatory activity, which has been shown to represent temporal regularities in the sensory input, as well as the prediction of upcoming sensory events (Friston, 2005; Jones, 2010; Arnal and Giraud, 2012; Fujioka et al., 2012, 2015; Henry and Herrmann, 2014; Morillon and Schroeder, 2015; Herrmann et al., 2016). While time domain event-related potential (ERP) analyses of electroencephalogram (EEG) waveforms in response to unpredicted stimuli have revealed aspects of neural processes underlying sensory prediction (e.g., Costa-Faidella et al., 2011; Schwartze and Kotz, 2013; Schröger et al., 2015), recent studies indicate that neural oscillatory activities obtained by decomposing EEG signals into frequency-specific bands reveal processes of communication between neural ensembles (Buzsaki, 2006) that are essential to sensory prediction (Arnal and Giraud, 2012).

Oscillatory activities in sensory cortices in both delta (1–3 Hz) and beta (15–25 Hz) bands are associated with temporal prediction (Henry and Herrmann, 2014). The phase of the delta oscillation shows entrainment to rhythmic sequences and it is reset by the onset of a stimulus and predicted (imagined) onset of a future stimulus. On this basis, it has been suggested that delta phase reflects an oscillatory time frame for parsing a continuous sensory stream into meaningful chunks for subsequent perceptual processing (Schroeder and Lakatos, 2009; Calderone et al., 2014). Neural responses to sensory inputs that occur at the time of the excitation phase of delta oscillations are enhanced compared to those that coincide with the inhibition phase (Schroeder and Lakatos, 2009). Local field potential recordings in primary visual and auditory cortices of macaque monkeys show that the delta phase entrains to the onsets of stimuli in rhythmic stimulus streams (Lakatos et al., 2008, 2013), consistent with

intracranial electrocortical and surface EEG recordings in humans (Besle et al., 2011; Gomez-Ramirez et al., 2011; Henry and Obleser, 2012; Herrmann et al., 2016), and it can be endogenously directed by selectively attending to one or the other of two simultaneous stimulus streams (Lakatos et al., 2008, 2013; Calderone et al., 2014).

The amplitude fluctuation dynamics of induced (non-phase-locked) beta band power also entrain to the tempo of events in an auditory input stream, as well as reflecting temporal prediction. EEG and MEG recordings of isochronous auditory sequences show that induced beta power decreases following each tone onset, and increases again prior to the onset time of the next tone, with the timing of the increase varying with tempo in a predictive manner (Snyder and Large, 2005; Fujioka et al., 2009, 2012, 2015; Iversen et al., 2009; Cirelli et al., 2014; Figure 1). Both delta phase angle and beta power in auditory and motor areas in the pre-stimulus onset period predict the accuracy of detecting a temporal delay in the stimulus (Arnal et al., 2015). Furthermore, in primary motor cortex, beta power is modulated by attention, and aligned with the delta phase, suggesting that beta power might reflect attentional fluctuation in time and delta phase an entrained internal clock that aids in the execution of a motor task (Saleh et al., 2010).

Although delta phase and induced beta power are both associated with temporal prediction, compared to the compelling evidence for delta oscillations, the functional significance of beta oscillations in perceptual processing remains less clear. We hypothesized that the entrainment of induced beta power in auditory cortex to an external stimulus might reflect more than predictive timing. Specifically, given that auditory cortex is sensitive to both spectral and temporal dimensions of the input (Fritz et al., 2003; Griffiths and Warren, 2004; King and Nelken, 2009), and auditory evoked ERP components can be interactively modulated by predictions of both pitch and time (Costa-Faidella et al., 2011), beta oscillations might also reflect predictive coding for specific content, such as pitch. In order to examine this hypothesis, we conducted two experiments in which we presented isochronous auditory oddball sequences containing occasional deviations in pitch at different presentation rates. If the induced beta power only reflects predictive timing, the occasional unpredicted pitch changes should not affect the ongoing beta entrainment behavior, given that the pitch deviants are presented at the predicted rhythmic time points. On the other hand, if the induced beta power is affected by the

unpredicted deviant pitches, it would suggest that beta power is associated with predictive perceptual processing for both what and when. In the case that induced beta power is affected by unpredicted deviant pitches, we examine further whether it is modulated by response to novelty (rare events in the preceding local context) or prediction error (the probability of encountering a deviant pitch under the statistical conditions of the context).

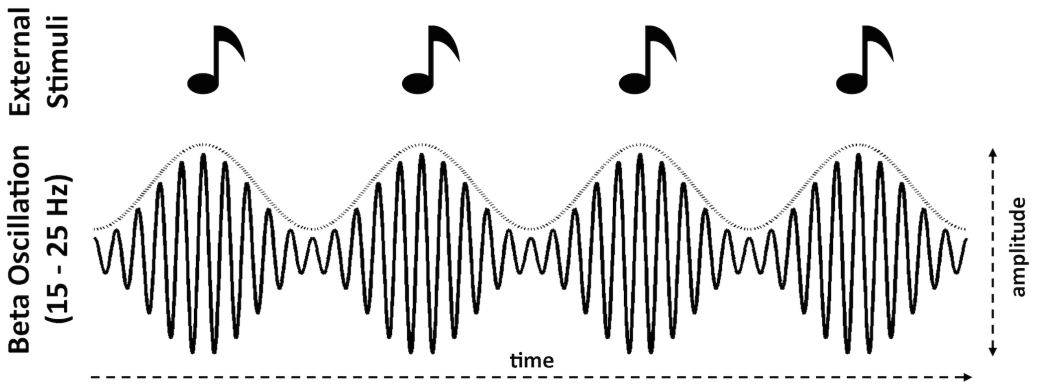

FIGURE II.1: Schematic illustrations of power modulation in induced (non-phase-locked) beta (15–25 Hz) entraining to the tempo of the stimuli. Specifically, power decreases following isochronous onsets and increases that predict the onset time of the next stimulus (e.g., Fujioka et al., 2012; Cirelli et al., 2014). The dotted curve above the beta waveform envelope represents this power modulation.

II.4 Materials and methods

II.4.1 Stimuli

Two recorded piano tones, C4 (262 Hz) and B4 (494 Hz), from the University of Iowa Musical Instrument Samples were used. The amplitude envelopes of the piano tones were percussive with 10 ms rise times. Tones were truncated to be 200 ms in duration, and a linear decay to zero was applied over the entire excerpt to remove offset artifact. The DC shift was removed for each tone. Sounds were converted into a monaural stream at 71 dB (C weighted), measured through an artificial ear (type 4152, Brüel & Kjær) with sound level meter (type 2270, Brüel & Kjær).

II.4.2 Procedure

The experiment was conducted in a sound-attenuated room. Each participant was presented with a continuous sequence of tones in two sessions, each lasting 30 min, while they watched a silent movie on a computer screen. Participants took a 3-min break between sessions. Sounds were delivered binaurally via ear inserts (Etymotic Research ER-2). All stimulus sequences were presented under the control of a digital signal processor (Tucker Davis RP2.1).

The tones were presented in an oddball sequence. The C4 tone was used as the standard and the B4 tone as the deviant. For the first group of participants, the inter-onset interval (IOI) was fixed at 500 ms. There were 3600 tones presented in each session, and the deviance occurrence rate was 10% in one session and 20% in the other session, with an equal number of participants completing the 10% or 20% session first. Within each session, tone order was pseudorandomized with the constraint that two deviant tones could not be presented sequentially, and each session started with five consecutive standard tones. Participants were instructed to sit comfortably and remain as still as possible during the experiment while watching a silent movie. They were not required to make any responses.

In order to replicate and to generalize the findings to a different presentation rate, for a second group of participants, we employed a longer IOI of 610 ms in an isochronous oddball sequence with the 10% deviant tones condition. Otherwise, the procedure for group two was the same as that for group one.

For convenience, we refer to the 500 ms IOI experimental sessions (10% and 20% deviance occurrence rates) as the Fast Experiment, and the 610 ms IOI experimental session (10% deviance occurrence rate only) as the Slow Experiment.

II.4.3 Participants

Sixteen participants (17–22 years old, mean age 18.93 ± 1.39; 12 female) for the Fast Experiment and a different thirteen participants (17–21 years old, mean age 18.62 ± 1.33, 10 female) for the Slow Experiment were recruited from the McMaster University community. Participants were screened by a self-report survey to ensure they had normal hearing, were neurologically healthy and were

right-handed. Signed informed consent was obtained from each participant. The McMaster University Research Ethics Board approved all procedures. Participants received course credit or reimbursement for completing the study.

II.4.4 Electroencephalographic Recording

The EEG was sampled at 2048 Hz (filtered DC to 417 Hz) using a 128-channel Biosemi Active Two amplifier (Biosemi B.V., Amsterdam). The electrode array was digitized for each participant (Polhemus Fastrak) prior to recording. EEG data were stored as continuous data files referenced to the vertex electrode.

II.4.5 Signal Processing of the EEG Data

Three stages of signal processing were conducted in order to examine the behavior of auditory evoked and induced oscillations in bilateral auditory cortices. In the first stage, we obtained a dipole source model based on auditory evoked responses, following Fujioka et al. (2012). The second stage segmented and categorized the source waveform into epochs based on the relative order of the presented auditory sequence. In the third stage, epochs containing excessive artifacts were rejected.

Stage 1: Dipole Source Modeling

The continuous EEG data was band-pass filtered 0.3–100 Hz for each participant for each session, and then segmented into epochs covering the time period -100 to 300 ms, time locked to stimulus onset. Epochs containing standard tones that preceded and followed other standard tones with amplitudes exceeding 150 μV were rejected as artifacts. The surviving standard epochs (89.6% $\pm$ 5.1% for 10% session and 89.5% $\pm$ 5.1% for 20% session of Fast Experiment, and 88.4% $\pm$ 5.5% of Slow Experiment) were averaged into ERP waveforms and band pass filtered between 1 and 20 Hz (Figure 2). To confirm that our oddball context was set up appropriately, a similar procedure was performed on the deviant epochs, and the average of the standard epochs subtracted from the average of the deviant epochs in order to produce difference waves. As can be seen in Figure 2, both mismatch negativity (MMN)

and P3a responses can be observed, consistent with the literature on ERP responses in oddball contexts (Friedman et al., 2001). Paired t-tests, performed on the average of channels in the mid-frontal area (F1, Fz, F2, FC1, FCz, and FC2), confirmed the presence of an MMN component between 100 and 120 ms; specifically, deviant trials were significantly more negative than standard trials in this time window in all sessions of both Fast and Slow Experiments ($ps < 0.001$). There was also a P3a component between 200 and 220 ms: deviant trials were significantly more positive than standard trials in this time window in all sessions of both Fast and Slow Experiments ($ps < 0.001$). It is worth noting that although the latencies of MMN and P3a observed in the current study were earlier than are sometimes reported (e.g., MMN: 150–250 ms, P3a: 250–300 ms; Friedman et al., 2001; Näätänen et al., 2007; Polich, 2007), our results are consistent with several previous studies showing that the latencies of MMN and P3a are as short as around 100 and 200 ms, respectively, when the stimuli are presented in a rhythmic context with IOIs less than or equal to 700 ms (e.g., Regnault et al., 2001; Jongsma et al., 2004; Pablos Martin et al., 2007; Matsuda et al., 2013).

We employed a dipole source model as a spatial filter for increasing the signal-to-noise ratio of the EEG signal generated from left and right auditory cortices for subsequent analyses. A previous study showed that beta activities generated in both auditory and motor cortices entrained to external auditory rhythms when participants passively listened to isochronous sequence of tones (Fujioka et al., 2012). In the present study, we were primarily interested in responses from auditory areas, so we analyzed the EEG signals in source space rather than from surface channels, to extract the oscillatory signals generated from auditory cortex while attenuating signals generated from other brain regions. The source modeling was performed on each participant's mean standard ERP waveform using the multiple source probe scan algorithm and the four-shell ellipsoid model included in the Brain Electrical Source Analysis (BESA) software package. Two auditory cortex sources were estimated for each participant for the auditory evoked P1 (60–100 ms; Figure 2) with the dipoles constrained to be symmetric across hemispheres in location but not orientation. P1 was chosen because it is the dominant peak at fast presentation rates (N1 peaks are strongly reduced at fast rates; Näätänen and Picton, 1987), and is generated primarily from primary auditory cortex (Godey et al., 2001). The mean locations of fitted dipoles across participants were at Talairach coordinates -45.0, -3.2, 16.2 with orientation (0.2, 0.6, 0.8) and 45.0,

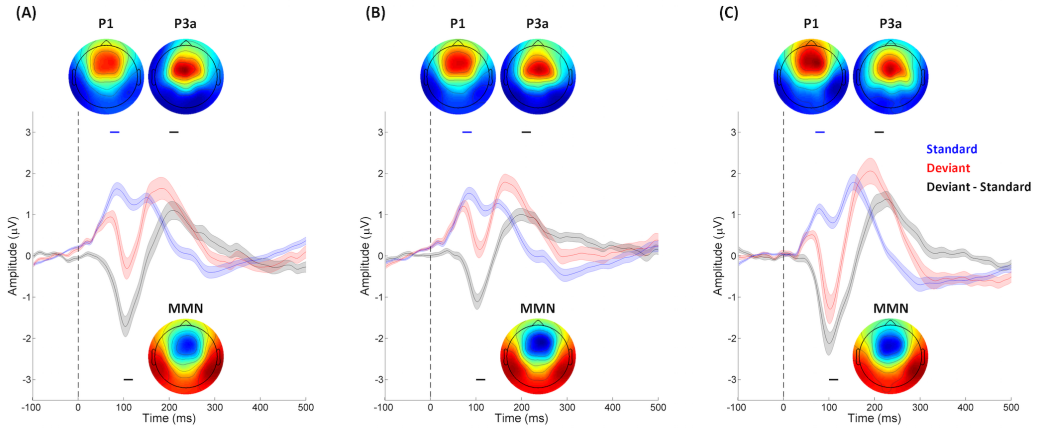

FIGURE II.2: Auditory evoked event-related potential (ERP) waveforms of mid-frontal electrodes from the (A) 10% session and (B) 20% session of the Fast Experiment, and (C) Slow Experiment. Waveforms were collected using 128 EEG channels, and averaged across channels located at the mid-frontal area (F1, Fz, F2, FC1, FCz, and FC2), with stimulus-onset at 0 ms (indicated by the vertical dashed line in each plot). The shaded areas indicate the SEMs of standard trial (blue), deviant trial (red), and the difference waveform of deviant minus standard trial (black). The ERP waveforms of standard trials show a prominent P1 component around 70–90 ms (indicated by the blue line above each waveform). P1 topography of each session (inset; red represents positive potential, blue negative) shows a mid-frontal focus, consistent with generators in primary auditory cortex. The ERP difference waveforms show significant MMN (100–120 ms, indicated by the black line below each waveform) and P3a (200–220 ms, indicated by the black line above each waveform) components. The topography of the MMN (inset) shows the typical frontal negativity of the MMN. The P3a is larger in deviant than standard trials (inset), with typical topography showing a frontal positivity.

-3.2, 16.2 with orientation (-0.1, 0.7, 0.7) in the 10% session of the Fast Experiment; and at -45.4, -3.1, 17.2 with orientation (0.3, 0.7, 0.7) and 45.4, -3.1, 17.2 with orientation (-0.1, 0.8, 0.6) in the 20% session of the Fast Experiment; and -44.9, -4.7, 16.4 with orientation (0.1, 0.7, 0.7) and 44.9, -4.7, 16.4 with orientation (-0.2, 0.7, 0.7) in the Slow Experiment, which are all closely located at bilateral primary auditory cortices with orientations toward the mid-frontal surface area (Figure 3). The residual variances of the source fittings for each session for each participant were between 5% and 10%.

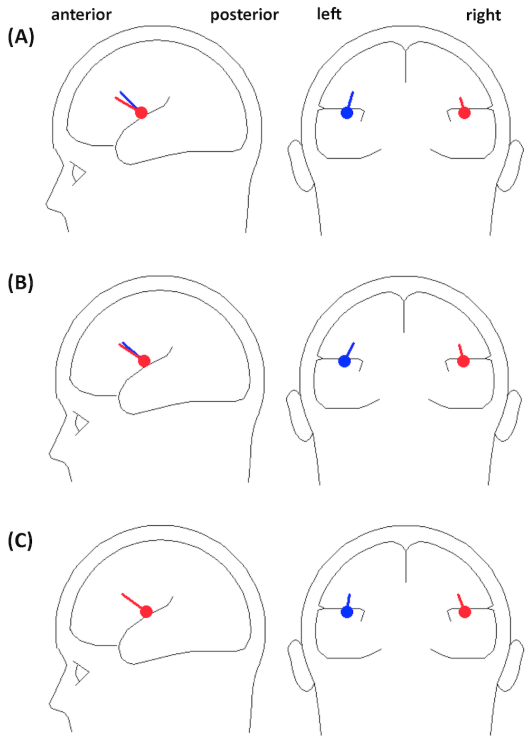

FIGURE II.3: The mean locations and orientations of dipoles. Dipole locations were symmetrically fitted for the auditory P1 ERP component across participants for each (A) 10% session and (B) 20% session of the Fast Experiment, and (C) Slow Experiment, presented in both sagittal and coronal planes. The fitted dipoles are closely located at bilateral primary auditory cortices with orientations toward frontal midline.

Stage 2: Epoching

Based on individual participant dipole model fits for each session, the source activities of single trials in auditory cortices were extracted for all epoch types using signal space projection following Fujioka et al. (2012). Because we were interested in the inter-stimulus neural responses, and to avoid edge effects in subsequent time-frequency analysis, the unfiltered EEG data of each session were segmented into relatively long -500 to 1000 ms epochs, where 0 ms represents a stimulus onset. The epochs were categorized based on the relative position of tones presented in the experiment, including standard (standard

tones between two standard tones), deviant (deviant tones between two standard tones) and SpreD (standard tones preceding a deviant tone and following a standard tone). The individual source waveform epochs as well as raw channel EEG data were exported from BESA to MATLAB for further processing.

Stage 3: Artifact Rejection

Another artifact rejection procedure was applied to the raw 128-channel data. Epochs identified to have artifacts were noted, and the corresponding source waveform epochs were eliminated from further analysis. Thus we made sure the source waveform epochs entered into the time-frequency analysis in the next stage were artifact-reduced and unfiltered, to maximize the signal-to-noise ratio. Because we aimed to reject epochs containing EOG or EMG responses, each raw channel EEG epoch was filtered by a third-order Butterworth band pass filter (1–60 Hz). The filtered EEG epochs that exceeded a threshold (40 μV, compared to the baseline mean voltage of -100–0 ms) for more than 10% of the epoch at any channel were excluded from further analysis. An additional seven participants' data were not included in the current data set because more than 50% of their epochs did not pass the criteria at this stage. For the remaining participants 66.18% $\pm$ 8.68% of the epochs in the Fast Experiment and 71.57% $\pm$ 10.54% in the Slow Experiment were accepted for further analysis.

II.4.6 Time-Frequency Decompositions

Time-frequency decompositions were calculated for each participant on each single-epoch source waveform in left and right auditory cortices and for each stimulus condition using a Morlet wavelet transform (Bertrand et al., 1994) for beta frequency band.

In order to remove the evoked (phase-locked) responses from the epoch and thereby obtain the induced (non-phase-locked) responses for subsequent analyses on beta band, we averaged the source waveform for each trial type (evoked response estimate), and then subtracted it from each source waveform epoch (Kalcher and Pfurtscheller, 1995; Fujioka et al., 2012).

The Morlet wavelet transformation was calculated for each time point for each induced epoch with 32 logarithmically spaced frequency bins between 15 and 25 Hz. The wavelet was designed such that the half-maximum width was equal to 3.25 periods of the lowest frequency while the width was equal to 3.56 periods of the highest frequency, linearly interpolated for each frequency bin in between. Subsequently, 300 ms at the beginning and ending of the epoch were eliminated to avoid edge effects. The induced oscillatory mean signal power was calculated by averaging the magnitude of each time-frequency point of wavelet coefficients across trials. Normalizing this to the mean value of the standard epochs across the whole epoch for each frequency resulted in relative signal power changes expressed as a percentage (Fujioka et al., 2012), and all types of epochs within the same session were compared to the same baseline (mean power in the averaged standard epoch between 0 and 500 ms). The fluctuation in power for each type of epoch at each frequency was visualized as a function of time and frequency in color-coded maps of event-related synchronization and desynchronization (Pfurtscheller and Lopes da Silva, 1999).

II.4.7 Discrete Fourier Transform for Neural Oscillation Entrainment

In order to examine whether the observed neural oscillation activity entrained to the presented stimulus rate, we analyzed the time series of each participant's normalized mean induced beta power (derived as above) via discrete Fourier transforms (DFT). For each participant, we took the -200 to 700 ms epoch for the averaged induced beta power from the wavelet transform, zero-padded to 5 s in order to increase the frequency resolution of the DFT to a bin size 0.2 Hz. For each of the beta power time series, the power spectrums revealed by the DFTs were averaged across participants at each of the left and right auditory cortices.

II.4.8 Data Analysis and Statistics

In order to examine whether the deviant tone affected the beta band induced power (1) we compared the standard and deviant trials for each individual participant for both the 10% and 20% deviance sessions to identify deviant-elicited prediction error responses, and (2) we compared this difference of "standard - deviant" between the 10% and 20% deviance rate sessions to investigate the effect of prediction precision, as deviants in the 10% session are less predicted than those in the 20%. We analyzed the window 0–500 ms for the Fast Experiment and 0–610 ms for the Slow Experiment, time-locked to stimulus onset. The standard and deviant trials of individual participants were then used for random effects analysis.

To assess the statistical differences between the induced beta band powers while controlling for multiple comparisons, we performed cluster-based permutation analyses on the two-dimensional time-frequency maps (Maris and Oostenveld, 2007). First, we used a Wilcoxon signed-rank test, a non-parametric paired difference test, to examine the mean power difference in the beta band between each paired time-frequency sample from 0 to 500 ms for the Fast Experiment or 0–610 ms for the Slow Experiment. Second, we grouped the time-frequency adjacent samples reaching a threshold of $p < 0.05$ into single clusters. Third, we summed the test statistics within each cluster into a cluster-level statistic, which became the observed value. Fourth, to build a permutation distribution, we randomly interchanged the experimental conditions for each participant, repeated the previous three steps 5000 times, and extracted the largest cluster-level statistics for each repetition. The final p-value was calculated by comparing the observed value of each cluster with the permutation distribution.

II.5 Results

We first tested whether the induced beta power entrainment phenomenon reported by Fujioka et al. (2012) was replicated in the standard trials. In the Fast Experiment, the induced power in the beta band of the standard trials showed a clear entrainment to the IOI rate (2.0 Hz). Specifically, the DFT analysis on induced beta band power showed the strongest power at 2.0 Hz for

both the 10% and 20% sessions at both left and right auditory cortices (Figures 4A–D). In the Slow Experiment, the induced power in the beta band of the standard trials showed a clear entrainment to the slower IOI rate ($\tilde{1}$.6 Hz) with the DFT analysis showing the strongest power at 1.6 Hz at both left and right auditory cortices (Figures 4E,F). These results replicate previous studies showing that induced beta band power entrains to the IOI of isochronous stimulus sequences (Fujioka et al., 2009, 2012, 2015; Cirelli et al., 2014).

We then examined whether trial type (deviant vs. standard) and session (deviant rate) modulate the induced beta power, in additional to the entrainment activities. In the Fast Experiment, the cluster-based permutation test identified one significant cluster in the 10% session at right auditory cortex, in which the mean induced power at 16–20 Hz, within the range of low-beta band (15–20 Hz), around 200–300 ms after stimulus onset was larger in the deviant trials than the standard trials ($p = 0.044$; Figure 5A) with a large effect size (rank correlation = 0.67). We did not identify any significant cluster at left auditory cortex. We examined the same contrast for the 20% session. Although we failed to identify any significant cluster at either left or right auditory cortex, the power difference of "deviant–standard" trials peaked around 200–300 ms in the low-beta band at right auditory cortex (Figure 5B), which is consistent with the results of the 10% session. We further compared the power difference of "deviant–standard" trials between the 10% and 20% sessions at the previously identified cluster. The Wilcoxon signed-rank test showed that the power difference was significantly larger in the 10% session than in the 20% session ($p = 0.026$), with a large effect size (rank correlation = 0.56). Taken together, this indicates that the induced power in low-beta band around 200–300 ms after stimulus onset was higher in deviant trials than in standard trials, and that this effect was larger in the 10% session than in the 20% session.

The results of the Slow Experiment replicated the results of the Fast Experiment. A cluster-based permutation test showed only one significant cluster around 200–300 ms after stimulus onset at 15–19 Hz at right auditory cortex ($p = 0.026$; Figure 5C), in which the mean induced power was larger in the deviant trials than the standard trials with a large effect size (rank correlation = 0.79).

To further distinguish whether the deviant-induced responses in low-beta band are associated with prediction error or response to novelty (rare events in the preceding local context), given that both processes can be engaged by

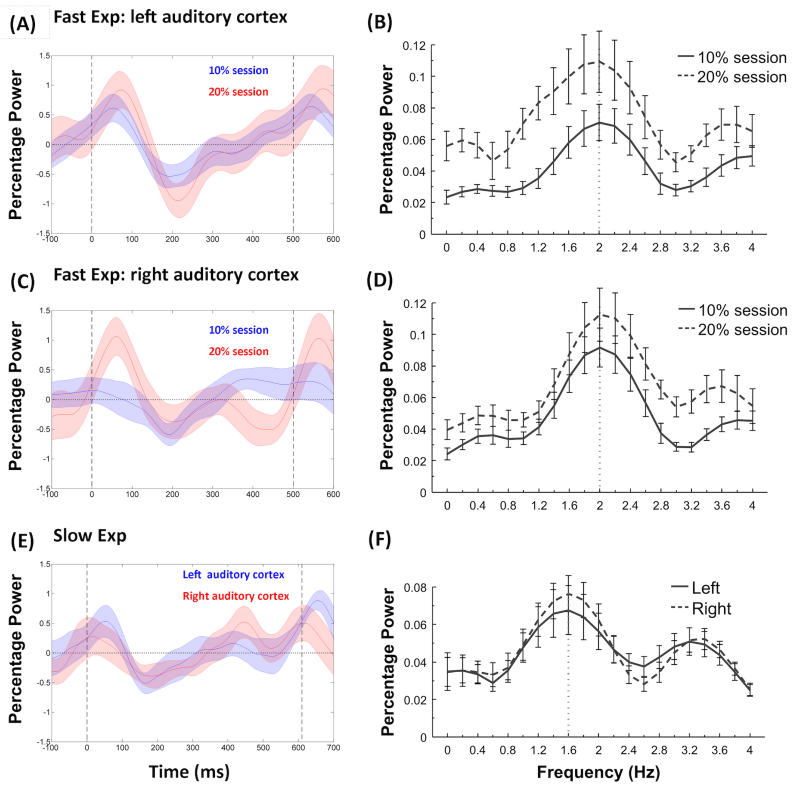

FIGURE II.4: Power fluctuations of induced beta (15–25 Hz) and associated discrete Fourier transformation (DFT) analyses. Fast Experiment: (A) shows the induced beta power fluctuations in the standard trials of the 10% and 20% sessions in left auditory cortex, with shaded areas indicating SEM and vertical dashed lines representing the onsets of tones at 0 and 500 ms. The induced beta power decreases after the onset of a standard tone, and increases (or "rebounds") again before the onset of the next tone. The DFT analyses (B) confirmed entrainment to the stimulus presentation rate (dashed lines) in each case, with maximum power at 2.0 Hz. The same results were replicated at the right auditory cortex (C,D) of the Fast Experiment. Slow Experiment: (E) shows the induced beta power fluctuations in both left and right auditory cortex, with the vertical dashed lines representing the onsets of tones at 0 and 610 ms. The DFT analyses (F)confirmed that the power entrained to the stimulus presentation rate (dotted lines), with maximum power at 1.6 Hz.

deviant stimuli in an oddball context (Friedman et al., 2001), we performed an additional analysis for standard tones occurring in different places in the sequence. This was based on the idea that in an oddball sequence, not only can

the presentation of a deviant tone violate a prediction for a standard tone, but also the presentation of a standard tone that follows several standard tones in a row can violate an expectation (prediction) for a deviant tone. Specifically, the more standards that occur in a row, the more likely it is that a deviant will occur next, given a fixed overall probability of a deviant. On the other hand, a standard occurring after several standards in a row would not elicit a novelty response, as there is no change in the stimulus. If the beta band response that we measured reflects prediction error and not response to novelty, then the response to standard tones should depend on how many standards occurred prior to the standard of interest (as each successive standard builds prediction for an eventual deviant), whereas if the response simply associates with novelty, there should be a larger response to standards in the 20% than 10% condition, but no effect of how many standards occur in a row. Given that a deviant tone must occur eventually along the time line (Luce, 1986; Nobre et al., 2007), the conditional likelihood of encountering a standard tone decreases with the number of repetitions of the standard tone in a row, and thus, on average, the prediction of standard tones preceding a deviant tone will be lower in the 10% than in the 20% session since there are on average more standards in a row before each deviant in the 10% condition.

We can compare responses to standards between 10% and 20% sessions that occur either immediately before a deviant (SpreD) or between two other standards in the sequence (here referred to as SbS). SbS trials occur earlier on average in the sequence compared to SpreD trials. This allows a test of the two alternative hypotheses. Specifically, if the induced low-beta power response at right auditory cortex results from prediction error, the power difference between SpreD trials (20% session–10% session) should be larger than the difference between SbS trials (20% session–10% session), because the prediction error (mismatch between standard and deviant tone) is modulated by conditional likelihood (the position of standard tones in a stimulus sequence). On the other hand, if the induced low-beta power response is modulated by the novelty in the preceding context, the power difference between SpreD trials (20% session–10% session) should be equal to the difference between SbS trials (20% session–10% session), because the conditional likelihood does not matter. Indeed, if anything, the SbS trials would be predicted to show a larger induced low-beta power difference than the SpreD trials because the SbS trials constitute a deviation from a more recently presented deviant tone whereas SpreD trials follow a larger number of standard trials. A cluster-based permutation

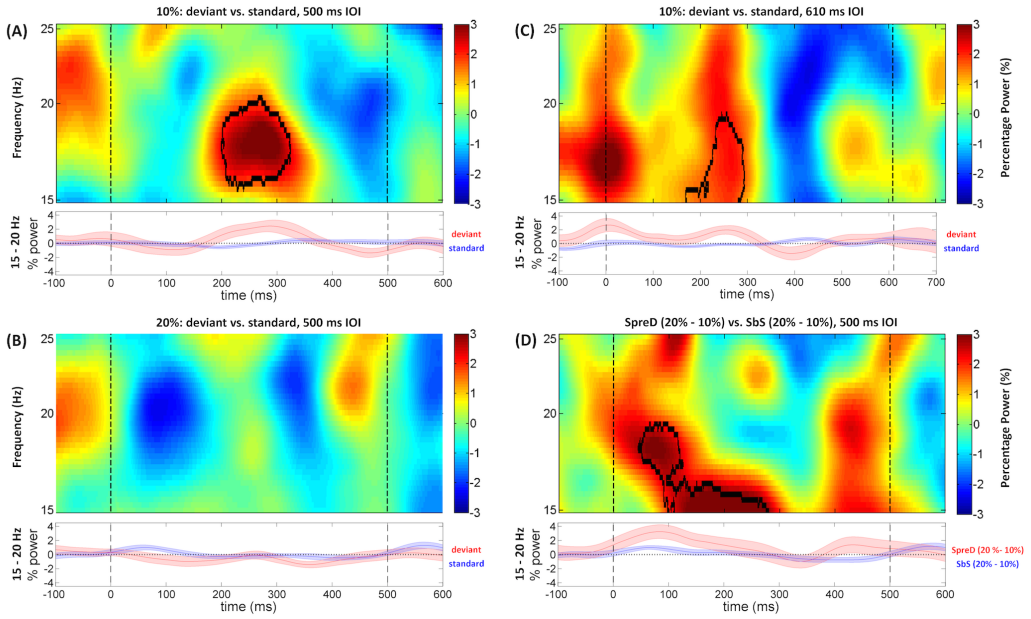

FIGURE II.5: Time-frequency maps of induced power difference between deviant and standard trials in the beta frequency range (15–25 Hz) at right auditory cortex of Fast and Slow Experiments. The shaded areas under each time-frequency map indicates SEM of the low-beta (15–20 Hz) power fluctuations. In the Fast Experiment, standard/deviant tones begin at 0 ms, and the onset time of the next tone is 500 ms (dashed lines). The black contours represent the significant time-frequency cluster. (A)The difference between time-frequency maps (deviant minus standard trials) shows that the deviant tone in the 10% sessions induced stronger power compared to the standard at right auditory cortex, around 16–20 Hz and 200–300 ms. (B) The difference between time-frequency maps (deviant minus standard trials) did not show any significant difference in the 20% session at right auditory cortex. (C) In the Slow Experiment, the standard/deviant tones begin at 0 ms, and the onset time of the next tone is 610 ms (dashed lines). The difference between time-frequency maps (deviant minus standard trials) shows that the deviant induced stronger low-beta (15–19 Hz) power compared to the standard at right auditory cortex, around 200–300 ms. (D) This shows the subtraction of the two difference maps SpreD trials (20% minus 10%) minus SbS trials (20% minus 10%) of Fast Experiment. The result showed that the power difference is larger between SpreD trials than between SbS trials across sessions, around 15–19 Hz and 50–250 ms.

test in low-beta band at right auditory cortex showed that the SpreD trials had a larger induced power difference than the SbS trials ($p = 0.045$; Figure

5D) around 50–250 ms at 15–19 Hz with a large effect size (rank correlation = 0.74). This suggests that the increased induced low-beta power is elicited by prediction error, modulated by conditional likelihood, rather than response to novelty, modulated by rareness of a pitch in the preceding context.

Another additional analysis was performed to investigate whether the current results were associated with the mechanism of auditory stimulus-specific adaptation (SSA) rather than sensory prediction. Auditory SSA refers to the phenomenon that the neural response to the same tone decreases as the number of times it is repeated increases, and raises the possibility that responses to rare tones in an oddball context reflect release from adaptation rather that prediction or response to novelty (e.g., Butler, 1968; Näätänen et al., 1988; Lanting et al., 2013). In the present study, it is possible that the magnitude of the low-beta response to pitch deviants reflects a release from adaptation to the repeated standard tones in our oddball context. Further, the finding that the low-beta power response was stronger on deviant trials in the 10% than 20% session might be due to the fact that there were on average more repeated standard tones preceding a deviant trial in the former case. In order to investigate whether the low-beta response was modulated by a predictive process, we compared conditions where the effect of SSA was constant, but prediction differed. Specifically, to accomplish this, we compared 10% and 20% sessions of the Fast Experiment where the number of standards since the previous deviant was held constant. Thus, we averaged separately deviant effects where there were two standards, three standards, four standards, five standards, or six standards since the last deviant. In each case, we took the low-beta power difference of deviant minus standard trials and compared between the 10% and 20% sessions. The critical point is that, for a given number of standard trials preceding a deviant, the sensory prediction hypothesis indicates that deviants are more expected in the 20% than 10% session because there is a generally higher probability of a deviant in the 20% condition. Specifically, the conditional likelihoods of encountering a deviant tone can be estimated by summing up the empirical occurrence rates of a deviant tone in the all the locations in a sequence following a deviant trial, until the current location (Figure 6A). We performed a cluster-based permutation test on the low-beta band at right auditory cortex. We did not find any cluster to be significant, but there was a trend for the power difference at the cluster at 200 to 300 ms to be larger in the 10% session than in the 20% session (Figure 6B) as predicted by the sensory prediction hypothesis. The fact that it did not reach conventional significance

levels is likely due to the small number of trials (in the 10% session, 141.0 ± 19.0 deviant trials were included in the current analysis, compared to the 244.6 ± 37.8 trials that were included in previous analyses). We compared the maximum deviant minus standard power difference of the averaged low-beta frequency band between 10% and 20% sessions in the time window 130–370 ms for each participant, time-locked to stimulus onset (Figure 6C). The Wilcoxon signed-rank test showed that the maximum low-beta power difference between deviant and standard trials was significantly larger in the 10% session than in the 20% session (2.96 ± 1.09 vs. 0.32 ± 0.45, p = 0.040) with a medium effect size (rank correlation = 0.53). This suggests that the increased induced low-beta power is associated with the degree of prediction error when we controlled the effect of SSA to be the same in both sessions.

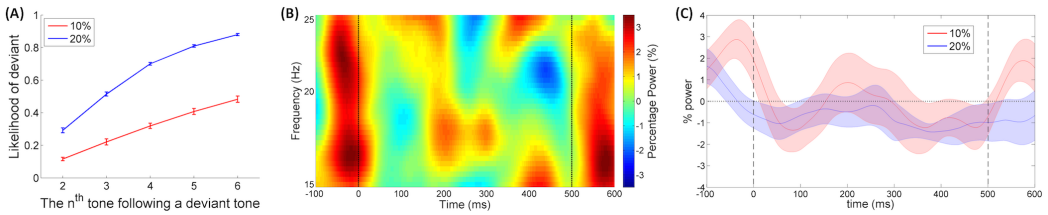

FIGURE II.6: The cumulative conditional likelihoods of encountering a deviant tone, and the time-frequency maps of induced difference (deviant minus standard) responses on matched trial locations in the beta frequency range (15–25 Hz) at right auditory cortex between the 10% and 20% sessions of Fast Experiment. (A) The cumulative conditional likelihoods of encountering a deviant tone as a function of the nth location following a deviant trial in 10% session (red) and 20% session (blue) with error bar indicating SEM. This was calculated by summing up the empirical occurrence rates of deviant tones at the current location and all preceding locations in the experiment. The likelihood of a deviant tone being presented at the nth location is the accumulation of the occurrence rate from the first to nth location following the previous deviant trial. (B) The subtraction of the two difference maps in the 10% session (deviant minus standard) minus the 20% session (deviant minus standard) at the second to the sixth trial following a deviant tone. Although the cluster-based permutation test did not find any cluster to be significantly different, the maximum of low-beta power difference (deviant minus standard) within the 130 to 370 ms window, time-locked to stimulus onset, was significantly larger in the 10% session than in the 20% session. (C) The shaded areas indicate SEM of the averaged low-beta (15–20 Hz) power difference (deviant minus standard) fluctuations of 10% session (red) and 20% session (blue).

In sum, we showed that the deviant tone induced an increase in power in the low-beta band around 200–300 ms following tone onset in right auditory

cortex, regardless of the presentation rate. Also, the effect was stronger when the deviance occurrence rate was lower. Furthermore, two additional analyses suggest that the induced low-beta power was higher for standard tones that violated a stronger prediction for a deviant tone, confirming that the low-beta response is more likely to reflect prediction error than response to novelty. Also, the induced low-beta power response was larger on deviant trials when they were less predictable, even when the effects of SSA were controlled, again suggesting that the low-beta response to deviant tones reflected processes associate with prediction.

II.6 Discussion

We sought to understand the roles of beta oscillations in entrainment to rhythmically predictable sequences by introducing occasional unpredictable pitch deviants. We replicated previous findings related to timing entrainment in induced beta power (Snyder and Large, 2005; Fujioka et al., 2009, 2012, 2015; Iversen et al., 2009; Cirelli et al., 2014), showing that fluctuations in beta power entrained to the rate of presented isochronous auditory stimulus sequences in both left and right auditory cortices. In addition, we found that induced beta band power at right auditory cortex increased around 200–300 ms after the onsets of deviant tones compared to standard tones, especially in the low-beta range (15–20 Hz). This effect was larger when the deviant pitch was less likely to occur (10% vs. 20%), suggesting it is related to prediction processes. The right lateralization of the beta response to pitch deviants is consistent with the idea that the right auditory cortex is more sensitive for processing spectral information than its left counterpart (e.g., Zatorre et al., 1992, 2002). To the best of our knowledge, this is the first study to show that induced beta power in auditory cortex is sensitive to an unpredicted pitch change, even when it is presented at the predicted time. This suggests that induced beta power plays a role in sensory prediction for both what will occur as well as when it will occur.

The increased beta response with decreased likelihood of deviance occurrence indicates that beta oscillations may associate with precision-weighted prediction error. It has been suggested that while prediction error signals do not necessarily involve attention, high precision-weighted prediction errors act

through attention to increase the gain of neural responses, acting as teaching signals for subsequent prediction updating (Friston, 2009; den Ouden et al., 2012; Hohwy, 2012; Schröger et al., 2015). According to predictive coding theory, prediction error is defined as the sensory mismatch between the predicted and perceived stimuli, and precision is the inverse of the input variance of the context which determines whether or not to deploy attention for updating future predictions (den Ouden et al., 2012). For example, prediction precision is higher for standard tones in the 10% than 20% session, because on average there are fewer deviant tone are intermixed in the same length of sequence in 10% than 20% sessions. Thus, larger beta power responses to deviants in the 10% compared to 20% session might indicate that the process involved is one of prediction precision. That beta oscillations are associated with deploying attention for improving perceptual performance is supported by attentional blink studies showing that enhanced phase synchronization in low-beta band among frontal–parietal–temporal regions involved in the attentional network is associated with improved behavioral performance for targets with abrupt onsets (Gross et al., 2004; Kranczioch et al., 2007). Further, it has also been suggested that gamma oscillations (>30 Hz) reflect feed forward prediction error signals (Herrmann et al., 2004) while beta oscillations represent a subsequent feed back processing stage for updating prediction (Arnal and Giraud, 2012), again consistent with the idea presented here that low-beta is sensitive to the precision of prediction, and associates with attention and prediction updating.

The latency of the low-beta response also implies that it is likely associated with attention and prediction updating. The low-beta response to pitch deviants in our data was around 200–300 ms after tone onset, which was later than the well-studied MMN prediction error response in the time waveform ERP, which was around 100 to 120 ms (Figure 2), consistent with other studies employing rhythmic sequences with relatively fast IOIs (Näätänen et al., 2007; Pablos Martin et al., 2007; Fujioka et al., 2008; Matsuda et al., 2013; Hove et al., 2014). This suggests that the low-beta response reflects a processing stage that is later than detecting prediction error. Interestingly, the 200–300 ms timing of the beta band power response occurs around the same time as P3a (Regnault et al., 2001; Jongsma et al., 2004, see Figure 2 for P3a latency), which is known to reflect exogenous attentional orienting and attentional updating (Friedman et al., 2001; Polich, 2007). The P3a and induced low-beta power likely reflect distinct neural responses because the P3a is

phase-locked to stimulus onset and originates in the anterior cingulate cortex and related structures (Polich, 2007) while, in contrast, the induced low-beta power response is not phase locked to stimulus onset and is observed with a spatial filter located in auditory regions. However, the overlapped response latencies are consistent with the idea that attentional processing in frontal areas, reflected by P3a, interacts with prediction precision, and is associated with induced beta power in auditory cortex.

To further evaluate the idea that beta is associated with precision-weighted prediction error, it is important to consider the alternative possibility that the beta band power increases we observed following pitch deviants are simply a response to novelty in the preceding local context rather than prediction error. Indeed, a number of studies in humans and other animals have shown effects of rare stimuli on both induced and evoked beta oscillations (Haenschel et al., 2000; Kisley and Cornwell, 2006; Hong et al., 2008; Fujioka et al., 2009; Pearce et al., 2010; Kopell et al., 2011). Our results strongly favor the idea that induced beta power associates with prediction rather than a simple response to rareness for two reasons. First, the induced power fluctuations of beta oscillation entrain to external isochronous tone sequences in the absence of deviants (Fujioka et al., 2012), which suggests that a primary function of induced beta power concerns temporal prediction rather than detecting rare events. Second, our analyses of standard tones showed that induced low-beta power responses were stronger after the onset of standard tones that were less likely to occur (i.e., the last standard tone occurring after an uninterrupted series of sequential standard tones, SpreD trials) than standard tones that were more likely to occur (i.e., standards occurring earlier in a sequence of standards, SbS trials). This confirms that increased induced low-beta power after tone onset reflects a process that is sensitive to the precision of prediction error.

Our results also suggest that the low-beta response is associated with precision weighted prediction error while controlling possible effects of SSA. Previous studies on adaptation show that the neural response decreases to repeated tones, and that an increased response to the presentation of a new (rare) tone in an oddball context could reflect a release from this adaptation (e.g., Butler, 1968; Näätänen et al., 1988; Lanting et al., 2013). By selecting the deviant trials in the 10% and 20% sessions that had a matched number of standard tones preceding them we equated any effects of SSA between sessions. The

results showed that the low-beta response to a deviant tone was larger in the 10% session than in the 20% session even after SSA was equated. Thus, the lower conditional likelihoods of encountering a deviant tone in the 10% than 20% session associate with a larger low-beta response on deviant trials. This analysis suggests that the low-beta response associates with precision-weighted prediction error although there may also have been a smaller effect of stimulus adaptation. Further research is needed on this question (e.g., see Herrmann et al., 2013, 2014, 2015).

A remaining question concerns the relation between prediction of rhythmic timing (Fujioka et al., 2012, 2015) and prediction precision for pitch, given that induced beta power is interactively modulated by both factors. Here we propose that timing and content (when and what) interact through attentional processing. Dynamic attending theory proposes that internal rhythmic entrainment to external temporal regularities is accomplished by a combination of self-sustained neural oscillation and the dynamic allocation of attention in the temporal dimension (Jones and Boltz, 1989; Large and Jones, 1999; Jones, 2010). The self-sustained oscillation acts as a time frame, and adapts its rate and phase to the external auditory rhythm. Attention increases at important time points such as the onset of beats, which is guided by the temporal prediction of the oscillatory time frame, and reflects temporal prediction for upcoming events during rhythmic entrainment. This attentional rhythmic entrainment is characterized as exogenous orienting (Jones et al., 2006; Nobre et al., 2007; Coull and Nobre, 2008; Jones, 2010), which is involuntary and automatic (Rohenkohl et al., 2011; Triviño et al., 2011; Correa et al., 2014). Further, an MEG study has shown that the mathematical model of dynamic attending theory predicts delta power activities generated in auditory cortex (Herrmann et al., 2016), suggesting that rhythmic attending modulates oscillatory activities in auditory cortex. In this way, it is possible that rhythmic beta power fluctuations representing attention to events with temporal regularity increase perceptual processing of the content of the input stream at predictable time points, such as beat onsets. The idea that beta oscillations reflect temporal attention is also consistent with converging evidence that similar processes occur in the motor system, where rhythmic temporal structure also plays a critical role (e.g., Nobre et al., 2007; Coull and Nobre, 2008; Morillon et al., 2015). This is particularly interesting given that an auditory rhythm sets up beta power oscillations not only in auditory cortex, but also in motor areas even though movement is not involved. Thus, beta power oscillations in

response to a rhythmic auditory input have also been interpreted as reflecting communication between auditory and motor system in the cortex (Jenkinson and Brown, 2011; Fujioka et al., 2012, 2015; Kilavik et al., 2013).

A lack of concurrent behavioral measurements to confirm whether induced beta power modulates perceptual sensitivity is a limitation of the current study. Further experiments are needed to examine this directly. However, the evidence to date shows that increased beta power before a stimulus onset reflects enhanced predictive readiness and improves perceptual performance. Studies using an auditory spatial temporal order judgments task (Bernasconi et al., 2011), an auditory temporal delay detection task (Arnal et al., 2015), intensity detection task (Herrmann et al., 2016), pitch distortion detection task during music listening (Doelling and Poeppel, 2015), or an audiovisual temporal integration task (Geerligs and Akyürek, 2012), all show that when the beta band power happened to be larger in the pre-stimulus period, participants made more accurate judgments or had enhanced audiovisual integration compared to when beta power was smaller. Together, the results of these studies are consistent with our speculation that beta oscillations reflect attention (Wróbel, 2000; Buschman and Miller, 2007, 2009).

II.7 Acknowledgments

This work was supported by a grant from the Canadian Institutes of Health Research (MOP 115043 to LT). AC was supported by a graduate student award from the Natural Sciences and Engineering Research Council of Canada CREATE grant in Auditory Cognitive Neuroscience (371324-2009).

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

We thank Dr. Ian C. Bruce for advice on signal processing, Dave Thompson for technical assistance and Alexandra Rice for assisting with data collection.

Chapter III

Beta oscillatory power modulation reflects the predictability of pitch change

Chang, A., Bosnyak, D.J., Trainor, L.J. (2018). Beta oscillatory power modulation reflects the predictability of pitch change. *Cortex*, 106, 248-260. doi: 10.1016/j.cortex.2018.06.008

III.1 Preface

The study reported in Chapter II cannot distinguish whether the beta power modulation following a pitch change reflects a process of pitch prediction per se or a reaction to a pitch change, because the neural modulation was measured following the pitch change. If beta power entrainment reflects prediction for pitch, predictable pitch changes should be preceded by modulated beta power that reflects the prediction. In Chapter III, two isochronous auditory tone sequences with infrequent pitch changes were presented, such that the only difference between them was that the pitch change was predictable (every 5th tone) in one sequence but unpredictable (pseudorandom) in the other sequence. University undergraduates were recruited to passively listen to these auditory sequences while the EEG activities were recorded. Results showed that beta

power entrainment activity prior to a pitch change is modulated by its predictability. Furthermore, trial-by-trial correlations showed that predictive beta power entrainment activity prior to a pitch change correlates with the size of a post-change prediction error neural response (P3a). Thus, this study demonstrates that beta power entrainment reflects the prediction of both when and what (pitch change).

III.2 Abstract

Humans process highly dynamic auditory information in real time, and regularities in stimuli such as speech and music can aid such processing by allowing sensory predictions for upcoming events. Auditory sequences contain information about both the identity of sounds (what) and their timing (when they occur). Temporal prediction in isochronous sequences is reflected in neural oscillatory power modulation in the beta band (~20 Hz). Specifically, power decreases (desynchronization) after tone onset and then increases (resynchronization) to reach a maximum around the expected time of the next tone. The current study investigates whether the predictability of the pitch of a tone (what) is also reflected in beta power modulation. We presented two isochronous auditory oddball sequences, each with 20% of tones at a deviant pitch. In one sequence the deviant tones occurred regularly every fifth tone (predictably), but in the other sequence they occurred pseudorandomly (unpredictably). We recorded the electroencephalogram (EEG) while participants listened passively to these sequences. The results showed that auditory beta power desynchronization was larger prior to a predictable than an unpredictable pitch change. A single-trial correlation analysis using linear mixed-effect (LME) models further showed that the deeper the pre-deviant beta desynchronization depth, the smaller the event-related P3a amplitude following the deviant, and this effect only occurred when the pitch change was predictable. Given that P3a is associated with attentional response to prediction error, larger beta desynchronization depth indicates better prediction of an upcoming deviant pitch. Thus, these findings suggest that beta oscillations reflect predictions for what in additional to when during dynamic auditory information processing.

III.3 Introduction

Humans need to process highly dynamic, fleeting incoming sensory information in real time, including speech and music. Structural regularities (e.g., beat, meter and pitch patterns in music; timing, stress, phonological, syntactic and semantic structure in speech) and the context in which sensory events occur allow predictability, which can greatly simplify the problem. Indeed, there is considerable evidence that sensory systems extract regularities from perceived external events and predictively encode upcoming sensory input, including what (predictive coding) it might be and when (predictive timing) it might happen (Arnal & Giraud, 2012; Friston, 2005; Heilbron & Chait, 2018; Nobre, Correa, & Coull, 2007; Schröger, Marzecová, & SanMiguel, 2015; Winkler, Denham, & Nelken, 2009), which facilitates perceptual processing in real time (Henry & Obleser, 2012; Henry, Herrmann, & Obleser, 2014; Hickok, Farahbod, & Saberi, 2015; Rohenkohl, Cravo, Wyart, & Nobre, 2012). Prediction is also essential for anticipatorily deploying auditory attention (Large & Jones, 1999; Schroeder & Lakatos, 2009) and coordinating motor actions (Chang, Livingstone, Bosnyak, & Trainor, 2017; Kragness & Trainor, 2016, 2018; Trainor, Chang, Cairney, & Li, 2018; Warren, Wise, & Warren, 2005).

Neural oscillatory activities are regarded as essential neural mechanisms for sensory prediction. The power of beta band (~20 Hz) has been shown to reflect predictive timing (Arnal & Giraud, 2012; Cirelli et al., 2014; Fujioka, Ross, & Trainor, 2015; Fujioka, Trainor, Large, & Ross, 2012; Morillon & Baillet, 2017), but whether it reflects predictive coding (the what domain) remains unclear. Previous electroencephalographic (EEG) and magnetoencephalographic (MEG) studies show that when listening to an isochronous tone sequence, induced (non-phase-locked) beta oscillation in bilateral primary auditory cortices desynchronizes following the onset of each tone, resulting in a power decrease, and then resynchronizes prior to the onset of the upcoming tone. This power fluctuation is proposed to reflect temporal entrainment and temporal prediction because the slope of the resynchronization depends on the presentation speed of the isochronous tones (Cirelli et al., 2014; Fujioka et al., 2012), and because beta power modulation is disrupted in non-rhythmic sequences or by occasionally omitted tones (Fujioka et al., 2012, 2009; Snyder & Large, 2005). Outside of a rhythmic context, beta power has also been shown to reflect the predictability of a temporal gap between tone pairs (Todorovic, Schoffelen, Van Ede, Maris, & de Lange, 2015). Furthermore, beyond the traditional view that

beta oscillation reflects cortical communication between sensory (especially, auditory) and motor regions (Fujioka et al., 2015, 2012; Merchant & Bartolo, 2017; Morillon & Baillet, 2017), accumulating evidence shows beta oscillation is associated with perceptual performance (Spitzer & Haegens, 2017), including detecting temporal or intensity deviations (Arnal, Doelling, & Poeppel, 2015; Herrmann, Henry, Haegens, & Obleser, 2016). Together, this suggests beta oscillation is important for forming temporal predictions, and may associate with the fidelity of perceptual processing.

Pitch is a fundamental perceptual feature of sound that is critical for identifying auditory objects and understanding speech and music, which are part of the what domain of auditory perception. Therefore, it is important to understand how auditory prediction for pitch is represented by neural oscillations.

We hypothesized that beta power modulation in auditory cortex corepresents both predictions of spectral frequency (what) and time (when), two essential dimensions of any auditory signal, consistent with the following evidence. First, neurons in primary auditory cortex are selectively tuned to both spectral and temporal aspects of sound (Boemio, Fromm, Braun, & Poeppel, 2005; Fritz, Shamma, Elhilali, & Klein, 2003; King & Nelken, 2009; Lakatos et al., 2013; Schönwiesner and Zatorre, 2009). Second, beta oscillation is associated with communication between auditory and motor areas (Fujioka et al., 2012), and premotor cortex is activated by spectral as well as temporal predictions (Schubotz, 2007). Third, induced beta power recorded from human auditory cortex is associated with updating pitch prediction, as beta power is modulated by prediction error or prediction updating after the onset of an unpredicted sensory event (Chang, Bosnyak, & Trainor, 2016; El Karoui et al., 2015; Sedley et al., 2016). However, these studies did not report whether there is an effect of prediction on beta power prior to the stimulus onset, and thus it is unclear whether beta power reflects anticipatory sensory prediction. The current study aimed to overcome this lack of knowledge by investigating beta power prior to a pitch change.

The P3a component of the event-related potential (ERP) is regarded as an index of prediction error, and thus it can be used to access whether an auditory event is successfully predicted. When participants passively listen to a sequence of repeated tones with infrequent deviants, both mismatch negativity (MMN) and P3a components of the ERP will be evoked by those deviants

(Polich, 2007; Schröger et al., 2015). Although unpredictability usually covaries with rareness in a typical oddball paradigm, these two features can be dissociated and related to MMN and P3a differentially. The P3a amplitude reflects the magnitude of prediction error, as it is elicited by the unpredictability, but not the rareness, of the deviants (Max, Widmann, Schröger, & Sussman, 2015; Sussman, Winkler, & Schröger, 2003). In contrast, the MMN amplitude reflects local rareness rather than unpredictability (e.g., Bekinschtein et al., 2009). Therefore, we hypothesized that, if beta power modulation in auditory cortex reflects sensory prediction, it should attenuate the prediction error response (i.e., P3a amplitude).

The current study investigated whether beta power modulation, which reflects predictive timing, is also modulated by the predictability of an occasional pitch change. We employed two isochronous (temporally predictable) auditory oddball sequences with 20% of the tones being at a different (deviant) pitch from the standard tones. The deviant pitches were presented every fifth tone in one sequence (Predictable sequence), but the deviant and standard pitches were pseudorandomly intermixed in the other sequence (Unpredictable sequence). We hypothesized that beta power modulation prior to deviant pitch onsets would be modulated by the predictability of the pitch change in the sequence. Furthermore, if beta power modulation reflects the predictability of pitch changes, a larger beta power modulation (i.e., desynchronization depth) should occur prior to the onset of more successfully predicted deviant pitches. Therefore, in single-trial analyses using linear mixed-effect (LME) models, larger beta desynchronization depth immediately prior to a deviant should be associated with smaller deviant-evoked P3a amplitude, a neural response reflecting prediction error (Friston, 2005; Polich, 2007; Schröger et al., 2015).

III.4 Materials and methods

III.4.1 Participants

Seventeen participants were recruited from the McMaster University community, and 14 of them (18 – 27 years old, mean age 19.9 ± 2.4; 10 female)

were used for data analyses (see 2.5.3 Artifact rejection for exclusion criterion). Participants were screened by a self-report survey to ensure they had normal hearing, were neurologically healthy and were right-handed. Signed informed consent was obtained from each participant. The McMaster University Research Ethics Board approved all procedures. Participants received course credit or reimbursement for completing the study.

III.4.2 Stimuli

Two recorded piano tones, C4 (262 Hz) and B4 (494 Hz), from the University of Iowa Musical Instrument Samples were used, with 10 msec rise times. Tones were truncated to 200 msec in duration, and a linear decay to zero was applied over the entire excerpt to remove offset artifact. The DC shift was removed for each tone. Sounds were converted into a monaural stream at 71 dB (C weighted), measured through an artificial ear (type 4152, Brüel & Kjær) with a sound level meter (type 2270, Brüel & Kjær). Sounds were delivered binaurally via ear inserts (Etymotic Research ER-2). All stimulus sequences were presented under the control of a digital signal processor (Tucker Davis RP2.1).

III.4.3 Procedure

The experiment was conducted in a sound-attenuated room. Each participant was presented with a continuous sequence of tones in each of two 30-min experimental sessions, while they watched a silent movie on a computer screen. Participants were instructed to sit comfortably and remain as still as possible during the experiment while watching a silent movie, and they were not required to make any responses, so that we could examine beta responses from auditory cortex without possible artifacts from muscle movement. The tones were presented in oddball sequences, with the C4 tone as the standard and the B4 tone as the deviant. There were two sequence types, Predictable and Unpredictable, such that the only difference between them was whether the change from a standard pitch to a deviant pitch was predictable or not within the sequence. In the Predictable sequence, each deviant pitch was preceded by exactly 4 consecutive standard pitches, and this 5-tone pattern was cyclically repeated throughout the entire sequence, making the pitch of each tone in the

sequence predictable. This manipulation has also been used in previous studies on the predictability of pitch changes (e.g., Dürschmid et al., 2016; Max et al., 2015; Schubotz, von Cramon, & Lohmann, 2003). In the Unpredictable sequence, the order of the standard and deviant pitches was pseudorandomized (with the only constraint being that two deviant pitches could not be presented sequentially), which made the pitch of subsequent tones largely unpredictable (only a standard immediately following a deviant was predictable, and these trials were excluded from analysis). The inter-onset interval (IOI) was fixed at 500 msec in both sequences, and each sequence contained 3600 tones, with a deviant occurrence rate of 20% (720 deviant tones). Although predictive coding has sometimes been studied in sequences with random IOI, we used a fixed IOI here as this reduces variability in beta modulation effects (e.g., Fujioka et al., 2012, 2009; Snyder & Large, 2005). An equal number of participants completed the Predictable or the Unpredictable sequence first. Participants took a three-minute break between experimental sessions.

III.4.4 EEG recording

The EEG was sampled at 2048 Hz (filtered DC to 417 Hz) using a Biosemi Active Two amplifier (Biosemi B.V., Amsterdam). The 128-channel Biosemi pin-type active electrodes (Ag-AgCl) were placed based on geodesic partitioning of the head surface. The electrode array was digitized for each participant (Polhemus Fastrak) prior to recording. EEG data were transformed to an average reference offline.

III.4.5 Signal processing for source-space neural oscillatory activity

Three stages of signal processing were conducted in order to examine the behavior of auditory induced oscillations in bilateral auditory cortices, following Fujioka et al. (2012) and Chang et al. (2016). In the first stage, we obtained a dipole source model based on auditory ERP responses. The second stage segmented source waveforms into trials, categorized by whether they were STD, preDEV or DEV trials (Fig. 1). In the third stage, trials containing excessive artifacts were rejected.

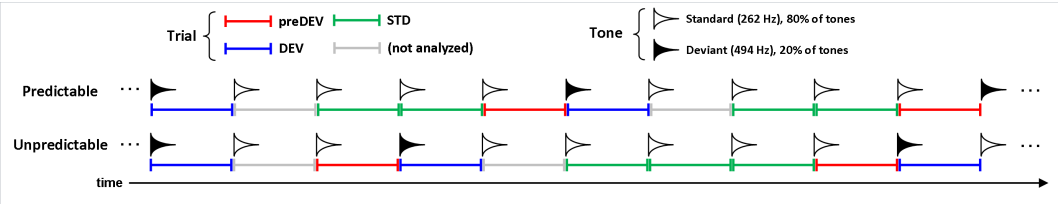

FIGURE III.1: The auditory sequences employed in current study. Two isochronous (IOI: 500 msec) tone sequences in which 80% of tones are at one pitch (standard F0: 262 Hz) and 20% are at a different pitch (deviant F0: 494 Hz). The only difference between sequences was that the deviant pitches were presented regularly every fifth tone in one sequence, making the pitch change predictable (Predictable condition), but the deviant and standard pitches were pseudorandomly intermixed in the other sequence, making the pitch change unpredictable (Unpredictable condition). According to their sequential position relative to the deviant pitch, the trials were categorized into preDEV (red interval), DEV (blue interval), and STD (green interval).

Stage 1: Modeling auditory dipole sources with P1 component

We employed a dipole source model as a spatial filter in order to increase the signal-to-noise ratio of the EEG signal generated from left and right auditory cortices for subsequent analyses. In the present study, we were primarily interested in responses from auditory cortex, so we analyzed the EEG signals in source-space rather than from surface channels, extracting the oscillatory signals generated from left and right auditory cortices while attenuating signals generated from other brain regions (Scherg & Von Cramon, 1985). The continuous EEG data was band-pass filtered 1–20 Hz and then segmented into time periods -100 to 300 msec, time locked to stimulus onset. STD trials on which the amplitude ranged below 150 μV (i.e., low artifact) were averaged into ERP waveforms and used to model dipole sources. Two auditory cortex sources were estimated for each participant for the auditory evoked P1 (60–100 msec; Fig. 3) with the dipole locations constrained to be symmetric across hemispheres in location but not orientation, using the multiple source probe scan algorithm and the four-shell ellipsoid model included in the Brain Electrical Source Analysis (BESA) software package. P1 was chosen because it is the dominant peak at fast presentation rates (N1 peaks are strongly reduced at fast rates; Näätänen & Picton, 1987), and is generated primarily from primary auditory cortex (Godey, Schwartz, De Graaf, Chauvel, & Liegeois-Chauvel, 2001). The mean locations of fitted dipoles across participants were at Talairach coordinates

[±45.8, -7.7, 17.7] with mean orientations [.3, .7, .7] and [-.2, .8, .6] for left and right dipoles, respectively. These locations are very close to bilateral primary auditory cortices with orientations toward the mid-frontal surface area, consistent with typical auditory evoked potentials. The residual variances of the source fittings for each participant ranged from 1.4 to 13.3%.

Stage 2: Trial segmentation

Based on individual participant dipole model fits, the auditory source waveforms of each trial at the left and right auditory cortical sources identified above were extracted for all trial types using signal space projection. Trials were segmented from -500 to 1000 msec, time-locked to the onset of the tone. Both the individual source waveform trials and raw 128-channel EEG data were exported from BESA to MATLAB for further processing.

Stage 3: Artifact rejection

For each segmented source waveform, the corresponding 128-channel data was examined for artifacts. The segments containing artifacts were eliminated from future analyses. The artifact rejection criteria (calculated on surface electrode waveforms) was to eliminate trials in which the amplitude range at any channel exceeded 40 μV for more than 10% of the 1500 msec trial length compared to the baseline mean voltage of the 100 msec prestimulus period for all trials. Three participants' data were excluded because too few trials (36–38%) were artifact-free. For the remaining participants (n = 14), 70.0 ± 13.8% (ranging from 50.6 to 90.9%) of the trials were accepted for further analyses.

III.4.6 Time-frequency decompositions

In order to remove the evoked (phase-locked) responses from the trial and thereby obtain the induced (non-phase-locked) responses for subsequent analyses on beta band, for each trial type, the source waveform was averaged as the estimation of the evoked response, and then subtracted from each single-trial source waveform (Chang et al., 2016; Fujioka et al., 2012;

Kalcher & Pfurtscheller, 1995; Mouraux & Iannetti, 2008; Pfurtscheller, 2001; Pfurtscheller & Da Silva, 1999).

Time-frequency decompositions in the beta frequency band were calculated for each participant on each unfiltered single-trial source waveform separately in left and right auditory cortices using a Morlet wavelet transform (Bertrand, Bohorquez, & Pernier, 1994). The Morlet wavelet transformation was calculated on a window centered around each time point for each trial with 67 logarithmically-spaced frequency bins between 12 and 35 Hz. The wavelet was designed such that the half-maximum width was equal to 2.4 periods of the lowest frequency while the width was equal to 3.4 periods of the highest frequency, linearly interpolated for each frequency bin in between. Subsequently, 300 msec at the beginning and ending of each 1500 msec long trial were eliminated to avoid edge effects, the effectiveness of which was confirmed by visual inspection. The induced oscillatory power was calculated by averaging the power at each time-frequency point across trials. Induced power changes were expressed as a relative percentage, by normalizing the power of a trial to the mean power of the 0–500 msec time window of the STD trials for each frequency (Chang et al., 2016; Fujioka et al., 2012). For the preDEV trials, in order to focus on event-related spectral dynamics in the -500 to 0 msec period (time-locked to the deviant pitch onset), the percentage power was baseline corrected to the mean power of each frequency bin in the -700 to -600 msec period on the preDEV trials (equivalent to the -200 to -100 msec period time-locked to the preceding standard pitch onset).

To quantify the amplitude of desynchronization of the mean beta band (15–25 Hz) power, we measured the size of the negative peak within the interval -350 to -150 msec, time-locked to the onset of the deviant pitch. We chose this time window because it is in the middle of the 500 msec IOI, and we expected the maximum negative peak occur within this time window.

III.4.7 Signal processing for ERPs

For the analyses on trial-averaged ERP waveforms, the continuous EEG data for each electrode was .3–40 Hz band-pass filtered for each participant for each sequence, converted to a standard 81-channel montage in BESA, and then segmented into trials covering the time period -100 to 500 msec, time locked to stimulus onset, for the DEV and STD trials. We averaged across frontal

midline (F1, Fz, and F2) and central midline (C1, Cz, and C2) montaged channels. The ERP waveform was baseline-corrected by subtracting the mean amplitude of the 100 msec prestimulus period.

In the ERP, we focused primarily on analyzing the amplitude of P3a, because previous studies have shown that it reflects the magnitude of prediction error (Max et al., 2015; Polich, 2007; Sussman et al., 2003) or a violation of the global regularity of a tone sequence (Bekinschtein et al., 2009; Faugeras et al., 2012). The P3a component is mainly observed over central midline (Polich, 2007), so P3a was analyzed at this location. In addition, we also analyzed the amplitude of the MMN at frontal midline, because it reflects an automatic preconscious response to occasional deviant pitches in a sequence, such as employed in the present study (Friedman, Cycowicz, & Gaeta, 2001; Näätänen, Paavilainen, Rinne, & Alho, 2007).

In order to accommodate potential individual differences in MMN and P3a latency (Conroy & Polich, 2007; Lieder, Daunizeau, Garrido, Friston, & Stephan, 2013), for each participant for each sequence, we first calculated the peak latency of the MMN and P3a in the difference ERP waveform (mean DEV – mean STD waveforms) in the time ranges 80–130 msec and 160–330 msec, respectively. In the second step, we estimated the amplitude of MMN and P3a by taking the area under the curve of the window of $\pm$10 msec and $\pm$20 msec, centered around the peak latencies of MMN and P3a, respectively. The ERP waveforms are corrected to the mean amplitude of the 100 msec prestimulus period.

The signal processing for extracting single-trial ERP waveforms was the same as the above with the following exceptions. First, the continuous EEG data was .3–7.0 Hz band-pass filtered, as a lowered cutoff frequency for the low-pass filter can increase the signal-to-noise ratio by attenuating the high frequency noise in single-trial ERP analysis (Heinrich et al., 2014; Spencer, 2005). Second, to quantify the single-trial ERP amplitude, the ERP difference waveform was extracted for each DEV trial by subtracting its preceding pre-DEV trial, so as to give a measure of the response to a deviant pitch relative its preceding context of standard pitches.

III.4.8 Single-trial correlation between beta desynchronization and P3a amplitude

To examine the trial-by-trial association between beta desynchronization depth on the preDEV trials and P3a amplitude on the DEV trials, we performed single-trial analyses using a LME model. LME modeling is an extension of linear regression modeling, which accesses the influence of predictors of interest (i.e., fixed effects), while taking into account variances across participants (i.e., random effects). In the current study, we were interested in whether P3a amplitude can be predicted by beta desynchronization amplitude (continuous variable) and sequence type (Predictable or Unpredictable; categorical variable). Both random intercepts and random slopes are included in our LME models, because this setup is known to have the best generalizability of the LME model (Barr, Levy, Scheepers, & Tily, 2013). Model fitting was implemented using the "lme4" package in R (Bates, Mächler, Bolker, & Walker, 2015), and the significance of the fixed effects was determined with type-II Wald tests using the "Anova" function in the "car" package in R (Fox & Weisberg, 2011).

III.4.9 Experimental design and statistical analysis

Participant (n = 14) was the random factor for the within-subject statistical tests in the current study. We employed planned non-parametric tests because the normality assumption was violated in many cases. The approximated test statistics (z-value) is reported for each test. The statistical tests were performed by MATLAB (2013a) or R (3.3.3). Statistical decisions were based on a two-tailed test.

III.5 Results

III.5.1 Predictability of pitch change modulates pre-deviant beta power

The main goal of this study was to test whether induced beta desynchronization is modulated by the predictability of the pitch of the upcoming tone. We analyzed the beta power activity across the interval preceding the onset of each deviant pitch (preDEV trials) in both the Predictable and Unpredictable sequences (Fig. 2), and we hypothesized that the predictability of the deviant pitch would modulate the induced beta power. We quantified the amplitude of desynchronization of the mean beta band (15–25 Hz) power by measuring the size of the negative peak within the interval of -350 to -150 msec, time-locked to the onset of the deviant pitch. In the right auditory cortex, Wilcoxon signed rank tests showed that the desynchronization was only significantly lower than 0 in the Predictable sequence ($z = -3.23$, $p < .001$), but not detectably different from 0 in the Unpredictable sequence ($z = -1.66$, $p = .104$). In the left auditory cortex, the desynchronization was not significantly different from 0 in either sequence (Predictable: $z = -1.73$, $p = .091$; Unpredictable: $z = -1.79$, $p = .079$). A Wilcoxon signed rank test showed that the desynchronization amplitude was more significantly negative in the Predictable sequence than the Unpredictable sequence in the right auditory cortex ($z = -2.17$, $p = .029$). However, there was no difference between Predictable and Unpredictable sequences in the left auditory cortex ($z = -.28$, $p = .808$).

We further explored whether the predictability of the pitch has different effects on low-beta (15–20 Hz) and high-beta (20–25 Hz) bands, as these two sub-bands could have different functions (e.g., Kilavik, Zaepffel, Brovelli, MacKay, & Riehle, 2013). However, we did not observe any differences between these two sub-bands. The detailed results are reported in Supplementary Materials: S.2 and Figure S2.

In sum, beta desynchronization was only observed prior to a deviant pitch in a predictable context but not in an unpredictable context. This is consistent with our hypothesis that predictability for what modulates beta power prior to a pitch change. In addition, the right-lateralization of the effect is consistent

with the view that right auditory cortex is more sensitive for processing spectral information than is the left hemisphere (e.g., Chang et al., 2016; Zatorre et al., 1992, 2002).

III.5.2 Predictability of pitch change modulates deviant ERP amplitudes

We analyzed whether the deviant pitch evoked larger amplitudes of MMN and P3a ERP components in the Unpredictable than in the Predictable sequence (Fig. 3). MMN and P3a responses to occasional deviant tones have been well studied using auditory oddball sequences such as those in the present study (see Friedman et al., 2001; Polich, 2007 for reviews).

We first showed that MMN and P3a responses were observed in both sequences, as expected in a typical auditory oddball paradigm (Näätänen et al., 2007; Polich, 2007). The analyses on the average of frontal midline electrodes using Wilcoxon signed rank tests showed that for both sequence types, the amplitudes of MMN ($z = -3.30$, $p < .001$ in both sequences) and P3a components ($z = 3.30$, $p < .001$ in Predictable sequences; $z = 2.86$, $p = .002$ in Unpredictable sequences), were significantly different from zero in the difference waveforms (DEV minus STD trials). Similar results were found for the same analyses on the average of central midline electrodes: the amplitudes of MMN ($z = -3.30$, $p < .001$ in both sequences) and P3a components ($z = 3.30$, $p < .001$ in both sequences) were significantly different from zero in the difference waves for both Predictable and Unpredictable sequences.

More interesting, we further investigated whether the predictability of the deviants modulated subsequent MMN or P3a. At the frontal midline average, Wilcoxon signed rank tests did not show any differences in either MMN ($z = 1.22$, $p = .241$, Fig. 3B) or P3a (peak-to-peak corrected to mean MMN amplitude, $z = -.09$, $p = .952$, Fig. 3C) between Predictable and Unpredictable sequences. At the central midline average, Wilcoxon signed rank tests did not show a significant difference in MMN amplitude between sequences ($z = 1.16$, $p = .268$, Fig. 3D). On the other hand, the mean P3a amplitude was larger in the Unpredictable sequence than the Predictable sequence ($z = 2.10$, $p = .035$, Fig. 3E; if the P3a Amplitude Difference outlier [-.61; exceeding 2.28 standard deviations from the mean] was excluded, then $z = 2.69$, $p = .005$).

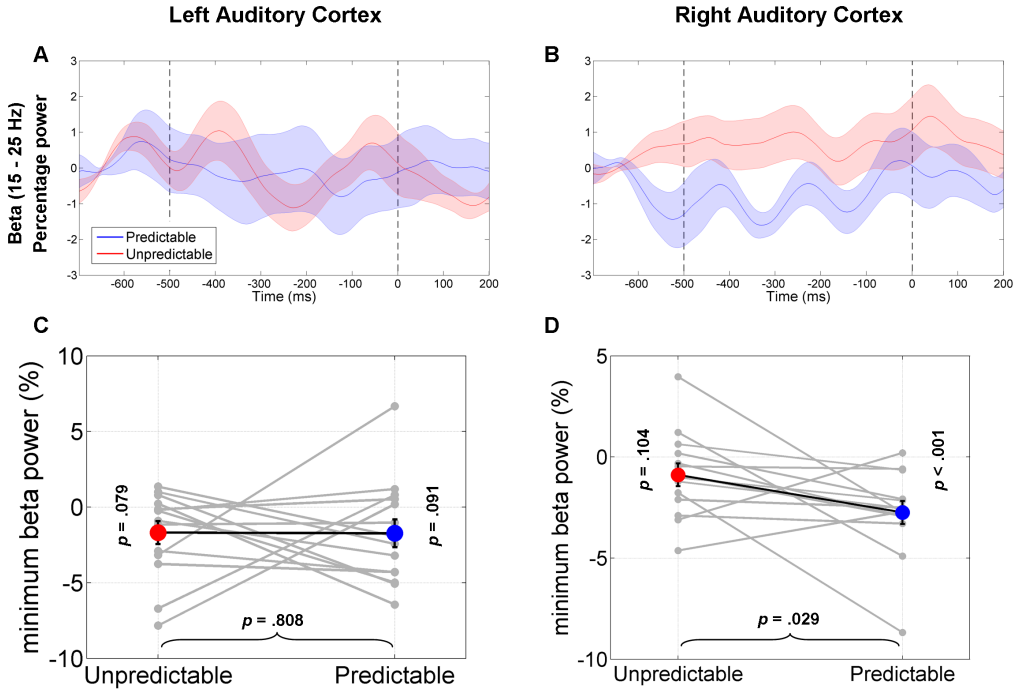

FIGURE III.2: Predictability of pitch change modulates pre-deviant beta power. The induced (non-phase-locked) oscillatory power was extracted from left (left panel) and right (right panel) auditory cortices using the dipole source model. (A and B) Beta (15–25 Hz) power was averaged across trials in each condition for each participant and are presented as the mean ± standard error across participants. The time-frequency maps of beta power on preDEV trials are shown in Figure S2. (C and D) The minimum beta power (desynchronization amplitude) in the -350 to -150 msec time window was extracted for each participant. The grey dots connected by a grey line represent the data points of each participant, and the red and blue dots and the error bars represent the conditional mean ± standard error. Wilcoxon signed rank tests showed that desynchronization was larger in Predictable than Unpredictable sequences in right auditory cortex, but not different in left auditory cortex.

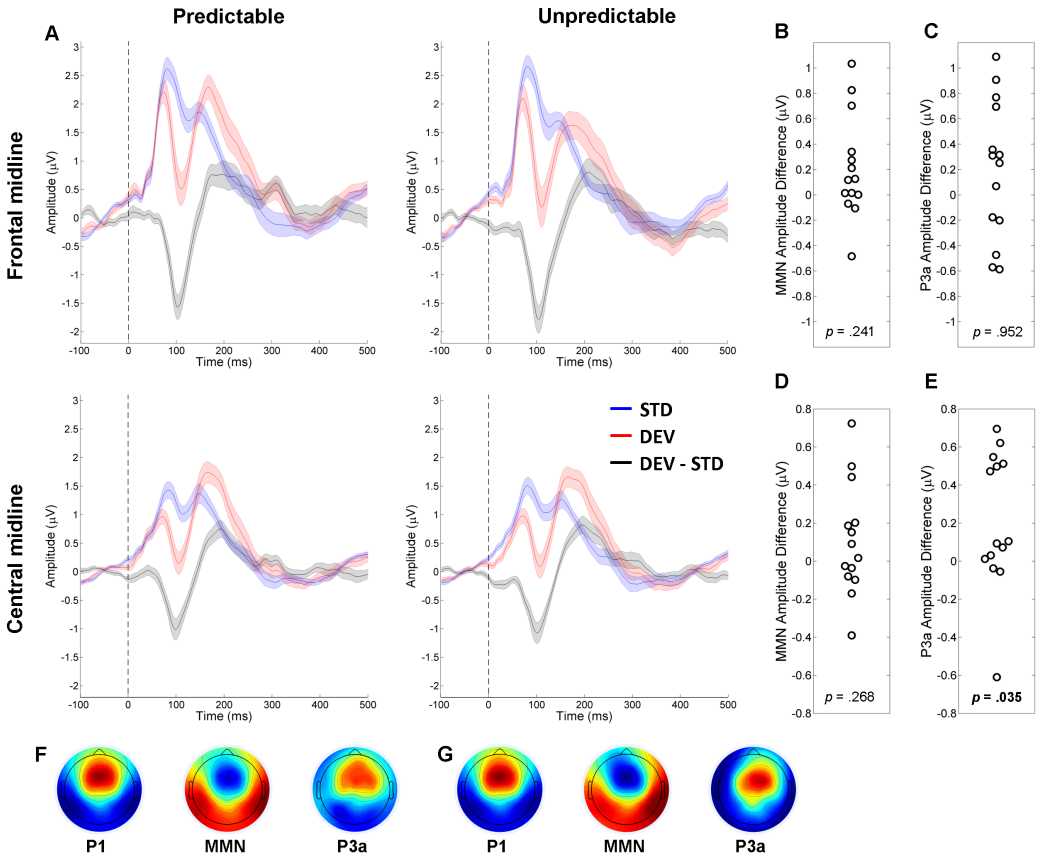

FIGURE III.3: Predictability of pitch change modulates deviant ERP amplitudes. (A) The ERP waveforms of STD (blue) and DEV (red), and the difference waveform (DEV – STD [black]) in Predictable and Unpredictable sequences at frontal midline and central midline electrodes. ERP waveforms are presented as the mean ± standard error across participants. (B, C, D and E) The amplitude differences (Unpredictable sequence minus Predictable sequence) for MMN (left panel) and P3a (right panel), where each circle represents the amplitude difference of one participant at frontal midline electrodes (upper panel) and at central midline electrodes (lower panel). Wilcoxon signed rank tests showed that only the P3a amplitude in central midline electrodes were significantly larger in the Unpredictable sequence than Predictable sequence. (F and G) ERP component topographies (grand-averaged across participants), including P1, MMN and P3a for the (F) Predictable and (G) Unpredictable sequences. The highest amplitude is presented as red, the lowest amplitude is presented as blue, with a symmetrical range around zero (green).

Together, these results suggest that the predictability of a deviant pitch modulates P3a amplitude but not MMN amplitude, such that unpredictable deviant pitches (generating a prediction error) elicit higher attentional capture compared to predictable deviant pitches, replicating previous studies (Max et al., 2015; Polich, 2007; Sussman et al., 2003). The results are also consistent with previous findings that MMN is sensitive to local deviance occurrence rates (e.g., Bekinschtein et al., 2009; Faugeras et al., 2012); in the current study, local deviance occurrence rates were equivalent in the two conditions (20%) and MMN did not differ between conditions.

III.5.3 Pre-deviant beta power is associated with deviant P3a amplitude

The above analyses, by comparing experimental conditions, showed that the predictability of a deviant pitch in a tone sequence context modulates both the pre-deviant (preDEV trial) induced beta power activity in the right auditory cortex and the deviant P3a amplitude at the central midline average. We thus examined whether these two neural activities are correlated at the trial-by-trial level.

We performed a single-trial analysis with linear mixed effect (LME) models of whether the P3a amplitude can be predicted by beta desynchronization amplitude (negative peak), and whether this relationship is modulated by Predictable or Unpredictable sequences (Fig. 4A). The initial LME model showed a significant interaction effect between beta desynchronization amplitude and the type of sequence [standardized β = -.029, standardized SE = .013, $\chi^2(1)$ = 4.56, p = .032]. We thus further analyzed the LME model within each Predictable or Unpredictable sequence. The results showed a significant positive relationship between beta desynchronization amplitude and P3a amplitude for the Predictable sequence [standardized β = .034, standardized SE = .014, $\chi^2(1)$ = 6.13, p = .013]; however, the effect was not significant for the Unpredictable sequence [standardized $\beta = -5.4 * 10^{-4}$, standardized SE = .013, $\chi^2(1)$ = .002, p = .968].

Together, the trial-by-trial analyses showed that deeper beta desynchronization depth (lower beta power amplitude) preceded trials with smaller deviant-evoked P3a amplitude in sequences where pitch changes were predictable (Fig.

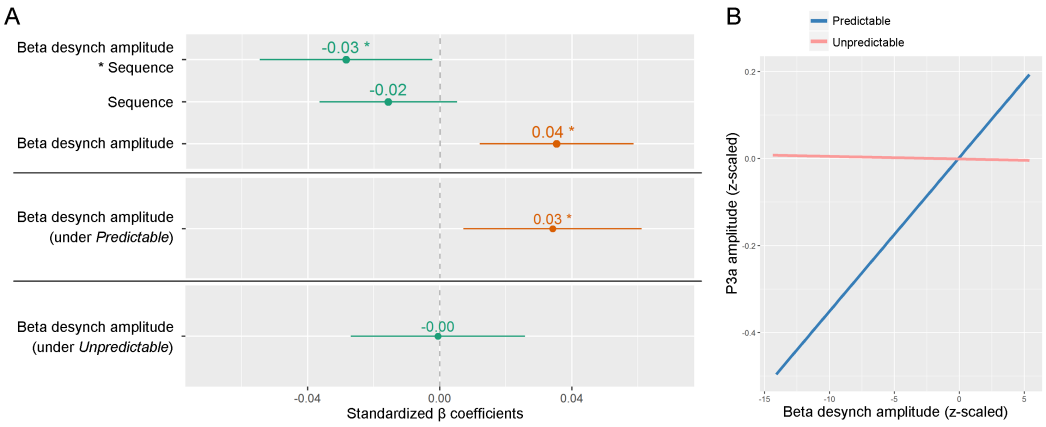

FIGURE III.4: Pre-deviant beta power depth correlates with deviant P3a amplitude when pitch change is predictable. We used LME models to examine the trial-by-trial relationship between beta desynchronization depth (negative peak) on preDEV trials and P3a amplitude on the following DEV trials. The visualization was implemented by the "sjPlot" package in R (Lüdecke, 2016). (A) The best fitted standardized β coefficients (fixed-effects only). In the upper panel, "Sequence" indicates differences in the effect of Unpredictable versus Predictable sequences on P3a amplitude, "Beta desynch amplitude" indicates the effect of beta desynchronization amplitude on P3a amplitude, and "Beta desynch amplitude * Sequence" indicates the interaction effect of these two variables on P3a amplitude. In the middle and lower panels, the effect of beta desynchronization amplitude on P3a amplitude was analyzed under each of Predictable and Unpredictable sequences. The dot and error bar represent the standardized β (fixed-effect coefficient) and 95% confidence interval, with * $p < .05$. (B) The predicted relationship between beta desynchronization amplitude and P3a amplitude, under each of Predictable and Unpredictable sequences, based on the LME model. The x and y axes are z-scaled for visualization. In sum, the individual trial analyses showed that deeper pre-deviant beta desynchronization depth (power decrease) was followed by lower deviant-evoked P3a amplitude in the sequence with predictable pitch changes.

4B). Increased post-processing (P3a amplitude) is indicative of a failure to predict the pitch change. Thus, greater beta desynchronization prior to a pitch deviant appears to enhance the predictability of the upcoming pitch deviant, resulting in a smaller prediction error response following the pitch deviant.

III.6 Discussion

The current study aimed to investigate whether power modulation of induced beta oscillation (15–25 Hz) in auditory cortices reflects spectral in additional to temporal prediction. To manipulate the predictability of pitch changes while controlling temporal predictability, we employed two isochronous auditory oddball sequences, where both sequences were presented at the same presentation speed and the same deviant pitch occurrence rate (20%). The only difference was that the deviant pitch occurred regularly every fifth tone in the predictable sequence but was pseudorandomly intermixed with standard pitches in the unpredictable sequence.

We report the novel finding that the predictability of a pitch change modulates beta power immediately prior to a deviant pitch onset. Specifically, we found that beta desynchronization was deeper prior to the onset of a predictable deviant pitch than an unpredictable one. Furthermore, a trial-by-trial LME analysis showed that within the predictable sequence, the deeper the beta desynchronization prior to a deviant pitch, the smaller the P3a amplitude evoked by that deviant pitch. This result indicates that even when the same stimulus is presented on different trials, deeper beta desynchronization depth prior to the stimulus is associated with better prediction of pitch change. To the best of our knowledge, while previous reports indicate that beta oscillations reflect temporal prediction, this is the first report that auditory beta modulation depth reflects sensory prediction of what is expected next.

We demonstrated that P3a amplitude, but not MMN amplitude, was larger in the Unpredictable than in the Predictable sequence, consistent with previous findings that predictability of a tone sequence modulate deviant-evoked P3a amplitude (Max et al., 2015; Sussman et al., 2003). Studies have shown that P3a reflects involuntary attentional capturing, elicited by unpredicted deviant stimuli, even when task-irrelevant (Friedman et al., 2001; Polich, 2007; Rinne,

Särkkä, Degerman, Schröger, & Alho, 2006; Schröger et al., 2015, 2000; Sussman et al., 2003; Wetzel, Schröger, & Widmann, 2013). For example, ERP studies have shown that unpredictable task-irrelevant pitch changes elicit P3a responses, but that such P3a responses are eliminated when the pitch change is predictable, either cued by the regularity of a sequence (e.g., every fifth tone) or by a visual cue (Max et al., 2015; Sussman et al., 2003). Also, other studies manipulating pitch regularity in tone sequences in global and local hierarchies showed that MMN amplitude responds to locally rare pitch and P3 amplitude responds to violations of global pitch regularity (e.g., Bekinschtein et al., 2009; Dürschmid et al., 2016; Faugeras et al., 2012). It's important to note that these ERP effects might depend on the listening tasks being used in the experiment, as the MMN amplitude can reflect the global pitch regularity (predictability) when participants are required to attend to the sound (e.g., Wacongne et al., 2011), which is different from the passive listening paradigm of the current study. Together, the association between beta desynchronization and P3a amplitude found in the current study is consistent with the interpretation that deeper beta desynchronization reflects more successful prediction of upcoming pitch changes, and thus reduces exogenous drawing of attention toward unpredicted events (Friston, 2005; Polich, 2007; Schröger et al., 2015).

There is a debate as to whether the desynchronization response of beta power modulation reflects predictive or a reactive mechanism (Teki & Kononowicz, 2016). The predictive coding framework hypothesizes that beta power modulation reflects prediction of the upcoming event (Arnal & Giraud, 2012) while the reactive event tagging framework hypothesizes that beta oscillation reflects the memory encoding of a perceived salient event (Hanslmayr & Staudigl, 2014; Teki & Griffiths, 2014, 2016). Previous studies were unable to distinguish these alternative hypotheses, given that they used metrically accented tones that occurred cyclically along the sequence (Fujioka et al., 2015; Iversen, Repp, & Patel, 2009; Snyder & Large, 2005), and the shortest temporal distance between two down-beats was 800 msec, making it even more difficult to distinguish whether the beta desynchronization reflected prediction of an upcoming stimulus or reaction to the preceding stimulus. We solved the cyclical issue by comparing sequences with and without pattern regularity. Under this manipulation, the current study showed that beta desynchronization is anticipatorily associated with pitch change only in the predictable context but not in the unpredictable context, which supports the idea that beta desynchronization reflects predictive coding. However, we do not exclude the possibility

that event tagging processes are represented as another pattern of activity in beta band. Studies to date suggest that beta oscillation can be subdivided into high (>20 Hz) and low (<20 Hz) bands, which might be related to different functions (Kilavik et al., 2013; Spitzer & Haegens, 2017). Specifically, novel auditory or visual stimuli elicit subsequent activity in the low-beta band (e.g., Chang, Ide, Li, Chen, & Li, 2017, Chang et al., 2016; Fujioka, Trainor, Large, & Ross, 2009; Haenschel, Baldeweg, Croft, Whittington, & Gruzelier, 2000), and these effects are consistent with the event tagging framework. Therefore, we suggest that the desynchronization of the entire beta band reflects predictive processes, while the low-beta power increase after an unexpected pitch might relate more specifically to an event-tagging process.

The view that beta oscillation reflects sensory prediction is consistent with predictive coding theory, as more salient beta modulation is associated with attenuated P3a amplitude. A successfully predicted sensory input elicits smaller subsequent attentional responses, which are responsible for propagating error signals in brain networks for further updating sensory prediction (Arnal & Giraud, 2012; Friston, 2005, 2009; Heilbron & Chait, 2018; den Ouden, Kok, & de Lange, 2012; Hohwy, 2012; Kopell, Whittington, & Kramer, 2011; Schröger et al., 2015). Also, beta oscillation has been proposed to reflect the cortical representation of maintenance or reactivation of the perceived sensorimotor or cognitive state (Engel & Fries, 2010; Spitzer & Haegens, 2017), which is the basis for predicting upcoming events. A limitation of the current study is that we did not orthogonally manipulate pitch prediction and prediction precision (inverse variance of prediction), which are dissociable factors in predictive coding theory, and thus we are unable to further distinguish whether beta power modulation reflects one or both of these processes. However, a recent study, which recorded local field potentials from human auditory cortex while participants listened to sequences with varied pitches, found that post-stimulus induced beta power was associated with updating spectral prediction but not prediction precision (Sedley et al., 2016). Although this association was only observed in the post-stimulus period and not in the pre-stimulus period (Sedley et al., 2016), it favors the explanation that the pre-deviant beta power modulation in the current study reflects prediction rather than prediction precision.

It is important to note that the standard and deviant pitches in the current study were 11 semitones apart. Considering the tonotopic organization of

auditory cortex, it is likely that the reported beta oscillations were generated by different neuron ensembles tuned to standard or deviant pitch (cf., Lakatos et al., 2013). Thus, because the pitch change occurred regularly every fifth tone of the Predictable sequence (.4 Hz), an alternative explanation of the beta power modulation effect is that the neuron ensemble tuned to the deviant pitch frequency independently entrained to the rhythmic inter-deviant interval in the Predictable sequence (.4 Hz). In other words, there is a potential confound between timing and pitch expectations, and the beta power modulation might reflect predictive timing of deviant tone onsets rather than, or in addition to, predictive coding of pitch changes. However, we argue that it is unlikely to be the case. Perceptual and sensorimotor evidence shows that rhythm tracking performance becomes imprecise for tempos slower than .5 Hz (see Repp, 2005 for a review), and neural oscillatory evidence shows that neural entrainment activities cannot be robustly observed for tempos slower than 1 Hz (Doelling & Poeppel, 2015). The inter-deviant interval rate (.4 Hz) in the Predictable sequence of the current study was slower than these limits, and thus the deviant tempo alone, if processed in a separate channel and not nested within the 2 Hz tone sequence, is unlikely to be tracked. Therefore, we propose that the neuron ensembles tuned to different pitch frequencies might be interconnected and form a predictive coding network in the auditory cortex, and beta oscillation might play an important role in this (cf., Wang, 2010). We acknowledge that the current evidence is only in favor of the explanation that beta oscillatory power reflects the predictability of a pitch change (the what domain), and the current experimental design is insufficient to completely rule out the alternative explanations, as the predictability of what and when are not fully dissociated. Further electrophysiological evidence is needed to fully test this hypothesis.

The current findings also extend our understanding of the function of beta oscillation in sensory processing. First, predictions for both what and when are essential for processing dynamic auditory information, and beta oscillation likely plays an important role in these processes as it co-represents both prediction types. In particular, prediction is essential for auditory stream segregation and chunking the input into hierarchical meaningful segments (Wacongne et al., 2011; Winkler et al., 2009), which are required for perceiving speech and music (Hickok, 2012; Hickok & Poeppel, 2007; Sridharan, Levitin, Chafe, Berger, & Menon, 2007) among other sounds. For example, speech must be hierarchically segmented into phonemes, words and phrases; music must be

segmented into subphrases and phrases. Second, sensory predictions reflected by beta power modulation might subsequently be associated with improved perception, via top-down controlling of attentional gain (Lee, Whittington, & Kopell, 2013; Morillon & Baillet, 2017). This is consistent with previous auditory studies showing that prestimulus beta power is higher when participants make correct compared to incorrect judgments at detecting temporal deviation (Arnal et al., 2015), detecting intensity deviation in tone sequences (Herrmann et al., 2016), detecting distorted pitch in music (Doelling & Poeppel, 2015), or discriminating frequency and intensity (Kayser, McNair, & Kayser, 2016). Third, the current finding of a relation between beta oscillation and sensory prediction can potentially be generalized to other sensory modalities. It has been proposed that beta oscillation plays a role in updating prediction in micro-scale neural circuits (e.g., Arnal & Giraud, 2012; Wang, 2010), and it is possible that such neural circuit exist in other cortical sensory regions. This idea is supported by previous studies showing that beta oscillation is associated with predicting tactile stimuli (van Ede, Jensen, & Maris, 2010), predicting spatial-temporal properties of the visual stimuli (Heideman, Ede, & Nobre, 2018), and motor planning for music performance (Bianco, Novembre, Keller, Villringer, & Sammler, 2018).

Previous studies indicate that the motor system is involved in predicting and perceiving temporal (e.g., Chen, Penhune, & Zatorre, 2008, 2006; Fujioka et al., 2012; Grahn, 2012; Grahn & Brett, 2007; Grube et al., 2010; Manning & Schutz, 2013; McAuley, Henry, & Tkach, 2012; Morillon, Schroeder, & Wyart, 2014; Phillips-Silver & Trainor, 2005; Schubotz, Friederici, & von Cramon, 2000; Schubotz & von Cramon, 2001; Teki, Grube, Kumar, & Griffiths, 2011, but see; Meijer, te Woerd, & Praamstra, 2016) and spectral (Schubotz, 2007; Schubotz & von Cramon, 2002; Schubotz et al., 2003) aspects of auditory sequences. Our finding that beta oscillation reflects pitch prediction as well as temporal prediction is consistent with the general idea that the sensory prediction process involves interactions between sensory and motor areas (Arnal, 2012; Grahn, 2012; Iversen & Balasubramaniam, 2016), given that beta oscillation has been suggested to reflect network communication between auditory and motor regions (Bartolo & Merchant, 2015; Fujioka et al., 2012; Kilavik et al., 2013; Merchant & Bartolo, 2017; Morillon & Baillet, 2017).

III.7 Conclusions

The current study shows that oscillatory beta power desynchronization is deeper prior to a predictable pitch change than an unpredictable change, indicating that beta power modulation reflects predictability for what is expected to happen as well as when it will happen. Our further trial-by-trial analysis showed that deeper beta power modulation prior to a pitch change is correlated with reduced attention-capturing prediction error responses (P3a amplitude) following the pitch change, demonstrating the link between sensory prediction and subsequent attentional processes. Together with previous studies, beta-frequency oscillations reflect predictive aspects of spectral and temporal processing that are critical for processing dynamic auditory streams, such as speech and music.

III.8 Supplementary Materials

III.8.1 Induced beta power fluctuates at the stimulus presentation rate

We analyzed the beta power on the STD trials in the Predictable sequence, where the same tone was presented isochronously, similarly as in previous studies (Cirelli et al. 2014; Fujioka et al., 2009; 2012; 2015; Iversen et al., 2009). As shown in Figure S1C and D, the beta power desynchonization occurred immediately after the tone onset (0 ms time point), and the power resynchronized prior to the onset time of the next tone at 500 ms point, consistent with previous studies.

We analyzed the time series of each participant's normalized mean induced beta power via short-time Fourier transforms. For each participant, we took the -200 to 700 ms period of the STD trial for the averaged induced beta power (mean power across 15 to 25 Hz) from the wavelet transform, zero-padded to 5 seconds in order to increase the frequency resolution to 0.2 Hz. The amplitude spectrum of the time series of induced beta band power (Figure S1E and F) showed the strongest amplitude at 2.0 Hz (the stimulus repetition rate), and Wilcoxon signed rank tests confirmed this observation that the mean power at 2.0 Hz was significantly larger than the mean power at 1.0, 3.0, 4.0, 5.0 and

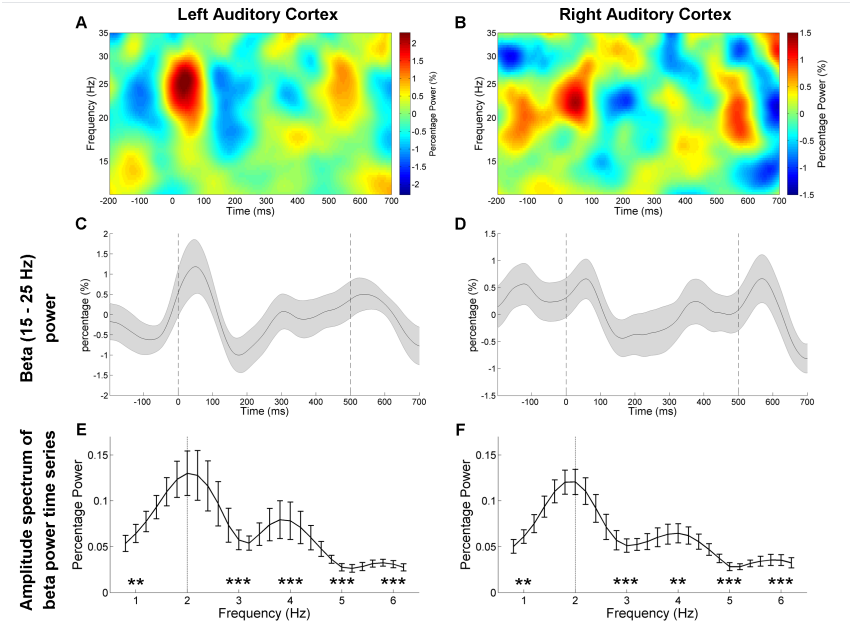

FIGURE III.S1: Beta power entrains to the presentation rhythm of tones. The induced (non-phase-locked) oscillatory power was extracted from left (left panel) and right (right panel) auditory cortices using dipole source modeling. (A and B) The time-frequency maps of induced power on STD trials in the Predictable sequence, grand averaged across participants. The onset time of standard pitches are at 0 and 500 ms. (C and D) Beta frequency (15-25 Hz) power was averaged across trials for each participant and is presented as the mean ± standard error across participants. The onset times of standard tones are at 0 and 500 ms. (E and F) The short-time Fourier transform for each beta power time series was averaged across trials for each participant. The amplitude spectrum is shown as the mean ± standard error across participants. Wilcoxon signed rank tests showed that the power at 2.0 Hz (the presentation frequency of the tones) was significantly greater than at 1.0, 3.0, 4.0, 5.0 and 6.0 Hz, showing the induced beta power entrained to the presentation rhythm of auditory tones. ** $p < 0.01$; *** $p < 0.001$.

6.0 Hz, with all $z > 2.60$ and $p < 0.007$. These results showed that induced beta band power entrains to the presentation rate (2.0 Hz in the present case) of isochronous stimulus sequences.

Although the current study cannot demonstrate beta power entrainment (e.g., Cirelli et al. 2014; Fujioka et al., 2012), due to the lack of variation in stimulus presentation rate and/or use of non-isochronous sequences, the beta power fluctuations observed in the current study are consistent with the

pattern of beta power entrainment found in previous studies.

III.8.2 Predictability of pitch change and pre-deviant low-beta/high-beta power

Given the effect of predictability of pitch change on pre-deviant beta power, we further explored whether the predictability of pitch change modulates low-beta (15-20 Hz) or high-beta (20-25 Hz) power differently (Figure S2), given that these sub-bands might related to different functions (Kilavik et al., 2013; Spitzer and Haegens, 2017). We quantified the amplitude of desynchronization of the mean low-beta or high-beta power in auditory cortex by measuring the size of the negative peak within the interval of -350 to -150 ms, time-locked to the onset of the deviant pitch.

We examined whether the Predictability of pitch change and high/low beta band have an interaction effect (Figure S2E and S2F). For each participant, we took the desynchronization amplitude difference between Predictable and Unpredictable sequences within low-beta and high-beta band. The Wilcoxon signed rank test did not show a significant effect of the desynchronization amplitude difference between the low-beta and high-beta bands in left auditory cortex ($z = -0.91$, $p = 0.391$) or in right auditory cortex ($z = 1.41$, $p = 0.173$).

Together, this exploratory analysis found that the effect of predictability of pitch change did not differ significantly between the high and low beta bands, suggesting low-beta and high-beta oscillatory activities reflect a similar function in current study.

III.9 Acknowledgements

This research was supported by grants to LJT from the Canadian Institutes of Health Research and by a Vanier Canada Graduate Scholarship to AC. We thank Dave Thompson for technical assistance, Alexandra Rice, Michael Ku, and Keeyeon Mark Hwang for assisting with data collection, and the excellent comments of two anonymous reviewers. The authors declare no competing financial interests.

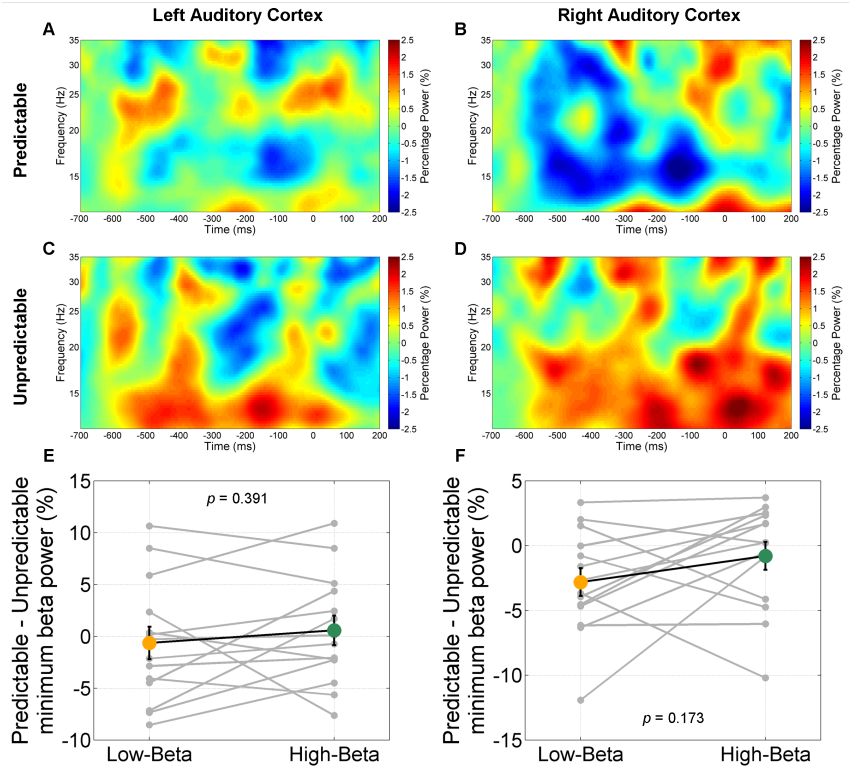

FIGURE III.S2: Pre-deviant beta power activities in Predictable and Unpredictable sequences. (A, B, C and D) The time-frequency maps of induced power on preDEV trials in the Predictable and Unpredictable sequences, grand averaged across participants. The onset time of standard pitch is at -500 ms, and the onset time of deviant pitch is at 0 ms. (E and F) The minimum beta power (desynchronization amplitude) in the -350 to -150 ms time window was extracted for each participant for each sequence, and then the difference (Predictable - Unpredictable) was estimated for each participant. The grey dots connected by a grey line represent the data points of each participant, and the colored dots and the error bars represent the conditional mean ± standard error. Wilcoxon signed rank tests did not show any significant effect of high/low beta band interacting with Predictable/Unpredictable sequence on the desynchronization amplitude.

Chapter IV

Rhythmicity facilitates pitch discrimination: Differential roles of low and high frequency neural oscillations

Chang, A., Bosnyak, D.J., Trainor, L.J. (2019). Rhythmicity facilitates pitch discrimination: Differential roles of low and high frequency neural oscillations. *NeuroImage*, 198, 31-43. doi: 10.1016/j.neuroimage.2019.05.007

IV.1 Preface

The studies of Chapters II and III did not directly measure the perceptual consequences of the neural entrainment activities, and so the different functions of the low-frequency phase and high-frequency power entrainment activities remained unclear. In Chapter IV, participants were requested to perform a pitch discrimination task, while the tones were embedded in a rhythmic (isochronous) or arrhythmic (non-isochronous) sequence. Both perceptual performance and EEG activities were measured. Perceptually, auditory temporal regularity (i.e., rhythm) facilitated pitch discrimination sensitivity. Neurally, the findings showed that low and high frequency entrainment and their coupling associated with different perceptual functions. Specifically, low

frequency entrainment associates with temporal regularity, high frequency entrainment associates with temporal attention, and their coupling associates with auditory-motor communication. Together, this study revealed that distinct perceptual functions of auditory neural oscillatory entrainment activities work in concert for rhythm tracking and proactively facilitate auditory perception.

IV.2 Abstract

Previous studies indicate that temporal predictability can enhance timing and intensity perception, but it is not known whether it also enhances pitch perception, despite pitch being a fundamental perceptual attribute of sound. Here we investigate this in the context of rhythmic regularity, a form of predictable temporal structure common in sound streams, including music and speech. It is known that neural oscillations in low (delta: 1–3 Hz) and high (beta: 15–25 Hz) frequency bands entrain to rhythms in phase and power, respectively, but it is not clear why both low and high frequency bands entrain to external rhythms, and whether they and their coupling serve different perceptual functions. Participants discriminated near-threshold pitch deviations (targets) embedded in either rhythmic (regular/isochronous) or arrhythmic (irregular/non-isochronous) tone sequences. Psychophysically, we found superior pitch discrimination performance for target tones in rhythmic compared to arrhythmic sequences. Electroencephalography recordings from auditory cortex showed that delta phase, beta power modulation, and delta-beta coupling were all modulated by rhythmic regularity. Importantly, trial-by-trial neural-behavioural correlational analyses showed that, prior to a target, the depth of U-shaped beta power modulation predicted pitch discrimination sensitivity whereas cross-frequency coupling strength predicted reaction time. These novel findings suggest that delta phase might reflect rhythmic temporal expectation, beta power temporal attention, and delta-beta coupling auditory-motor communication. Together, low and high frequency auditory neural oscillations reflect different perceptual functions that work in concert for tracking rhythmic regularity and proactively facilitate pitch perception.

IV.3 Introduction

Predictive temporal structure is essential across many aspects of human cognition and behaviour, ranging from proactively optimizing perceptual processing (Haegens and Zion Golumbic, 2018) and speech communication (Giraud and Poeppel, 2012) to facilitating interpersonal coordination (Chang et al., 2017a, 2019b; Savage et al., 2015) and prosocial behaviours (Cirelli et al., 2014, 2018). Rhythm is a form of predictable temporal structure, in which events tend to occur at regular time intervals or have recurring characteristics (Nobre and van Ede, 2018). Given that speech and music unfold over time, and sound events are fleeting, it is perhaps not surprising that speech and music are temporally structured (Ding et al., 2017) as the resulting rhythmic regularity (musical beat; syllable timing) enables prediction of when important upcoming information is likely to occur. In the present paper we manipulate rhythmic predictability to investigate whether temporal predictability enhances pitch perception.

Sounds can vary over time in multiple aspects. The temporal aspect involves variations in sound onsets, durations and rhythm; the intensity aspect involves variations in sound pressure; and the spectral aspect involves variations in frequency (e.g., resulting in pitch changes). Previous studies show that rhythmic regularity in auditory sequences facilitates detection of near-threshold timing and intensity deviations (e.g., Henry and Herrmann, 2014; McAuley and Fromboluti, 2014). However, to our knowledge, none have investigated the effect of rhythmic regularity on detecting near-threshold pitch changes, despite pitch being a fundamental perceptual attribute of sound crucial for speech, music and object identification. It is not obvious a priori that rhythmic regularity would enhance pitch perception similarly to how it enhances time and intensity perception because pitch perception can be dissociated from time and intensity perception. For example, people with amusia or tone deafness typically have auditory perceptual deficits in the spectral domain but not in the temporal or intensity domains (Peretz, 2016; Zendel et al., 2015). Therefore, it is important to investigate whether rhythmicity facilitates near-threshold pitch discrimination in order to determine whether rhythmicity enhances all of the major dimensions of auditory perception.

Neural oscillations in the brain synchronize their activity to rhythmic regularities in sensory input, which has been termed "neural entrainment", and

this is regarded as an important neural mechanism for tracking rhythmicity (Haegens and Zion Golumbic, 2018). In monkeys, electrophysiological recordings showed that the phase of low frequency oscillations (delta band: 1–3 Hz) time-lock and entrain to external rhythmic sensory input (Lakatos et al., 2013, 2016), and sensory input coincident with the excitatory phase is selectively facilitated (Schroeder and Lakatos, 2009; Calderone et al., 2014). In humans, electroencephalography (EEG) and magnetoencephalography (MEG) studies show that fidelity of delta phase entrainment associates with perceptual facilitation in the auditory temporal and intensity domains (e.g., Arnal et al., 2015; Bauer et al., 2018; Henry and Obleser, 2012; Henry et al., 2014; ten Oever et al., 2017), but it remains uncertain whether delta phase associates with facilitated pitch perception. Again, rhythmic facilitation might operate differently at the neural oscillatory level for perceiving pitch changes compared to temporal and intensity changes. Therefore, it is important to investigate whether delta phase associates with perceptual facilitation in the spectral domain as it does in the temporal and intensity domains.

Power modulations in high frequency oscillations (beta band: 15–25 Hz) also entrain to rhythmic regularity and reflect temporal prediction, but perceptual consequences have not been investigated. In response to a rhythmic tone sequence, beta power decreases following each tone onset, and then increases with the appropriate slope to anticipate the predicted onset time of the upcoming tone, dependent on sequence tempo (Fujioka et al., 2012; Cirelli et al., 2014). Beta modulation is disrupted in non-rhythmic sequences (Fujioka et al., 2009, 2012) and modulated by hierarchical timing structures (e.g., waltz, march) (Snyder and Large, 2005; Iversen et al., 2009; Fujioka et al., 2015). Beyond time, a few previous studies hinted at the role of beta in prediction of pitch (Chang et al., 2016, 2018; Sedley et al., 2016); however, perceptual performance was not measured or related to beta oscillations in these studies. Here we hypothesized that the magnitude of beta power entrainment in a rhythmic context associates with increased pitch sensitivity.

A fundamental question concerns why delta phase and beta power both entrain to external rhythms. Do they implement different perceptual functions? One MEG study showed pre-target delta phase, beta power, and delta-beta coupling all predict detection of a timing deviation (Arnal et al., 2015), but did not differentiate them functionally. Another MEG study suggested delta

and beta entrainments represent opposite directions of auditory-motor communication (Morillon and Baillet, 2017), but how they differ in perceptual functions remained unclear.

The current study investigates (1) whether rhythmic regularity facilitates near-threshold pitch discrimination and (2) how neural oscillations associate with such facilitation. Participants performed a pitch discrimination task embedded in either rhythmic (regular, isochronous) or arrhythmic (irregular, non-isochronous) tone sequences. Behaviourally, we hypothesized that rhythmic regularity would facilitate pitch discrimination sensitivity and reaction time. Neurally, we focused on the oscillatory signals generated from bilateral auditory cortices, as they are regarded to be the major sources of auditory neural entrainment activities (Fujioka et al., 2012; Morillon and Baillet, 2017) and the primary regions for processing pitch (McDermott, 2018). Specifically, we used dipole models as spatial filters to extract the source waveforms that focused on activity in auditory cortices. We hypothesized that rhythmicity would modulate oscillatory activities, including not only the phase of low-frequency oscillatory activity, but beta power and delta-beta coupling as well. Further, we explored how pre-target delta phase, beta power modulation and delta-beta coupling would predict different aspects of perceptual performance at a trial-by-trial level.

IV.4 Materials and methods

IV.4.1 Participants

The 16 participants (18–27 years old, mean age 19.3 ± 2.3, 11 females) were students at McMaster University and received course credit for completing the study. Participants were screened by a self-report survey to ensure they had normal hearing, were neurologically healthy and were right-handed. Signed informed consent was obtained from each participant. The McMaster University Research Ethics Board approved all procedures.

IV.4.2 Stimuli

Auditory stimuli were computer-generated complex tones. Each tone was composed by summing sinusoidal waves at the fundamental frequency (F0) and two overtones (F1 and F2) with slope -6 dB/oct and frequency components added in random phases. Tones had 10 ms rise and fall times and a 40 ms steady-state duration. Sounds were presented at 60.6 dB SPL (C weighted), measured through an artificial ear (type 4152, Brüel & Kjær) with a sound level meter (type 2270, Brüel & Kjær). All stimulus sequences were presented under the control of a digital signal processor (Tucker Davis RP2.1) and delivered binaurally via ear inserts (Etymotic Research ER-2).

IV.4.3 Procedure

Initially, a calibration session was conducted to estimate individual pitch discrimination thresholds in order to present near-threshold stimuli for each participant in the main experiment. We used a 2-alternative forced-choice procedure. On each trial, participants heard a standard tone (F0 fixed at 500 Hz) and a target tone (8.4 dB louder than the standard tone, as presented during the main experiment) that was either higher or lower in pitch (random presentation order), with a 500 ms inter-onset interval (IOI). Participants were required to judge which tone was higher in pitch by pressing one of two buttons. The target tone started at 500 * $e^{\pm 0.02}$ Hz (≈510.1 and 490.1 Hz), and the 1-up/2-down procedure adjusted the F0 of the target tone according to the performance on previous trials (Levitt, 1971), with step size $e^{0.002}$ (≈1 Hz). Over 40 trials, the 70.7% correct pitch discrimination thresholds were estimated for target tones higher and lower than the standard tone. The logarithmic midpoint between these two thresholds (one higher and one lower than 500 Hz) was the point of subjective equality (PSE), and the just-noticeable difference (Δ) was the logarithmic distance between the PSE and a threshold. This calibration session took 5–10 min. The mean estimated PSE was 498.07 ± 0.45 Hz and the estimated Δ was 3.69 ± 0.40 Hz across participants (mean ± standard error).

To manipulate rhythmic regularity (Rhythmicity), the experiment included Rhythmic (isochronous) and Arrhythmic (non-isochronous) sequences (Fig. 1A). Each sequence was composed of 100 tones, 90% of which were standard

tones (F0 = 500 Hz), and 10% were target tones. Two adjacent target tones were separated by 6–12 standard tones. Target tone F0s were presented at 7 different levels, specifically PSE + [-5Δ, -3Δ, -Δ, 0, +Δ, +3Δ, +5Δ] (logarithmic scale), and there were 8 target tones at the -5Δ level and 7 target tones at each of the other levels in each condition of each run. A run consisted of 5 Rhythmic and 5 Arrhythmic sequences, randomly intermixed. Each participant completed 5 runs for a total of 50 sequences, for a total of 250 target tones in each of the Rhythmic and Arrhythmic conditions. The tones in Rhythmic sequences were presented isochronously with 500 ms IOI. In contrast, the tones in Arrhythmic sequences were presented with IOI uniformly random between 250 and 750 ms. The length of each Arrhythmic sequence was controlled to be between 48 and 52 s, matching the 50 s length of Rhythmic sequences. Importantly, the IOI before and after each target tone was fixed at 500 ms in both Rhythmic and Arrhythmic conditions to control for any possible effects of foreperiod or masking on behavioural or EEG responses (Rohenkohl et al., 2012; Cravo et al., 2013; Herrmann et al., 2016).

Target tones were presented 8.4 dB louder than standard tones, which served as a cue for the participant to respond. All participants were able to clearly detect this loudness change. Participants were instructed to judge whether the pitch of each loud (target) tone was higher or lower than the other (standard) tones, by pressing one of two buttons with their thumbs (right for higher and left for lower) on a controller, as accurately and as quickly as possible. Participants practiced this task in both conditions prior to the experimental session. The button pressed and reaction time (RT) of the first response within 1600 ms was recorded for each target tone. Ten participants who missed more than 3% of targets were considered to be not following the instructions (Fig. S1), and were not included in the n = 16 of the subsequent analyses. We used this strict exclusion criterion because most participants tended to miss responses when the target levels were difficult (around PSE). Therefore, a strict criterion made sure the participants included in the analyses had relatively low bias. After finishing a sequence, the next sequence started when the participant was ready and pressed a button. After every run, there was a mandatory break of at least 1 min before starting the next run. The experimental session lasted approximately 1 h.

IV.4.4 Psychometric model fitting

Trials with premature response (RT < 200 ms) were excluded as missing responses. The proportion of "higher" response was calculated based on all trials that were responded to for each target level for each condition for each participant.

The psychometric data from each participant and condition was fitted by a function ψ, which combines a logistic model (F) and additional parameters for controlling lapse rate (Kingdom and Prins 2010):

$$F(x;\alpha,\beta) = \frac{1}{1+e^{-\beta(x-\alpha)}}$$

$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)F(x;\alpha,\beta)$$

which was defined by x (the frequencies of target tones on a logarithmic scale) and parameters α (the x point at which 50% of responses are "higher") and β (the slope of psychometric function, or discrimination sensitivity) of the logistical model, and additional parameters γ and λ (lapse rate, i.e., the probability of making an incorrect response independent of the stimulus intensity). In the first step, each of γ and λ was empirically fixed to the mean response at [-5Δ, -3Δ] or [+3Δ, +5Δ] levels for each model (lapse rate was measured at the extreme target levels but lapses can occur at any target level). Second, the fitting of free parameters α and β of function ψ was implemented with PAL_PFML_Fit function in Palamedes toolbox (RRID: SCR_006521) (Prins and Kingdom, 2009), which iteratively searches the best-fitted parameters across parameter space using a maximum likelihood criterion. Goodness of fit for each participant was estimated by comparing predicted and observed accuracy, with R2 > 95%. The individual perceptual judgment performances and the fitted psychometric functions are shown in Fig. S2.

IV.4.5 Electroencephalographic recording and preprocessing

The EEG was sampled at 2048 Hz (filtered DC to 417 Hz) using a 128-channel Biosemi Active Two amplifier (Biosemi B.V., Amsterdam). The electrode array was digitized (Polhemus Fastrak) for each participant prior to recording. EEG

data were stored as continuous data, and transformed to an average reference offline. In subsequent steps, the EEG data was processed in MATLAB using the FieldTrip toolbox (RRID: SCR_004849) (Oostenveld et al., 2011).

We employed independent component analysis (ICA) to remove artifact signals (Jung et al., 2000). For performing ICA, we filtered the continuous EEG data 0.7–100 Hz and segmented it into -3 to 54 s time windows, time-locked to the first tone of each sequence. The ICA was performed on all sequences of each participant, and the components reflecting artifact, including eye blinking, eye movement, electrocardiogram, and 60-Hz powerline noise, were identified by visual inspection. Once obtaining the unmixing and mixing matrices from ICA, we went back to the unfiltered data, projected it to ICA space, removed the artifact components as identified above, and then the data was projected back to 128-channel space for all subsequent analyses. We used a subsequent 120 μV criterion on epochs to exclude the ones containing artifacts which were not removed by ICA, with removal rate ranging between 0.2 and 17.0% across participants. To facilitate the speed of subsequent processes, the data was downsampled to 256 Hz.

IV.4.6 Modeling dipole sources for auditory cortex

We employed a dipole source model as a spatial filter in order to extract the EEG signal generated from left and right auditory cortices, following previous studies (Fujioka et al., 2012; Chang et al., 2016, 2018). In the present study, we were primarily interested in responses from auditory cortex, so we analyzed the EEG signals in source space rather than from surface channels, by extracting the oscillatory signals generated from left and right auditory cortices while attenuating signals generated from other brain regions.

Using dipole source modelling enabled us to perform hypothesis driven analyses on the EEG signal generated primarily from bilateral auditory cortices. Most previous entrainment studies that analyze surface electrodes either select the electrodes that have maximum auditory ERP responses or are known to reflect auditory responses, such as frontocentral electrodes, and these studies interpret the results as generated from auditory cortex (e.g., Henry et al., 2014; ten Oever et al., 2017). The dipole approach has advantages in that it is a linear weighting of all the surface electrodes that emphasizes the response from auditory cortex. It also enables assessment of hemispheric effects. Using

auditory cortical sources is appropriate as electrophysiological evidence from monkeys shows that delta entrainment is observed in auditory cortex (Lakatos et al., 2013; Calderone et al., 2014), giving a physiological basis for focusing on the auditory EEG source.

We estimated the locations of bilateral auditory cortices by localizing the P1 ERP (event-related potential) component. First, the continuous EEG data was band-pass filtered 0.7–40 Hz and then segmented into -100 to 200 ms epochs, time locked to standard tone onset in all sequences, which were averaged into ERP (event-related potential) waveforms and used to model dipole sources. Second, two auditory cortex sources were estimated for each participant for the auditory evoked P1 (∼60–90 ms) with the dipole locations constrained to be symmetric across hemispheres in location but not orientation, using the multiple source probe scan algorithm and the four-shell ellipsoid model included in the Brain Electrical Source Analysis (BESA, RRID: SCR_009530) software package. P1 was chosen because it is the dominant peak at fast presentation rates (N1 peaks are strongly reduced at fast rates; Näätänen and Picton, 1987), and is generated primarily from primary auditory cortex (Godey et al., 2001). The mean locations of fitted dipoles across participants were at Talairach coordinates [±45.6, -14.9, 23.8] with approximate mean orientations [0.0, 0.7, 0.7] and [0.1, 0.8, 0.6] for left and right dipoles. These locations are close to bilateral primary auditory cortices (far below the 1 cm averaged error range of four-shell head model; Slotnick, 2004) with orientations toward the mid-frontal surface area, consistent with typical auditory evoked potentials. The mean residual variances of the source fittings was 9.7% (range 2.4–19.6%). Finally, using the dipole model, the unfiltered continuous 128-channel EEG was projected into source-space EEG for further time-frequency analyses. Note that the current procedure only ensures that the source-space EEG signals are dominantly generated in the auditory cortex, given the low residual variances of the source fitting. Considering the spatial specificity of EEG, brain signals generated from the neighboring areas may contribute to the source waveforms, but they should have a relatively small contribution compared to the auditory brain signals.

IV.4.7 Delta band (1–3 Hz) analyses

We were interested in the delta phase during the pre-target interval, which has been shown to phase-lock with the stimulus presentation rate and thus reflect temporal entrainment. To initially identify whether low-frequency oscillations in the EEG entrained to the presentation rate of the auditory sequence, we frequency transformed each unfiltered 50 s sequence, using 0.1 Hz frequency bins ranging 1–7 Hz and a Hanning taper to avoid edge artifacts.

To isolate the delta band, we filtered the continuous source-space EEG at 1–3 Hz with Butterworth (zero-phase, third order) high-pass and low-pass filters, performed a Hilbert transform to obtain the instantaneous phase, and then segmented into -1.5 to 1.5 s epochs, time-locked to target onsets. We used the Butterworth filter because we aimed to create a narrow band-pass filter while minimizing ripple effects in the time domain signal (Widmann et al., 2015), which follows previous studies with similar purposes (Besle et al., 2011; Ng et al., 2012a, 2012b; Cravo et al., 2013; Tal et al., 2017).

The inter-trial phase clustering (ITPC), sometimes referred to as the phase-locking value or phase coherence, was used to measure the phase distribution of delta band at each time point across trials (Cohen, 2014a), which ranges from 0 (circular-uniformly distributed phases) to 1 (perfectly identical phases). Higher ITPC of delta phase represents more precise phase entrainment (Tal et al., 2017; ten Oever et al., 2017).

To correlate delta phase and pitch discrimination sensitivity, we first ordered the trials based on the sum of absolute phase deviations from the mean phase (the aligned phase of delta entrainment) at each time point across the pre-target interval (-0.5 to 0 s), and then binned the ordered trials into 10 bins (with 50% of trials overlapped across adjacent bins), within each target level, condition, and participant. Second, we estimated the pitch discrimination sensitivity (β or slope) by modeling the behavioural performance with psychometric function (as described in Psychometric model fitting) for each bin across all target levels. Third, we performed a Spearman rank correlation between ranked absolute phase deviation and pitch discrimination sensitivity, and obtained a Spearman rank correlation coefficient for each participant and condition. Finally, each correlation coefficient was further z-normalized to a bootstrapped null distribution for each condition and participant. Each null distribution was built by performing the same preceding steps on randomly

paired behavioural response and delta phase deviation across trials for 1000 iterations within each target level, which represented the null distribution of an uncorrelated effect. In sum, more negative z-values (z-normalized Spearman rank correlational coefficient) indicate a higher association between better pitch discrimination sensitivity and pre-target delta phase being closer to the entraining phase (optimal phase).

To correlate delta phase and RT, we performed a linear correlation between RT and the absolute phase deviation in the pre-target interval across all correct trials, within each target level, condition, and participant. Each observed correlation coefficient was further z-normalized to a bootstrapped null distribution. Each null distribution was built as above by randomly pairing RT and delta phase deviation across trials for 1000 iterations, which represented the null distribution of an uncorrelated effect. A mean z-value for each condition and participant was obtained by averaging z-normalized correlational coefficients across target levels. We only include the PSE $\pm\Delta$ and $\pm3\Delta$ target levels, because the levels at PSE $\pm5\Delta$ were very easy to discriminate, and the pitch of target tones at the PSE level was not perceptually discriminable from the pitch of standard tones. The mean standard deviation of RT across participants at $\pm5\Delta$ levels was 113.81 $\pm$ 9.99 ms, which is smaller than the mean standard deviation of RT at $\pm3\Delta$ and $\pm\Delta$ levels, which was 146.86 $\pm$ 11.44 ms ($t(15) = -5.23$, $p < 0.001$). The smaller RT variance at the $\pm5\Delta$ levels likely reflects near ceiling performance. In sum, higher mean z-values indicate that pre-target delta phases that are closer to the entraining phase (optimal phase) are associated with faster (shorter) RTs to the target.

IV.4.8 Beta band (15–25 Hz) analyses

We were interested in the induced (non-phase-locked) power in beta band in the pre-target interval. Previous studies show that power decreases (desynchronization) and increases (resynchronization) to entrain to the rate of the beat of an externally presented tone sequence, and thus reflects temporal prediction (Snyder and Large, 2005; Iversen et al., 2009; Fujioka et al., 2009, 2012; Cirelli et al., 2014; Fujioka et al., 2015, but see Meijer et al., 2016). To obtain induced activities, the trial-averaged ERP waveform of -1.5 to 1.5 s around target tone onset was subtracted from each target epoch within each condition and participant (Fujioka et al., 2012; Cohen, 2014a; Chang et al.,

2016, 2018). The Morlet wavelet transformation was performed on each target epoch of induced activities (frequency bin size = 1 Hz; 10 cycles). The beta power was baseline corrected by percentage change to the mean power in the pre-target interval for each frequency bin, and the beta power time series was obtained by averaging induced power across 15–25 Hz. Edge artifacts did not affect the time-frequency activities at the pre-target interval.

Following our previous studies (Fujioka et al., 2012; Chang et al., 2018), for each condition, we first identified the latency of beta desynchronization (decrease in power) as the time point with minimum beta power within the pre-target interval and the resynchronization (power rebound following the desynchronization) as the time point with beta power closest to the baseline (0% power change) between the latency of desynchronization and time point 0 s.

To describe the shape of beta power modulation for each trial, we fitted a quadratic (parabola) function ($y = ax^2 + bx + c$) on the induced beta power time series in the pre-target interval (data was centered and scaled, and fitted with bisquare weights method). The fitted quadratic coefficient (a) reflects the modulation peak-trough shape of beta power, with more positive values representing U-shaped modulation with deeper troughs, 0 representing linear shape, and more negative values representing ∩-shaped modulations with higher peaks. The vertex position ($-b/2a$) estimates the beta desynchronization latency. The mean R2 of fitting across participants was 59.6 ± 0.6%. This approach allows us to simply use one parameter to quantify the peak-trough shape, and one parameter to quantify the vertex latency of the beta power fluctuation, which is suitable to model single-trial neural activities and perform trial-by-trial EEG-behaviour correlations. This approach is better than a peak-searching approach at a single-trial level as it acts like a low-pass filter; it avoids identifying spurious noise peaks while extracting the slow modulation shape and vertex.

To correlate the size of the beta modulation trough or desynchronization latency with pitch discrimination sensitivity, following the procedure described above, we ordered and binned the trials into 10 bins (with 50% of trials overlapped across adjacent bins), based on each fitted beta modulation index (quadratic coefficient or vertex position), within each target level, condition, and participant, and then estimated the pitch discrimination sensitivity (β,

slope) for each bin across target levels. We performed a Spearman rank correlation between each ranked beta modulation index and pitch discrimination sensitivity across target levels, and further z-normalized to a bootstrapped null distribution for each participant and condition. In sum, for the beta power modulation shape, higher z-values show a greater association between deeper pre-target beta modulation trough (ideal U-shaped beta power) and better pitch discrimination sensitivity. For the beta desynchronization latency, higher z-values show a greater association between later beta desynchronization latency and better pitch discrimination sensitivity.

To correlate beta modulation shape with RT, following the procedure described above, each beta modulation index was linearly correlated with RT across trials within each target level, normalized to z-value (the bootstrapped null distribution of randomly paired RT and quadratic coefficients across trials). A mean z-value was obtained for each condition and participant by averaging z-values across target levels. In sum, for the beta power modulation shape, lower mean z-values indicate a greater association between deeper pre-target beta modulation trough (ideal U-shaped beta power) and faster (shorter) RT. For the beta desynchronization latency, lower mean z-values indicate a greater association between later beta desynchronization latency and faster (shorter) RT.

IV.4.9 Delta-beta coupling analyses

Given that both delta phase and beta power modulation are affected by rhythmic regularity, we were interested in the cross-frequency phase-amplitude coupling (PAC) between delta and induced beta activities. Phase-amplitude coupling is thought to result from the excitatory or inhibitory phase of neural circuits oscillating at low-frequencies affecting the power of high-frequency oscillations (Hyafil et al., 2015b). To estimate the delta-beta coupling, we employed the dPAC index, which was calculated for each trial across time points,

$$dPAC == |\frac{1}{n}\sum_{t=1}^{n} A_t(e^{i\varphi_t} - \frac{1}{n}\sum_{t=1}^{n} e^{i\varphi_t})|$$

where the n signifies the total number of time points, A_t the power amplitude of modulated frequency, and φ_t the phase of the modulating frequency at time point t. The dPAC is the debiased version of the widespread PAC method (Canolty et al., 2006) for solving phase clustering bias (van Driel et al., 2015). In the present study, we estimated dPAC with reversed beta power polarity to reflect the modulation trough rather than peak. Each observed dPAC was further transformed into a z-value (dPACz) by comparing with 1000 time-shuffled dPAC (null distribution).

To correlate delta-beta coupling with pitch discrimination sensitivity, following the procedure described above, the dPACz values were rank-correlated with discrimination sensitivity across 10 dPACz-ranked bins (with 50% of trials overlapped across adjacent bins), and then a z-normalized correlational coefficient (relative to the bootstrapped null distribution) was obtained for each condition and participant. In sum, a higher z-value indicates a greater association between better delta-beta coupling and better pitch discrimination sensitivity.

To correlate delta-beta coupling with RT, the dPACz scores were linearly correlated with RT across trials within each target level, normalized to z-values (relative to the bootstrapped null distribution). A z-value was obtained for each condition and participant by averaging z-values across target levels. In sum, lower mean z-values indicate a greater association between better delta-beta coupling and faster (shorter) RT.

IV.4.10 Partial out covariances among EEG indexes for EEG-behaviour correlations

The covariances were relatively small among the EEG indexes (delta phase deviation, quadratic coefficient of beta power modulation peak-trough shape, vertex latency of beta desynchronization, and dPACz) across the two hemispheres (see Supplementary Material and Table S1). However, to exclude any potential confounding effects among these covariances on the EEG-behaviour correlations, each EEG index of each hemisphere was partialled out by all remaining indexes across both hemispheres (taking the residuals from a regression model in which each EEG index was predicted by all other indexes), prior to each above-mentioned correlational analysis. The only exception was

that the quadratic coefficient and the vertex latency of the beta power modulation shape were not partial out from each other, because these two indexes are mathematically dependent in the quadratic function.

IV.4.11 Experimental design and statistics

Participant (n = 16) was the random factor for the within-subject statistical tests in the current study. The assumptions of parametric tests were checked (Kolmogorov-Smirnov test with threshold $p < 0.01$), and the alternative test or non-parametric permutation tests (1000 iterations) were used when the assumptions were violated. The test being used is specified below for each p-value. The statistical tests were performed by MATLAB (2015b) or R (3.3.3). Statistical decisions were based on two-tailed tests. Multiple comparisons of all the planned statistical tests of the current study were controlled by experiment-wise false discovery rate (FDR) (Benjamini and Hochberg, 1995; Luck and Gaspelin, 2017), and each corrected p-value was reported as p_{FDR}.

To access the EEG-behaviour trial-by-trial correlational effects, we performed statistical tests on the z-normalized correlational coefficients across participants. It is important to note that we did not test the statistical significance of each within-participant correlational coefficient, but we tested whether the within-participant correlational effect was consistent across all participants. In other words, for each condition, we used a one-sample t-test to assess whether the z-normalized correlational coefficients across participants were significantly different from 0 in the same direction. This approach is statistically similar to a linear mixed-effect model with random intercepts and slopes, which assesses the influence of predictors of interest while taking into account the variance across participants (Barr et al., 2013). The current approach fits our purpose better than a linear mixed-effect model because it can incorporate non-parametric rank correlation and bootstrapping z-normalization. Note that the trial-by-trial correlational analyses were only performed within the Rhythmic condition, as we aimed to investigate how neural activities anticipatorily associate with perceptual performance, when rhythmic temporal regularity is available. An extended question is whether rhythmicity moderates these EEG-behaviour relationships, that is, whether such a relationship would be

diminished in the Arrhythmic compared to Rhythmic conditions. We therefore performed post-hoc analyses to investigate this question, and the results are reported in the Supplementary Material.

IV.5 Results

IV.5.1 Behavioural performance and psychometric modeling

The psychometric modeling results showed that rhythmic regularity facilitated pitch discrimination sensitivity (Fig. 1B). The frequency of 50% responding "higher" (converted from α of logistic model) was not significantly different between Rhythmic and Arrhythmic conditions ($t(15) = -0.16$, $p_{FDR} = 0.945$, paired t-test; Fig. 1C). The discrimination slope (β of logistic model) was significantly higher in the Rhythmic than Arrhythmic condition ($t(15) = 2.97$, $p_{FDR} = 0.029$, paired t-test; Fig. 1D), suggesting that participants have better pitch discrimination sensitivity in the rhythmic than arrhythmic context. This is equivalent to saying the required frequency difference to achieve 70.7% pitch discrimination accuracy decreased 11.5% ± 5.3% on a logarithmic frequency scale, comparing the Rhythmic to Arrhythmic conditions. This effect size is at the same level (approximately 10%) as other similar behavioural studies (Haegens and Zion Golumbic, 2018).

The RT results are shown in Fig. 1E. A repeated measure MANOVA (alternative ANOVA without sphericity assumption, O'Brien and Kaiser, 1985) on RT (correct trials only) with variables Rhythmicity (Rhythmic, Arrhythmic) and Target Levels (PSE ± [Δ, 3Δ, 5Δ]) showed significant main effects of Rhythmicity (Wilks(1) = 0.49, approximate $F(1,15) = 15.55$, $p_{FDR} = 0.010$) and Target Level (Wilks(1) = 0.12, approximate $F(5,11) = 15.83$, $p_{FDR} = 0.002$), but no interaction effect (Wilks(1) = 0.95, approximate $F(5,11) = 0.11$, $p_{FDR} = 0.988$). A linear trend analysis (collapsing Rhythmicity) further showed that RTs are longer when the target level is closer to the PSE ($t(15) = 9.39$, $p_{FDR} < 0.001$) as expected. As for the main effect of Rhythmicity, the averaged RT across target levels (excluding the PSE level) in the Rhythmic condition was 16.3 ms shorter than in the Arrhythmic condition, indicating that temporal regularity facilitates behavioural response speed. This effect

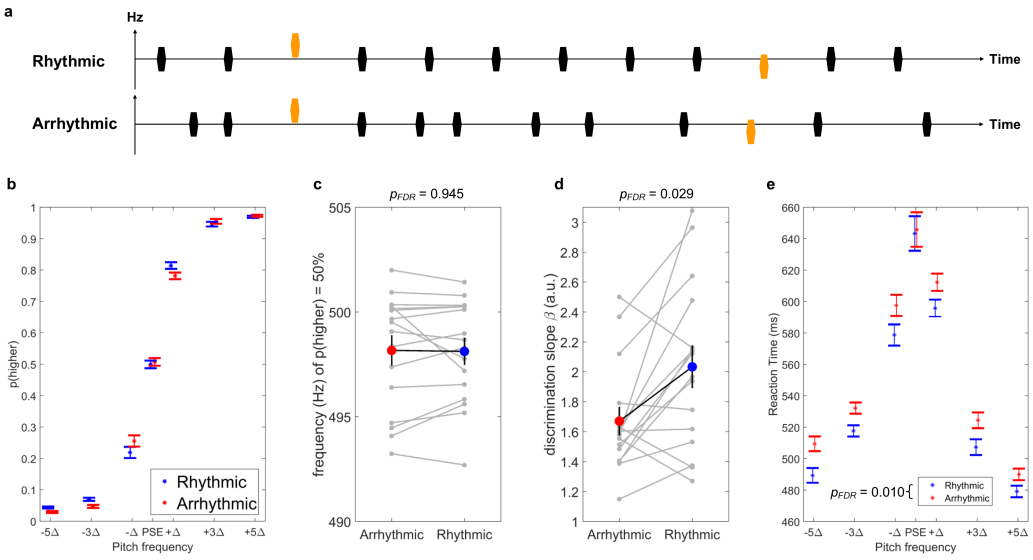

FIGURE IV.1: Experimental design and behavioural performance. (A) Stimuli and task. In both conditions, 10% of tones were targets (yellow), which were higher or lower in pitch than standards (black) at a near-threshold. Participants judged whether target tones were higher or lower than standards. Tones in the Rhythmic sequence were isochronous with 500 ms IOI (inter-onset interval), and tones in the Arrhythmic (non-isochronous) sequence were presented with IOI uniformly random between 250 and 750 ms. Importantly, the IOI before and after each target tone was fixed at 500 ms in both conditions. (B) Psychometric performance on the perceptual judgment. For purposes of visualization, the error bars represent within-subject error (Cousineau, 2005) for the p(higher) at each level. Both PSE and Δ were individually determined by an initial adaptive psychophysical procedure. We fitted a logistic psychometric model for each participant's performance under each condition. Fitted parameters are presented in (C) and (D); each connected dot represents one participant's data, and the colored dots and error bars represent the mean ± standard error across participants. (C) The frequency of 50% responding "higher" (converted from α of logistic model) was not significantly different between Rhythmic and Arrhythmic conditions. (D) Discrimination sensitivity. The discrimination slope (β of logistic model) was significantly higher in the Rhythmic than Arrhythmic condition, suggesting participants have better pitch discrimination sensitivity in the Rhythmic than Arrhythmic context. (E) Response Time (RT). The correct trial RT on PSE ± [Δ, 3Δ, 5Δ] levels was shorter in the Rhythmic than Arrhythmic condition, suggesting rhythmic regularity facilitates behavioural response speed. Error bars represent within-subject error. For visual representation, RTs on PSE level were based on both correct and error trials, and RTs on other levels were based on only correct trials.

size is similar to previous studies reporting that rhythmicity facilitates RT (e.g., Morillon et al., 2016).

IV.5.2 Beta power modulation is affected by rhythmicity and associates with pitch discrimination sensitivity

We were interested in the induced (non-phase-locked) power activities in beta band (15–25 Hz) in the pre-target interval, as previous studies showed that beta power decreases (desynchronization) and increases (resynchronization) so as to entrain to the rate of the beat of an externally presented tone sequence, and thus reflects temporal prediction (Snyder and Large, 2005; Iversen et al., 2009; Fujioka et al. 2009, 2015, 2012; Cirelli et al., 2014). Replicating these previous studies, the induced beta power fluctuations in bilateral auditory cortices (Fig. 2A) showed that power decreased following the onset of a standard tone, and then increased in anticipation of the upcoming target tone, prior to the 0 ms point. In the left auditory cortex, the desynchronization latency was earlier in the Arrhythmic than Rhythmic condition ($t(15) = -2.88$, $p_{FDR} = 0.032$, permutation test), as was the resynchronization latency ($t(15) = -2.70$, $p_{FDR} = 0.046$, permutation test) (Fig. 2B). There were no significant latency effects in the right auditory cortex (t-values < 1.26 and p-values$_{FDR}$ > 0.347, permutation test; Fig. 2B). The resynchronization latency difference is consistent with a previous study (Fujioka et al., 2012) that found that resynchronization latency is earlier in arrhythmic than rhythmic conditions. It was suggested that the beta power resynchronization latency adapts to the onset time of the next tone when it is temporally predictable, but when timing is uncertain, beta resynchronization reflects preparation for the possibility of an early tone onset (Fujioka et al., 2012). The current study was not a full replication of the previous study, as we did not include variation in the stimulus presentation rate. Nevertheless, the current findings are consistent with the previous study, suggesting that beta power modulation in the Rhythmic condition reflects temporal prediction and thus is consistent with beta power reflecting entrainment.

Beyond replicating the effects of latencies, additional exploratory analyses did not showed any differences on beta desynchronization depth between Rhythmic and Arrhythmic conditions in either left (-6.61 ± 0.74 vs. $-5.16 \pm$

0.48, t(15) = -1.38, p_{FDR}= 0.338, permutation test) or right (-6.65 ± 0.65 vs. -6.41 ± 0.61, t(15) = -0.32, p_{FDR}= 0.869, permutation test) auditory cortex.

We further investigated whether pre-target beta power modulation indexes are anticipatorily associated with perceptual performance, as our previous study showed that the depth of the beta power modulation trough reflects expectations for pitch (Chang et al., 2018). In the present paper, at a trial-by-trial level, we correlated both pitch discrimination sensitivity and RT with the depth of the beta trough (the depth of the U-shaped fluctuation) or beta desynchronization latency in the Rhythmic condition, where the rhythmic temporal regularity is available. The power modulation was modeled by a quadratic (parabola) function, and the U-shaped fluctuation depth and vertex (desynchronization) latency are represented by fitted coefficients (see Method: Beta band (15–25 Hz) analyses for details).

The correlational analyses on beta modulation depth and discrimination sensitivity (Fig. 3A) showed that, in left auditory cortex, deeper beta modulation troughs are associated with increased pitch discrimination sensitivity (t(15) = 3.10, p_{FDR}= 0.025, one-sample t-test). The same analyses in right auditory cortex did not show a significant result (t(15) = -0.55, p_{FDR}= 0.741, one-sample t-test). This effect was significantly stronger in left than right auditory cortex (t(15) = 2.62, p_{FDR}= 0.045, paired t-test). On the other hand, depth of the beta modulation trough was not significantly related to RT (Fig. 3B) for either left (t(15) = 1.05, p_{FDR}= 0.474, one-sample t-test) or right (t(15) = 0.10, p_{FDR}= 0.969, one-sample t-test) auditory cortex.

The correlational analyses on beta desynchronization latency and discrimination sensitivity did not show any significant results (left: t(15) = 0.51, p_{FDR}= 0.750; right: t(15) = -0.96, p_{FDR}= 0.480; one-sample t-tests; Fig. 3C). The correlational analyses on beta desynchronization latency and RT did not show any significant results either (left: t(15) = -0.23, p_{FDR}= 0.909; right: t(15) = -0.94, p_{FDR}= 0.480; one-sample t-tests; Fig. 3D).

In sum, induced beta power modulation is affected by the rhythmic regularity of the auditory sequence, and trial-by-trial analyses show that a deeper beta power modulation trough prior to a deviant pitch is anticipatorily associated with better pitch discrimination sensitivity.

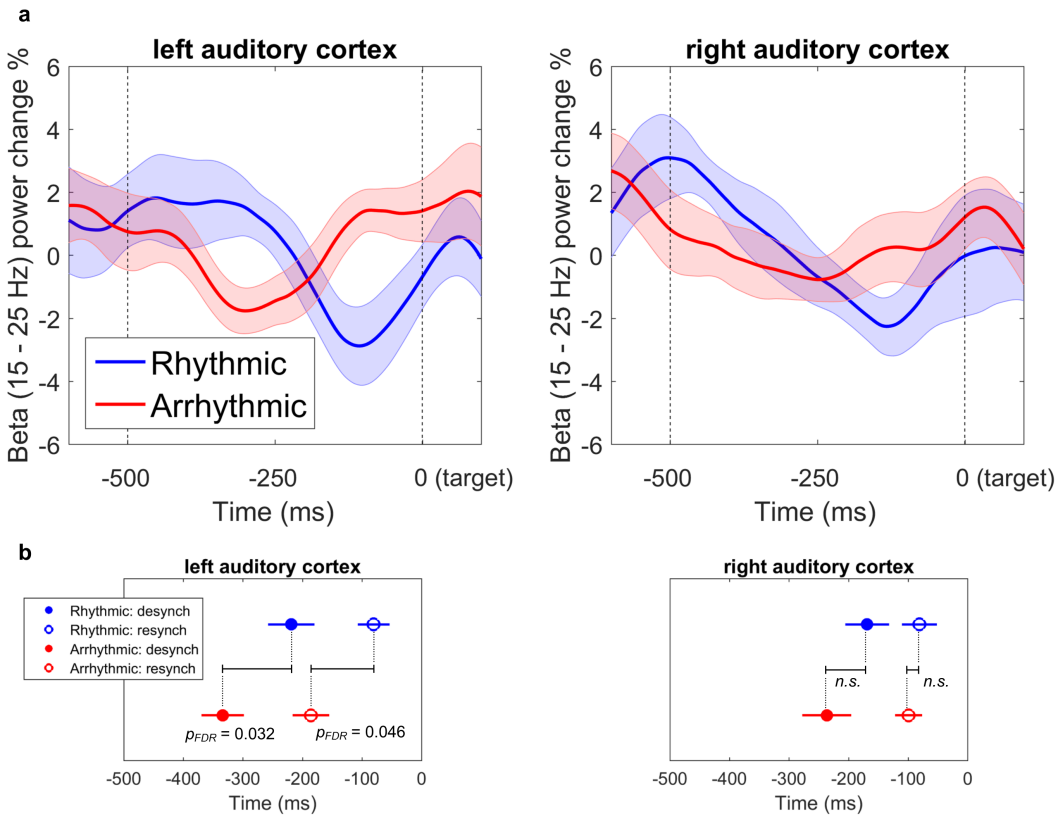

FIGURE IV.2: Beta power modulation is affected by rhythmicity. (A) Beta power (15–25 Hz) decreases (desynchronization) following the onset of a standard tone, and then increases (resynchronization) again in anticipation of the upcoming target tone. (B) Beta desynchronization and resynchronization latencies were both earlier in the Arrhythmic than Rhythmic condition in left auditory cortex (but not in the right auditory cortex). Error bars represent standard error of mean. (n.s.: non-significant).

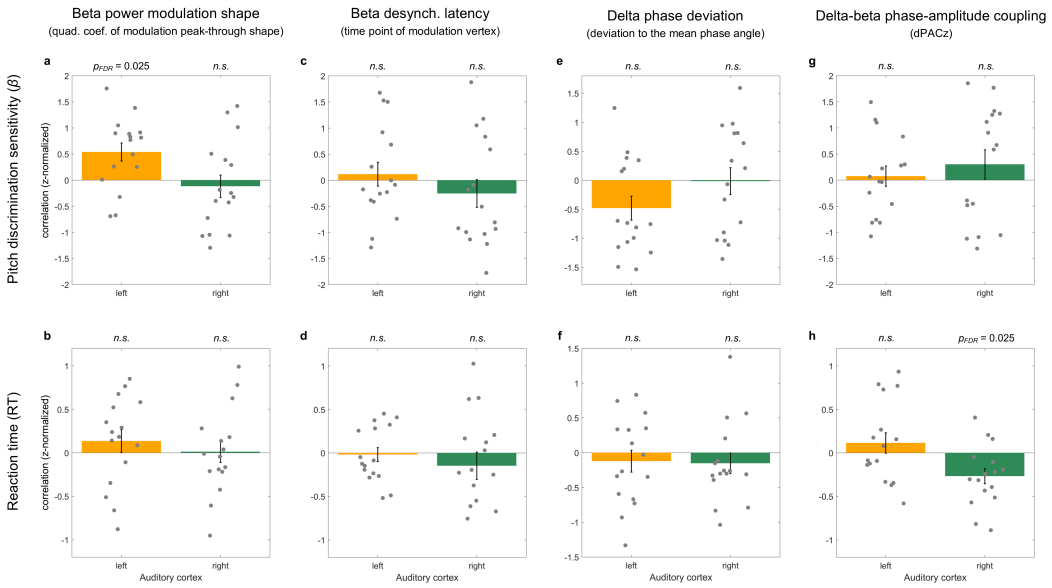

FIGURE IV.3: Distributions of trial-by-trial correlational strength between pre-target neural oscillatory activities and behavioural performance on the target tone in the Rhythmic condition. For the beta power on each trial, we modeled the pre-target beta power modulation peak-trough shape with the quadratic coefficient (quad. coef.) of a quadratic (parabola) function. The beta desynchronization (desynch.) latency was also estimated by the vertex of the same quadratic function. For the delta phase on each trial, we extracted the absolute phase deviation from the mean phase (the aligned phase of delta entrainment) across the pre-target interval. The dPACz index was used to estimate the delta-beta phase-amplitude coupling strength on each trial. For each participant, trial-by-trial neural activities were correlated with discrimination sensitivity and with RT, after the covariances among these EEG indexes were partialed out (see Methods for details). (A–H) The distribution of correlational strengths across participants is presented as dots, and the mean ± standard error across participants is presented as bar graphs ± error bars. Results showed that (A) deeper pre-target beta power U-shaped modulation trough in left auditory cortex predicts better pitch discrimination sensitivity, and (H) better pre-target delta-beta coupling in right auditory cortex predicts faster RT. (n.s.: non-significant).

IV.5.3 Delta phase is modulated by rhythmicity but not associated with perceptual performance

We were interested in the delta (1–3 Hz) phase in the pre-target interval, which has been shown to phase-lock with stimulus presentation rate and thus reflect temporal entrainment (e.g., Tal et al., 2017; ten Oever et al., 2017). We first examined the frequency content of the unfiltered time-domain waveform in response to each stimulus sequence to confirm that delta oscillations entrained to the auditory presentation rate (2 Hz). The power spectrum (Fig. 4A) shows a clear peak at 2 Hz in the Rhythmic but not in the Arrhythmic condition (left auditory cortex: $t(15) = 2.80$, $p_{FDR} = 0.035$; right auditory cortex: $t(15) = 3.63$, $p_{FDR} = 0.020$; paired t-tests), as well as at harmonic frequencies of 4 and 6 Hz (t-values > 3.07, p-values$_{FDR} < 0.025$, paired t-tests). Thus this initial analysis shows delta oscillations entrain to the 2 Hz stimulus presentation rate.

The ITPC (inter-trial phase clustering) was calculated on band-pass filtered waveforms (delta band 1–3 Hz) for both Rhythmic and Arrhythmic conditions (Fig. 4B). The group-mean ITPC for each condition in our study was around the same level reported in previous studies (e.g., Henry and Obleser, 2012; ten Oever et al., 2017), suggesting that we observed a robust delta-band ITPC signal at auditory dipoles of similar magnitude as in previous studies. The mean ITPC across the pre-target interval was higher in Rhythmic than Arrhythmic conditions in both left ($t(15) = 1.98$, $p_{FDR} = 0.046$, permutation test) and right ($t(15) = 3.67$, $p_{FDR} = 0.023$, permutation test) auditory cortices. This finding replicates previous studies showing that the delta phase phase-locks to the rhythmic regularity in the auditory input. The reason we averaged the ITPC across the entire pre-target interval rather than reporting the ITPC time series results was because the beta power modulation and dPACz indexes considered the entire pre-target interval, and we wanted the ITPC index to represent the information over the same time-interval.

An interesting observation is that some previous studies reported a stronger rhythmicity effect on ITPC using surface electrodes than our findings using auditory dipoles (e.g., ten Oever et al., 2017). However, as we aimed to investigate delta entrainment activities generated by auditory cortex, analyzing ITPC using surface electrodes is a suboptimal approach for the current study, because surface electrodes will likely include delta entrainment signals generated from other brain regions (Besle et al., 2011; Morillon and Baillet, 2017).

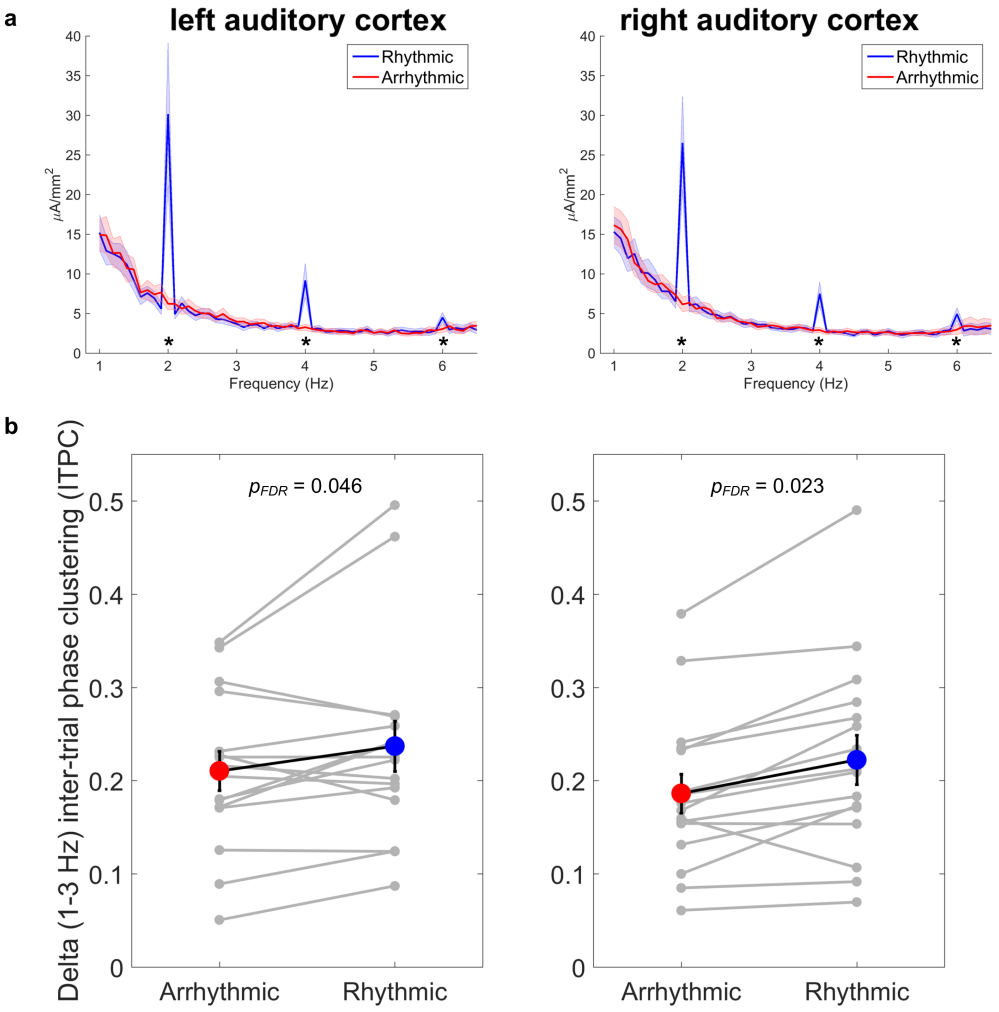

FIGURE IV.4: Delta phase entrains to rhythmic regularity. (A) A frequency transformation of the unfiltered auditory EEG signal (extracted from auditory cortex) for each individual sequence (∼50 s) for each condition showed that the power at 2 Hz was stronger in the Rhythmic than Arrhythmic condition in both cortices. (*: $p_{FDR} < 0.05$) (B) We further band-passed filtered the signal at the delta range (1–3 Hz) and then obtained the instantaneous phase with a Hilbert transform. The inter-trial phase clustering (ITPC) averaged over the pre-target interval was obtained for each condition and participant. ITPC ranges between 0 and 1, where 1 represents perfect phase-locking across trials, and 0 represents uniformly distributed phase across trials. The results showed that ITPC was higher in Rhythmic than Arrhythmic condition in both cortices. Each connected dot represents one participant's data, and the colored dots and error bars mean ± standard error across participants.

We further investigated whether the pre-target delta phase was anticipatorily associated with subsequent perceptual performance. On a trial-by-trial level, we separately correlated pitch discrimination sensitivity and RT with the phase deviation on each trial, relative to the mean phase in the Rhythmic condition (see Method: Delta band (1–3 Hz) analyses for details). If the mean (entrained) phase prior to a target represents the optimal neural status for anticipatorily facilitating perceptual performance, then pre-target intervals with delta phase closer to the mean phase should associate with better subsequent perceptual performance (e.g., Henry and Obleser, 2012). However, our results did not show such an effect. Phase deviation was not significantly correlated with discrimination sensitivity, although there was a trend in the left auditory cortex (Fig. 3E; left auditory cortex: $t(15) = -2.32$, $p_{FDR} = 0.067$; right auditory cortex: $t(15) = -0.07$, $p_{FDR} = 0.970$, one-sample t-tests). Phase deviation was not significantly correlated with RT either (Fig. 3F; left auditory cortex: $t(15) = -0.77$, $p_{FDR} = 0.583$; right auditory cortex: $t(15) = -1.03$, $p_{FDR} = 0.474$, one-sample t-tests).

Together, the analyses replicated previous studies in that the 2 Hz delta oscillations entrained to the rhythmic regularity of the tone sequence, and the ITPC of the delta phase was higher in the Rhythmic than Arrhythmic condition. However, unlike previous studies showing that delta phase is associated with perceptual performance for auditory intensity and timing (Henry and Obleser, 2012; Henry et al., 2014; Arnal et al., 2015; ten Oever et al., 2017; Bauer et al., 2018), the present results suggest that delta phase entrainment is not strongly related to near-threshold pitch perception performance.

IV.5.4 Delta-beta phase-amplitude coupling is modulated by rhythmicity and associates with RT

Given that both delta phase and beta power modulation are affected by rhythmic regularity, we were interested in, first, whether the degree of delta-beta phase-amplitude coupling (indexed by dPACz, where higher dPACz values reflect better delta-beta coupling; see Method: Delta-beta coupling analyses for details) related to the rhythmicity of the input sequence. Results showed that the dPACz was higher in the Rhythmic than Arrhythmic condition in both left ($t(15) = 4.85$, $p_{FDR} = 0.003$, paired t-test) and right ($t(15) = 3.69$, $p_{FDR} = 0.016$, paired t-test) auditory cortices (Fig. 5).

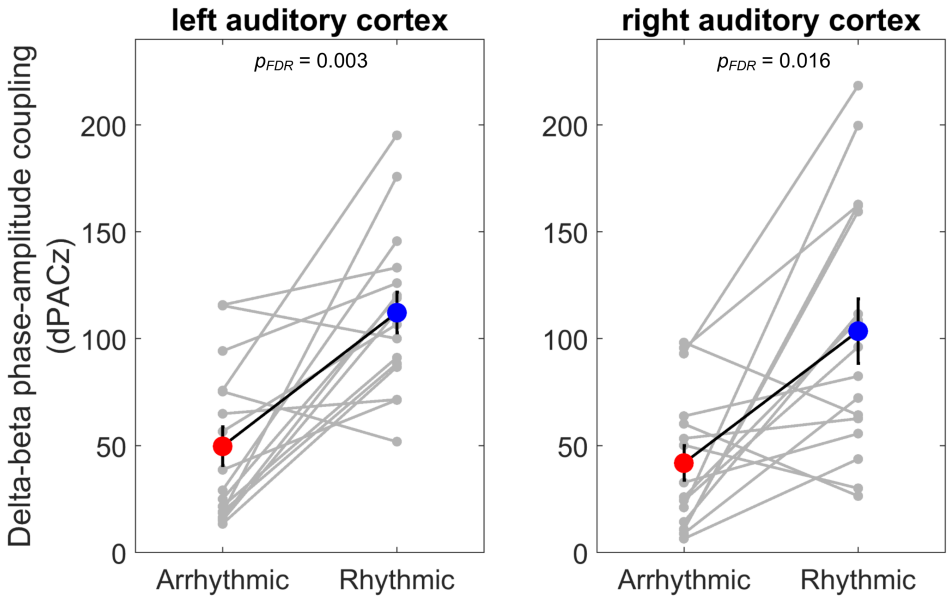

FIGURE IV.5: Pre-target delta-beta phase-amplitude coupling is modulated by rhythmicity. Higher dPACz values represent higher phase-amplitude coupling strength between the phase of low-frequency oscillations (delta: 1–3 Hz) and the power of high-frequency oscillations (beta: 15–25 Hz). The dPACz value of each participant is shown as connected dots, and the mean ± standard error across participants by colored dots. The result showed dPACz is higher in the Rhythmic than Arrhythmic condition in both cortices.

We further investigate whether delta-beta coupling in the pre-target period is anticipatorily associated with subsequent perceptual performance, by correlating dPACz separately with pitch discrimination sensitivity and RT on a trial-by-trial level in the Rhythmic condition (Fig. 3H). The results showed that dPACz and RT are negatively correlated in the right auditory cortex ($t(15) = -3.06$, $p_{FDR} = 0.025$, one-sample t-test), indicating that the higher the coupling strength, the shorter the RTs. The same analyses did not show any significant effects in left auditory cortex ($t(15) = 0.97$, $p_{FDR} = 0.480$, one-sample t-test). This effect was stronger in right than left auditory cortex ($t(15) = 2.55$, $p_{FDR} = 0.046$, paired t-test). For pitch discrimination sensitivity, there were no significant correlations in either left ($t(15) = 0.39$, $p_{FDR} = 0.826$, one-sample t-test) or right ($t(15) = 1.08$, $p_{FDR} = 0.474$, one-sample t-test) auditory cortices (Fig. 3G).

Together, the analyses of delta-beta phase-amplitude coupling showed, first, that coupling strength was higher for Rhythmic than Arrhythmic sequences and, second, that higher coupling strength is anticipatorily associated with shorter RT in a subsequent pitch discrimination judgment on a trial-by-trial level.

IV.6 Discussion

Accurate pitch perception is crucial for identifying objects in the world and perceiving speech and music. We showed that rhythmic regularity facilitates fine pitch discrimination, and revealed how it is implemented in neural entrainment activities. Participants discriminated near-threshold pitch deviations (targets) in contexts where tones were either rhythmically or arrhythmically sequenced. The behavioural results showed rhythmic regularity facilitates both psychophysical pitch discrimination sensitivity and reaction time (RT). The EEG analyses showed that delta (1–3 Hz) phase, beta (15–25 Hz) power, and the degree of delta-beta coupling were all modulated by rhythmicity. Interestingly, we further showed that these neural activities immediately prior to target tones have differential behavioural contributions to pitch discrimination performance at a trial-by-trial level. Specifically, (1) deeper U-shaped beta power modulation predicts higher discrimination sensitivity, (2) higher delta-beta coupling strength predicts shorter RT, but (3) delta phase alone was not

related to either discrimination sensitivity or RT. Note that the covariances among the EEG indexes were partialed out prior to these EEG-behaviour correlational analyses.

Delta phase entrained (phase-locked) to the rhythmicity in the input sequences, consistent with previous studies (Calderone et al., 2014). However, in contrast with beta modulation depth and delta-beta coupling (discussed below), degree of delta entrainment did not associate significantly with perceptual discrimination of pitch or RT on a trial-by-trial basis, although there was a trend for an association between higher pitch discrimination sensitivity and smaller delta phase deviation in the left hemisphere. In contrast, previous studies have found robust associations between delta phase and the perception of timing or intensity deviations (Henry and Obleser, 2012; Henry et al., 2014; Arnal et al., 2015; ten Oever et al., 2017; Bauer et al., 2018). Thus, there might be a distinction in how delta phase entrainment relates to spectral-based (involving frequency and pitch) and temporal/intensity-based (involving onsets, intensity changes, duration and rhythm) perceptual sensitivity, although additional studies need to be done that directly compare associations between delta phase and pitch, timing and intensity. If found, such a distinction would be interesting as, at a more abstract level, pitch is more related to the content or identity of a sounding auditory object, whereas timing and intensity are more related to an object's location in time and space, suggesting that delta phase may relate more to timing than to perceptual facilitation of a sounding object's identity. Note, of course, that these post-hoc speculations are based on null statistical differences and should therefore be taken with caution. It is possible that a correlation between delta phase and perceptual facilitation might be observed with a larger sample size. However, given the same statistical power, beta modulation shape and delta-beta coupling were more strongly associated with pitch discrimination performance than delta phase.

With respect to beta, we found that beta power modulation latencies are affected by rhythmicity, and that U-shaped modulation depth predicts pitch perception. The former is consistent with previous findings on the role of beta in temporal prediction (Snyder and Large, 2005; Iversen et al., 2009; Fujioka et al. 2009, 2015, 2012; Cirelli et al., 2014; Morillon and Baillet, 2017). Furthermore, we found that the depth of U-shaped beta power modulation predicts pitch discrimination sensitivity on a trial-by-trial basis. Thus, our data indicate that beta power modulation shape affects sensory predictions,

and anticipatorily facilitates perceptual performance. A few previous studies hinted at the role of beta in prediction of pitch. Specifically, beta power was shown to increase after unexpected pitch changes (Chang et al., 2016; Franken et al., 2018), to reflect the magnitude of pitch prediction updating (Sedley et al., 2016), and to relate to smaller event-related potential pitch prediction error responses (Chang et al., 2018). However, the lack of behavioural tasks in these studies makes it challenging to link the neural findings with perception. Our finding in the current study that entrained beta power modulation shape prior to a pitch change affects pitch perception adds to previous literature indicating that non-entrained beta power reflects information integration leading to improved perceptual performance across many domains (e.g., Arnal et al., 2015; Herrmann et al., 2016; Kayser et al., 2016; Florin et al., 2017; Pefkou et al., 2017; Spitzer and Haegens, 2017). Note that we found no association between beta desynchronization latency and either pitch discrimination sensitivity or RT on a trial-by-trial basis. It is unclear why rhythmicity affected beta desynchronization latency, but beta modulation depth predicted the discrimination sensitivity at the trial-by-trial level under the Rhythmic condition. Nevertheless, considering that rhythmicity of the sensory input modulates whether the temporal attending mode is rhythmic or not (Nobre and van Ede, 2018), a possible explanation is that the beta desynchronization latency effect reflects the difference between distinct attending modes and beta modulation depth reflects the trial-by-trial neural mechanism of facilitating pitch discrimination under the rhythmic attending mode. In sum, data across studies suggest that for audition, both delta phase and beta power modulation latencies might relate to the prediction of rhythmic temporal regularity; but that delta phase may proactively facilitate perception in the temporal and intensity domains, whereas the U-shaped beta power modulation depth may relate to the spectral domain, including prediction of pitch change and the quality of pitch perception.

Behaviourally, despite the fact that pitch is a fundamental aspect of the perception of sound and critical for both speech and musical processing, we provide novel evidence that rhythmicity improves near-threshold pitch sensitivity. It adds to previous literature indicating that rhythmicity affects sensitivity to deviations in timing and intensity and detecting signals in noise (e.g., Henry and Herrmann, 2014; Hickok et al., 2015). Evidence that rhythmic regularity facilitates detection of large (supra-threshold) pitch changes is inconsistent (Jones et al., 2002; Morillon et al., 2016, but see Bauer et al., 2015), so further

research is needed to determine whether near-threshold and supra-threshold pitch discrimination are differentially affected by rhythmic context. In the absence of a predictable rhythmic context, one study found that predictability of a timing delay between isolated cue-target pairs was associated with decreased RT in a pitch discrimination task (Herbst and Obleser, 2017). However, this study did not show any effects on accuracy, and thus it cannot distinguish whether the effects of shortened RT were caused by facilitated perception or simply better motor preparation due to the predictable intervals. The current psychophysical results further showed that perceptual pitch sensitivity was improved in a rhythmic context. Together our study and that of Herbst and Obleser (2017) suggest that predictable temporal structure in general, whether it arises from a rhythmic context or memory for a time interval, facilitates fine pitch discrimination.

Finally, the present results show that delta-beta phase-amplitude coupling predicts RT but not sensitivity of pitch discrimination, suggesting that cross-frequency coupling reflects auditory-motor interaction. Phase-amplitude coupling is thought to result when the excitatory or inhibitory phase of low-frequency oscillations modulates the power in high-frequency oscillations (Hyafil et al., 2015b), such that processing is optimized for rhythmic input (Lakatos et al., 2005; Schroeder and Lakatos, 2009). Only a few studies have reported that delta-beta coupling affects perceptual performance. For example, delta-beta coupling was reported to associate with the accuracy of detecting an auditory temporal delay (Arnal et al., 2015) and comprehending speech (Keitel et al., 2018). In a visual task, delta-beta coupling in primary motor cortex associated with enhanced visual performance (Saleh et al., 2010). In auditory tasks, delta-beta coupling reflects the communication between motor and auditory cortical regions and might modulate perceptual and behavioural selection (Morillon and Baillet, 2017). Our finding is consistent with this idea that delta-beta coupling reflects the efficiency of auditory-motor communication, as shorter RTs would indicate faster information transfer between auditory and motor regions, resulting in faster response selection.

It is an open question as to why we observed that U-shaped beta power modulation depth correlated with pitch discrimination at the left auditory cortex but delta-beta coupling with RT at right auditory cortex. We did not have an a priori hypothesis about lateralization, as the vast majority of previous studies did not examine hemispheric differences. To our knowledge,

there are only two previous studies using MEG that looked at hemispheric lateralization. Our lateralization pattern is in line with these previous studies in that they found that the beta effect is localized to the left hemisphere and the delta effect is localized to the right hemisphere (Morillon and Baillet, 2017; Tal et al., 2017). Together, although further investigation is needed, these lateralization findings suggest that left and right auditory cortices associate with different neural entrainment mechanisms and perceptual performances.

The current study focused on neural activities generated from bilateral auditory cortices because they are regarded to be the primary regions for processing pitch in the brain (McDermott, 2018) and are typically reported to be an important source of neural entrainment activities in response to auditory sequences (Fujioka et al., 2012; Morillon and Baillet, 2017). Indeed, most studies investigating auditory neural entrainment activities focus on the neural signal from auditory cortex, either through direct electrophysiological recording, source modeling, or analysis of surface electrodes that predominantly record activity from auditory cortex (e.g., Lakatos et al., 2013; Henry et al., 2014; Morillon and Baillet, 2017; ten Oever et al., 2017). However, neural entrainment activities in response to auditory sequences can also be observed in other cortical regions, including sensorimotor, premotor, supplementary motor and inferior-frontal regions (Besle et al., 2011; Fujioka et al., 2012; Morillon and Baillet, 2017). Examining the activities from those regions is beyond the scope of the current study, given the lack of individual structural scans or sufficient spatial specificity of EEG to reliably extract the neural activities from motor and frontal regions. Nevertheless, we argue that the oscillatory activities from auditory cortex should be most relevant for studying rhythmic facilitation of pitch discrimination. Further studies are needed to examine the perceptual functions of neural entrainment activities beyond auditory cortex.

Many behavioural and neural studies suggest that auditory and motor systems cooperate in processing rhythm (e.g., Phillips-Silver and Trainor, 2005; Chen et al., 2008; Grahn, 2012; Merchant et al., 2015), and beta oscillation is commonly hypothesized as the auditory-motor communication channel (Fujioka et al., 2012; Morillon and Baillet, 2017). Our data suggest delta-beta coupling is also critical for auditory-motor interactions. It would be interesting for future studies to examine delta-beta coupling in populations with motor deficits, such as Parkinson's patients, who are known to have reduced beta power entrainment to auditory rhythms (te Woerd et al., 2018). Indeed,

we have hypothesized that children with developmental coordination disorder will show reduced beta and delta-beta entrainment (Trainor et al., 2018), and others have proposed that interpersonal auditory-motor synchronization (e.g., joint music performance) may involve beta coupling between brains (Novembre et al., 2017).

It is important to note that our Rhythmic condition used the same repetitive isochronous intervals while the Arrhythmic condition used random intervals, and thus the perceptual facilitation effect could be caused by either rhythmicity or by memorizing the predictable fixed single interval or both (Breska and Deouell, 2017; Herbst and Obleser, 2017). Neurally, it has been shown that a predictable fixed single interval without rhythmicity can increase low-frequency phase clustering, so this index might reflect non-oscillatory neural ramping activity (e.g., contingent negative variation) rather than neural oscillation (Breska and Deouell, 2017, but see Obleser et al., 2017). For the current study, we did not aim to eliminate the potential effect of a predictable fixed single interval on perceptual performance, as a fixed interval is inevitably embedded in any isochronous rhythmic sequences. The rhythmic/arrhythmic manipulation was an approach for examining temporal prediction, whether of a rhythmic or fixed interval nature. The important behavioural contribution of the current study is in demonstrating that rhythmic temporal predictability facilitates pitch perception in audition. Future studies are needed to examine whether the effects reported in the current study can be generalized to other forms of temporal predictability, such as cue-target association, hazard rate and repeated intervals (e.g., Breska and Deouell, 2017; Herbst and Obleser, 2017; Nobre and van Ede, 2018). As for delta phase, the current time-frequency neural measurement cannot dissociate possible ramping activities from oscillatory activities. It is possible that the null correlational results between delta phase and behavioural pitch performance were due to confounding ramping neural activities; however, many previous studies have reported an association between delta phase and rhythmic perceptual facilitation, although it remains unclear whether ramping neural activities played a role. In general, separating oscillatory and non-oscillatory neural activities in low-frequency neural oscillations is challenging and needs to be addressed in further careful studies (Doelling et al., 2019).

In a broader theoretical context, the composition of two coupled neural entrainment activities with dissociable functions is consistent with dynamic

attending theory. It posits two components: oscillatory temporal expectation and temporal attention (Large and Jones, 1999) in an auditory-motor network (Large et al., 2015). The alignment between these two components enables attention to be anticipatorily deployed for tracking rhythmic sensory sequences and enhancing attention, and therefore perceptual sensitivity, at critical points in time. Our post-hoc speculation is that delta phase might reflect oscillatory temporal expectation and beta power might reflect temporal attention. In particular, our finding that delta phase is affected by rhythmicity but is not related to perceptual pitch enhancement suggests that the primary role of delta oscillations might concern temporal structure prediction rather than facilitating fine spectral perception (i.e., pitch discrimination). Our finding that depth of U-shaped beta power modulation predicts pitch discrimination sensitivity might be explained as beta reflecting enhanced auditory attention at particular time points, where perceptual gain is increased. This speculative framework is consistent with previous studies showing that delta phase entrainment (temporal expectation) is critical for time and intensity tasks. This idea is also consistent with converging evidence that the phase of low-frequency oscillation reflects top-down sensory selection among multiple sensory streams, that the power of high-frequency oscillation implements endogenous perceptual processes or attentional gain, and that the alignment between these two optimizes perceptual performance (Arnal and Giraud, 2012; Henry and Herrmann, 2014; Lakatos et al., 2013, 2016; Morillon and Baillet, 2017; Schroeder and Lakatos, 2009; Saleh et al., 2010). Additional evidence comes from a study that mathematically modeled MEG recordings in the context of dynamic attending theory, to show that delta oscillations link to temporal expectations (Herrmann et al., 2016). The current study extends this model to beta oscillations, but only shows a single-dissociation with correlational evidence; the differential roles of delta, beta and delta-beta coupling suggested by the present data would be strengthened by further studies showing causal evidence of a triple-dissociation.

The findings of the current study are relevant to speech perception. First, the proposal that the theta phase (4–8 Hz) tracks speech envelope, and nested gamma power (∼25–35 Hz) aligns neuronal excitability to acoustic structure for perceiving speech (Giraud and Poeppel, 2012) is similar to our proposal that delta phase reflects temporal prediction and U-shaped beta power modulation depth associates with temporal attention. Second, rhythm facilitates

both speech perception (Haegens and Zion Golumbic, 2018) and pitch discrimination (present findings), suggesting beta modulation reflects the "content" of auditory perception. This relationship might also reflect that frequency modulation is crucial for speech recognition (Zeng et al., 2005).

IV.7 Conclusion

We present the novel finding that the temporal regularity of rhythmic sound sequences facilitates perceptual processing of pitch, and that neural oscillatory entrainment activities from auditory cortex, including delta phase, beta power modulation, and delta-beta coupling, are all modulated by rhythmicity. At the same time, trial-by-trial analyses showed that U-shaped beta power modulation depth predicts discrimination sensitivity whereas delta-beta coupling strength predicts reaction time. Future studies should investigate whether temporal predictability in other contexts such as cue-target association, hazard rate, and repeated intervals also enhance pitch perception. With respect to a neural instantiation of dynamic attending theory, our findings suggest a speculative interpretation that delta phase maps onto oscillatory temporal expectations, beta power onto temporal attention, and delta-beta coupling onto the efficiency of auditory-motor communication. Understanding how these neural oscillations work together is critical for uncovering the auditory-motor network and the neural basis of the perception of dynamic auditory inputs such as speech and music.

IV.8 Supplementary Material

IV.8.1 Shared covariances among EEG indexes

Four EEG indexes in the current study (delta phase deviation, quadratic coefficient of beta power modulation shape, temporal position of beta power modulation vertex, and delta-beta dPACz) might potentially share high covariances, as previous studies reported these indexes are likely to be coupled when averaged data for each condition is examined (e.g., Arnal et al. 2015; Keitel et al. 2018; Saleh et al. 2010). Therefore, to assess their covariances

at the single-trial level, we performed pairwise Pearson correlation analyses among EEG indexes at the single-trial level for each participant.

The pairwise correlation coefficients are listed in Table S1. In general, the covariances are low. The highest unsigned mean coefficient across participants was only 0.181, and thus the R2 was only 3.3%, which is very small. Therefore, we argue that these EEG indexes have low covariances and are largely mutual independent at the single-trial level. They are therefore likely to reflect different neural and perceptual functions in the current study.

IV.8.2 Rhythmicity moderating EEG-behaviour associations

The trial-by-trial EEG-behaviour correlational analyses in the Rhythmic condition showed that beta modulation depth in left auditory cortex predicts discrimination sensitivity and delta-beta coupling in right auditory cortex predicts RT. An extended question is whether rhythmicity moderates this EEG-behaviour relationship. Therefore, we calculated the same trial-by-trial EEG-behaviour correlations for each participant in the Arrhythmic condition as we did in the Rhythmic condition, and then we performed a post-hoc comparison of the z-normalized correlational coefficients across participants between conditions. We performed this test for the significant EEG-behaviour correlations reported in the Fig. 3. Results showed that the association between beta modulation depth and discrimination sensitivity in left auditory cortex was stronger in the Rhythmic than Arrhythmic condition ($t(15) = 2.52$, $p = 0.023$, Fig. S3A). However, the association between delta-beta coupling and RT in right auditory cortex was not significantly different between conditions ($t(15) = -1.88$, $p = 0.080$, Fig. S3B).

TABLE IV.S1: Pairwise Pearson correlation coefficients among EEG indexes

		Delta phase deviation		Beta power modulation				Delta-beta dPACz	
				quadratic coefficient		vertex time point			
		L	R	L	R	L	R	L	R
Delta phase deviation	L	–	–	–	–	–	–	–	–
	R	0.149 ± 0.044	–	–	–	–	–	–	–
Beta power modulation quadratic coefficient	L	-0.010 ± 0.014	-0.016 ± 0.013	–	–	–	–	–	–
	R	-0.008 ± 0.012	0.009 ± 0.015	0.103 ± 0.020	–	–	–	–	–
Beta power modulation vertex time point	L	-0.014 ± 0.015	-0.028 ± 0.017	0.030 ± 0.040	0.027 ± 0.015	–	–	–	–
	R	-0.009 ± 0.015	-0.028 ± 0.023	-0.011 ± 0.022	0.024 ± 0.023	0.050 ± 0.020	–	–	–
Delta-beta dPACz	L	-0.071 ± 0.018	-0.005 ± 0.018	0.148 ± 0.017	0.014 ± 0.016	-0.008 ± 0.016	-0.028 ± 0.011	–	–
	R	-0.039 ± 0.016	-0.028 ± 0.016	0.031 ± 0.015	0.181 ± 0.021	0.019 ± 0.019	0.016 ± 0.026	0.026 ± 0.021	–

Note: L for left auditory cortex; R for right auditory cortex. The values are the mean $\pm$ standard error of correlation coefficients across participants.

IV.9 Acknowledgements

This research was supported by grants to LJT from the Canadian Institutes of Health Research (grant numbers MOP 115043, 153130) and by a Vanier Canada Graduate Scholarship to AC. We thank Dave Thompson for technical assistance, Michael Slugocki for the suggestions on psychophysical analyses, and Katherine Clayworth and Jennifer Chan for assisting with data collection. The authors declare no competing financial interests. The datasets and the code generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

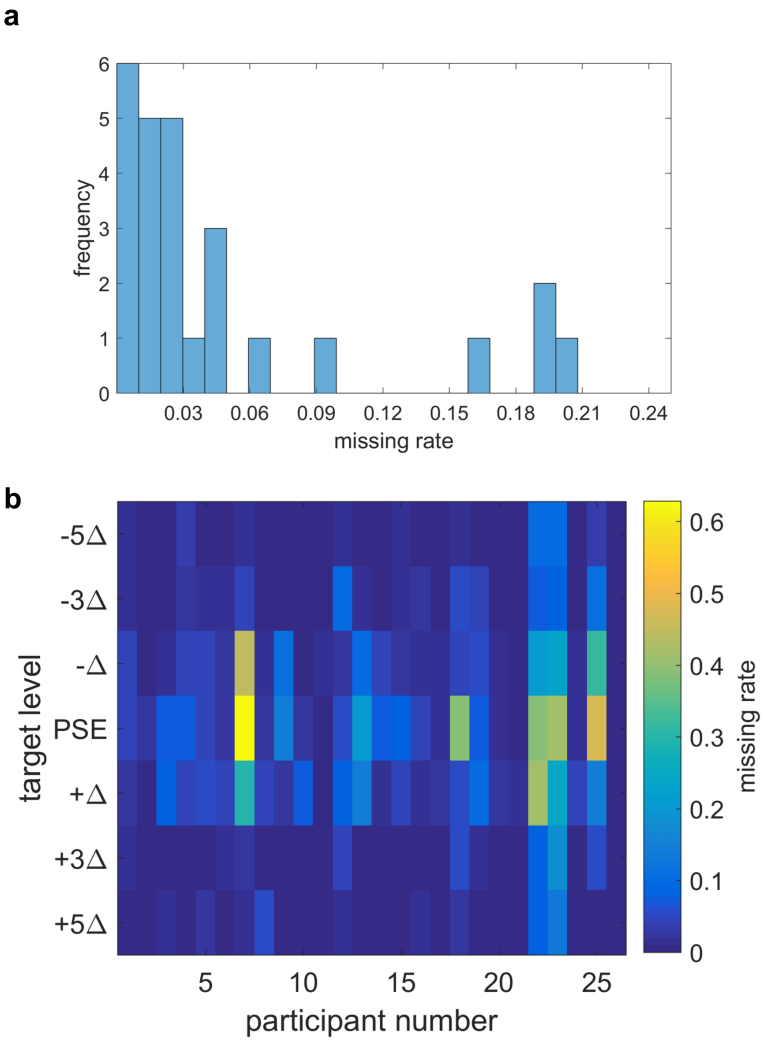

FIGURE IV.S1: Distribution of missing rates. (A) The histogram of the missing rates across participants. Ten out of twenty-six participants missed responding for more than 3% of the target tones, and thus they were excluded for further analyses. (B) The distribution of missing rates across target levels of each participant. Participants tended to miss responses when the target levels were difficult (around the PSE).

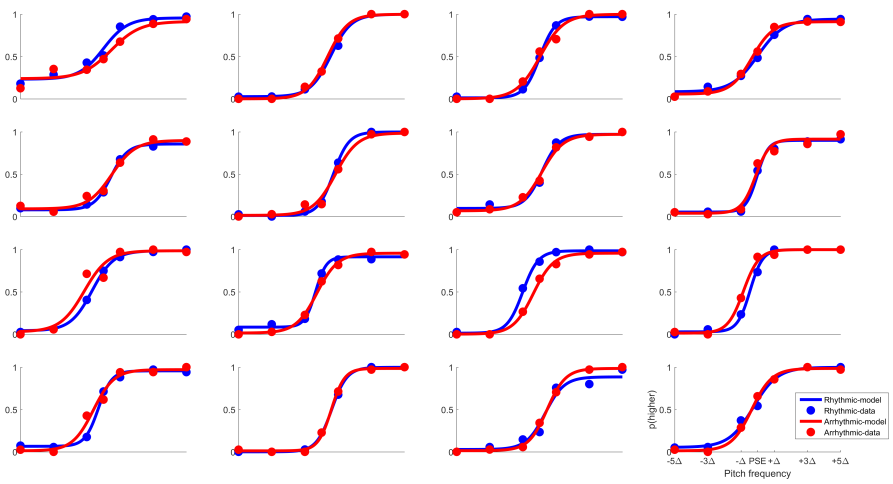

FIGURE IV.S2: Individual perceptual judgment performances and the fitted psychometric functions.

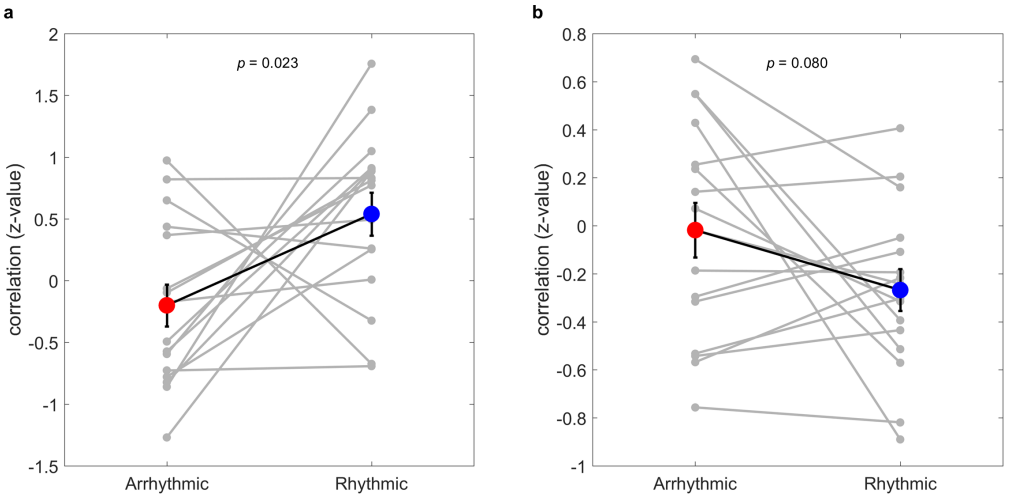

FIGURE IV.S3: Rhythmicity modulating EEG-behaviour associations. (A) The association between beta modulation depth and discrimination sensitivity in left auditory cortex was stronger in the Rhythmic than Arrhythmic condition. (B) The association between delta-beta coupling and RT in right auditory cortex was not significantly different between conditions. Each connected dot represents one participant's data, and the colored dots and error bars mean ± standard error across participants.

Chapter V

General discussion

Temporal regularity is common among the acoustic waveforms of our everyday life, such as the beat of music and the quasi-isochronous syllable onsets of speech. Rhythmic regularity simplifies the complexity and the burden of perceiving fleeting auditory information because this regularity enables sensory systems to generate temporal predictions and thus proactively optimize perceptual processing at predicted time points (Large & Jones, 1999; Haegens & Zion Golumbic 2018; Nobre & van Ede 2018). Indeed, previous perceptual studies showed that humans are able to track rhythm, and that rhythmic temporal regularity enhances perception, as reflected in increased discrimination sensitivity and decreased perceptual thresholds (e.g., Cravo et al., 2013; Henry and Herrmann 2014; Rohenkohl et al., 2012; ten Oever et al., 2017). According to dynamic attending theory, the mechanisms for rhythm tracking include the alignment between an internal oscillator (representing the periodicity of temporal expectation) and a regular attentional pulse in time (Large & Jones, 1999). This theory also explains the rhythmic facilitation effect on perception as the internal oscillator periodically allocates the attentional pulse at particular time points, which results in enhanced perceptual performance.

At the neural level, oscillatory entrainment activities, including both low-frequency phase, high-frequency power, and their phase-amplitude cross-frequency coupling, have been associated with rhythmic temporal prediction and perceptual facilitation. Electrophysiological measures, including local field potentials (LFP), electroencephalography (EEG), and magnetoencephalography (MEG) in human and non-human animals, have shown that the rhythmic temporal regularity in acoustic waveforms entrains neural oscillatory activities

(see Arnal & Giraud, 2012 for a review). Furthermore, perceptual performance can be predicted by these neural activities even prior to the onset of a target signal (Haegens & Zion Golumbic 2018). This suggests that neural entrainment activities set up rhythmic temporal prediction and thus proactively optimize perceptual processes.

Given this background, the current thesis advances our understanding of auditory sensory prediction and neural entrainment by addressing several critical but previously unsolved issues: (1) Why do humans have multiple neural oscillation entrainment mechanisms for tracking auditory rhythm? Do these mechanisms have different perceptual functions? (2) How do temporal and spectral aspects of sound processing interact and relate to neural oscillations? Does temporal regularity (isochronous rhythm) enhance prediction and/or perception in the orthogonal spectral domain, given that spectral information is essential for audition (e.g., speech, music, object identification). Time and frequency are the two fundamental dimensions of audition necessary for tracking fleeting auditory information, so understanding how they work together is essential for understanding auditory perception.

V.1 Unique contributions and limitations of each chapter

V.1.1 Chapter II

Previous studies on beta power neural oscillation entrainment showed that it reflects the tempo of isochronous sequences composed of identical tones of the same pitch (e.g., Fujioka et al., 2012, 2015). However, it remained unclear whether beta power fluctuations reflect only temporal prediction, regardless of the spectral content. If yes, then this entrainment activity should not be disrupted by a tone with unexpected pitch (deviant stimulus) as long as it is presented at the expected temporal position. Contrary to this, however, this study found that beta power entrainment activity is disrupted by unexpected spectral (pitch) information, even when it is presented at a temporally expected rhythmic position.

The novel contribution of this study is in demonstrating that beta power entrainment activity reflects both rhythmic temporal prediction and the prediction violation related to unexpected spectral information. Typically, predictive timing (when) and predictive coding (what) have been studied separately (Arnal & Giraud, 2012), so it was unclear until the current study whether neural oscillatory activity in one frequency band (beta) reflected both of these perceptual processes.

However, this study only measured neural activity following a spectral prediction error (i.e., following the presentation of a tone with unexpected pitch), but it did not measure predictive neural activity, that is, activity occurring prior to the onset of an unexpected pitch. Thus, the evidence that beta oscillations reflect predictive processing in the spectral domain was indirect in this study. Because a prediction error was measured, it was assumed that there must have been a prediction, but this needed to be tested directly. Also, the deviant occurrence rate of the oddball paradigm was confounded with predictability, as this study manipulated predictability by contrasting 10% versus 20% deviance occurrence rates, rather than manipulating predictability while keeping occurrence rate constant. These issues were addressed in Chapter III. Finally, this study did not measure the behavioural consequences of predictions, and thus it was hard to interpret the perceptual functions of the observed oscillatory neural activities. This issue was addressed in Chapter IV.

V.1.2 Chapter III

This study showed that auditory beta power entrainment is modulated by the predictability of a pitch change. Specifically, isochronous auditory tone sequences were presented with infrequent pitch changes at either predictable (every fifth tone) or unpredictable (random, but with the same overall deviance rate as in the predictable condition) positions. The results showed that beta was better entrained prior to a predictable than unpredictable pitch change. Furthermore, the depth of beta power entrainment activity preceding a pitch change correlated with the prediction error neural response (i.e., P3a) following a pitch change at the trial-by-trial level. This suggests that beta power entrainment depth reflects spectral prediction. In contrast to most of the previous neural scientific studies on either predictive timing (predicting when) or predictive coding (predicting what), the current study showed that beta power

entrainment activity can reflect both predictive timing and predictive coding processes.

This study provides stronger evidence than Chapter II in suggesting that beta power entrainment prior to the onset of a tone reflects the prediction of both what (frequency) and when (time) events are expected. The findings are consistent with the fact that, although what and when in audition are physically orthogonal, the information from these two dimensions often covaries. In the case of music, for example, inter-note intervals tend to be longer between two phrases than within a phrase (Kragness et al., 2016; Palmer, 1989). Also, the tempo of a typical music performance is not completely steady, and musicians often make deliberate timing deviations for expressive purposes (Nakata & Trainor, 2015; Rankin et al., 2009). In the perceptual domain, pitch information can also influence time perception. For example, an unpredictable infrequent deviant stimulus will be perceived as longer than a frequent standard stimulus (e.g., Tse et al., 2004; McAuley & Fromboluti, 2014). In sum, the findings of Chapter III suggest the possibility that the sensory predictions in the what and when domains might have interactive influences on each other.

This study extended the results from the study of Chapter II by measuring the predictive neural activity prior to a pitch change and controlling the occurrence rate while manipulating predictability. However, an alternative interpretation for the results of the studies in both Chapters II and III is that beta power entrainment reflects the prediction of which tone in the sequential position is going to be changed to a different pitch, which can still be regarded as a type of temporal prediction (i.e., when the change will happen). To fully show that beta power entrainment reflects spectral prediction, an experiment should investigate not only whether beta power reflects prediction of an upcoming a pitch change, but also how well the pitch is perceived. For this, it is necessary to measure the perceptual consequences of knowing when a pitch change is expected. This was done in Chapter IV.

V.1.3 Chapter IV

This study made novel contributions to understanding both perceptual consequences and neural mechanisms of rhythm tracking. Participants were asked to report whether pitch changes at predictable or unpredictable times in a

tone sequence were higher or lower than standard pitches. Perceptually, auditory temporal regularity (i.e., rhythm) facilitated pitch perception, with lower thresholds for detection of pitch changes in isochronous than non-isochronous sequences. Neurally, low and high frequency neural entrainment and their coupling activities associated with different perceptual functions. Specifically, beta power entrainment associated with discrimination sensitivity, and delta-beta phase-power coupling associated with reaction speed. Mapping these findings to dynamic attending theory (Large & Jones, 1999), I argue that the phase of low-frequency oscillation (i.e., delta band) reflects temporal expectation, the power of high-frequency oscillation (i.e., beta band) reflects temporal attention, and phase-amplitude cross-frequency coupling reflects the alignment of these two perceptual mechanisms and is associated with auditory-motor communication. This study is the first to map multiple neural oscillation entrainment activities onto the different perceptual functions of dynamic attending theory. The distinct but coordinated perceptual functions of the different auditory neural oscillatory entrainment activities revealed in this study contributes to our understanding of how auditory rhythm tracking enables processing streams of phonemes in speech and chord progressions in music.

This study makes contributions beyond those of the two previous studies. It measured the perceptual consequences of the neural entrainment activities and revealed different functions of high and low frequency oscillations. However, this study has limitations. First, the EEG-behaviour relationships were based on correlational evidence with a single dissociation. Future experiments are needed to independently modulate each EEG-behaviour association to fully reveal the distinct perceptual functions of these EEG activities. Second, while the neural evidence was consistent with dynamic attending theory, much more work needs to be done to directly prove or disprove this theory. Specifically, temporal expectation and temporal attention would need to be orthogonally manipulated in order to determine whether they work as predicted by dynamic attending theory.

V.2 Theoretical contributions and future directions

This thesis has made two major theoretical contributions. First, it revealed that neural beta power entrainment is involved in both temporal and spectral predictions in audition. Previous studies showed that beta power fluctuations entrain to the tempo of presented isochronous auditory tone sequences, and thus reflect temporal prediction (e.g., Fujioka et al., 2012). The present thesis showed that beta power, both before and after a pitch change, can be modulated by the spectral predictability of tones and that temporal regularity enhances spectral processing. A question for future research is whether the predictability of the spectral information can also affect the temporal prediction at the perceptual level. I hypothesize that unpredictable pitch changes in an isochronous auditory sequence will distort the temporal prediction of the upcoming tone. Two remotely relevant studies have shown consistent evidence that deviation in the what (content) domain can influence temporal perception. Specifically, the presentation duration (onset to offset) of an infrequent visual or auditory deviant stimulus will be perceived to be longer than the frequent standard stimulus (McAuley & Fromboluti, 2014; Tse et al., 2004).

The second major theoretical contribution of the present thesis is showing that low-frequency phase, high-frequency power, and their coupled oscillation entrainment activities have different perceptual functions, which can be mapped onto distinct components of dynamic attending theory. To the best of our knowledge, only a few studies have examined multiple frequency oscillation entrainment activities at the same time (e.g., Arnal et al., 2015; Keitel et al., 2018), and only a couple of studies have distinguished perceptual functions of different entrainment activities (e.g., Morillon et al., 2017). Regarding linking dynamic attending theory to entrainment activities, most previous studies focused on low-frequency phase entrainment (e.g., Henry & Herrmann et al., 2014; Herrmann et al., 2016), and discussions on high-frequency power and the cross-frequency phase-amplitude coupling entrainment activities were rare. The current thesis showed distinct perceptual functions for each entrainment activity, that are consistent with multiple components of dynamic attending theory. Two critical future research goals are as follows. The first goal is to examine whether each entrainment activity will be individually modulated by temporal expectation and temporal attention, two distinct components of

the dynamic attending model. Such a study could provide stronger double-dissociative evidence for distinct perceptual functions. The second future goal is to examine the behaviours of oscillation entrainment activities when the temporal regularity of the stimuli is disturbed (e.g., phase shift, change in tempo), and whether the mathematical model of dynamic attending theory can predict these activities (Large & Jones, 1999). Such a study could provide a better assessment of how well dynamic attending theory can explain neural entrainment activities.

Another future direction is to extend the current findings to situations of predictable but aperiodic (non-isochronous) temporal regularity. The current studies, and most of the existing literature, investigated temporal prediction and neural entrainment by using isochronous rhythms. However, isochronous rhythms oversimplify the range of regularity possibilities in the real-world. The AM rates of sound in the real-world (e.g., speech and music) often speed up and slow down and are unlikely to be perfectly isochronous at a specific frequency. Nevertheless, humans are still able to track such acoustic signals (Rimmele et al., 2018). Therefore, it is necessary to investigate how humans track aperiodic but predictable auditory signals, such as accelerating and decelerating non-isochronous rhythms. To date, only a few behavioural studies have investigated humans' sensorimotor mechanisms for tracking predictable aperiodic sound sequences. For example, participants can form temporal predictions for accelerating or decelerating auditory sequences (Cope, Grube, & Griffiths, 2012), and participants can synchronize their finger tapping with auditory sequences containing predictable tempo changes (Loehr, Large, & Palmer, 2011). Similar to isochronous (predictable periodic) regularity, tone sequences with regular temporal change (predictable aperiodic) can also facilitate perceptual detection performance, compared to random (aperiodic unpredictable) sequences (Morillon et al., 2016). Future studies are needed to reveal the underlying neural oscillation mechanisms of this important ability.

V.3 Neural signal processing challenges

The studies of the current thesis faced a number of neural signal processing challenges. First, the current thesis used dipole models as spatial filters to emphasize the EEG signals generated from auditory cortex while attenuating the

signals from other areas, but this approach cannot completely isolate auditory cortical activity. Second, neural entrainment activities from areas beyond auditory cortex were not examined. Although the auditory cortex is arguably the most dominant and important region of neural entrainment activity for the questions addressed in this thesis (making our approach hypothesis-driven), other brain regions (such as motor and frontal areas) also generate entrainment activities (e.g., Besle et al., 2011; Morillon et al., 2017). However, this approach could not reliably investigate entrainment activities beyond the auditory cortex as spatial specificity in EEG is limited and due to the lack of individual MRI scans. Therefore, it remains unclear whether the features of neural entrainment showed in this study are specific to the auditory cortex, whether they generalize to other brain regions, and how oscillations in different regions may interact. Further studies with individual MRI scans and using MEG, which has higher spatial specificity than EEG, could potentially lead to a better understanding of entrainment activities across brain networks.

Third, a serious confounding factor for our study and others in the field is the influence of evoked neural responses to sound on measurements of low-frequency neural entrainment activity. Many neural entrainment studies are interested in understanding the proactive and predictive feature of the neural oscillations in the low-frequency range (1-8 Hz). However, evoked neural activities that occur in response to a sound (i.e., event-related potentials or ERPs), occur at the same time as entrainment activities, and typically also have spectral power distributed in the low-frequency range ($<$ 15 Hz) (Cohen, 2014a). Thus, the time-frequency analyses for extracting low-frequency neural oscillations related to entrainment and proactive prediction are difficult to separate from the reactive evoked neural activities (Novembre & Iannetti, 2018), and how much of the low-frequency neural activities measured are actually contributed by neural oscillations rather than ERPs remains unknown. This concern is relevant to the delta band findings of the current thesis, but is of minimal concern for the main beta band findings of the current thesis, because the range of beta oscillation (15-25 Hz) is higher than typical ERP activities, and the averaged evoked activities were removed from our data prior to analyzing induced beta oscillations.

Fourth, several time-frequency signal processing concerns have been raised recently in the literature. The Fourier-based time-frequency transformation (including wavelet transformations used in the current studies) assumes that

the waveforms are sinusoidal and stationary (time invariant). These assumptions are not always met, because EEG contains non-sinusoidal neural oscillations, nonstationary oscillatory bursts, and even aperiodic 1/f pink noise. Fourier-based time-frequency transformation cannot reliably distinguish or measure these signals (Cole & Voytek, 2017; Jones, 2016; van Ede et al., 2018). These concerns are relevant to both the current studies and a very large portion of the existing neural oscillation literature. Although the current thesis does not aim to solve time-frequency signal processing concerns, several advanced signal processing approaches are currently being developed to deal with this issue. For example, a cycle-by-cycle analysis approach has been proposed to analyze the amplitude, period, and waveform symmetry of nonsinusoidal neural oscillations, and it can also handle nonstationary oscillatory bursts (Cole & Voytek, 2018). Another approach is to eliminate aperiodic pink noise by parameterizing the neural power spectrum (Haller et al., 2018). Traditionally, neural oscillatory power has been examined within frequency bands, and has assumed that pink noise was equivalent across individuals. But this approach ignores the influence of individual differences in pink noise, which might be confounded with the estimations of the oscillatory power spectrum. This parameterizing approach simultaneously estimates the pink noise spectrum and the neural oscillatory power to separate these activities. Furthermore, to deal with the time-variant frequency of oscillations, it has been proposed to use the Hilbert transformation to obtain the phase-angle time series of the time-domain signal, and then use the first order derivative of the phase-angle time series to estimate the oscillatory frequency fluctuations (Cohen, 2014b). This approach is based on the idea that fluctuation of oscillatory frequency will result in changes in the speed of the phase time series. In sum, although these new approaches are still under development, they will likely improve the analyses of neural oscillations and result in more reliable findings in the future.

Last but not least, another limitation of the current thesis is in establishing a causal relationship between the neural oscillation activities and the perceptual effects. The studies of the current thesis, as well as most modern neuroimaging studies, cannot tell whether any behavioural effects (e.g., perceptual sensitivity) of the experimental manipulations were causally mediated by the neural activities measured. Modern cognitive neuroscience often assumes a hierarchical causal relationship in that the experimental manipulation affects the neural activities, and the neural activities further determine the behavioural outcomes

(e.g., perception). A causal relationship such as this can be supported by studies on patients with focal lesions or dysfunctions, but this assumption often is not examined in neuroimaging studies on healthy participants. Because the measured behavioural outcomes and the neural activities merely statistically covary across the experimental conditions, they should be treated as parallel dependent variables at the same hierarchical level rather than at different hierarchical levels. Therefore, any causal interpretation between the behavioural effects and neural activities should be taken with caution. Another related concern is that hypothesis-driven neuroimaging studies often greatly reduce and simplify data complexity, and then only neural activities brought to the spotlight are examined. This approach enables us to analyze the data within a theoretically interpretable framework, and it also greatly reduces the multiple comparison concern of examining highly complex data. However, this approach very likely ignores other neural activities or more complicated (high dimensional) activities that might be the "true cause" of the behaviours of interest. Although it is a very difficult challenge for cognitive neuroscience, integrating multiple research approaches can be a step forward. For example, transcranial magnetic or direct current stimulation methods have been used to directly influence neural activities and enable researchers to examine causal relationship between neural activities and behaviours. A few studies showed that transcranial current stimulation can causally affect neural entrainment activities – resulting in modulating speech perception – thus demonstrating the causal role of neural entrainment on perception (Riecke et al., 2018; Zoefel et al., 2018). Another data-driven approach is to employ machine learning and big data methods to examine high-dimensional neural activities, which are too complicated for most hypothesis-driven approaches. Note that using machine learning approaches does not imply that future cognitive neuroscience research has to be data-driven and hypothesis-free; conversely, these approaches may play a role in liberating researchers from existing theories that sometimes (perhaps often) corner and limit our scope. Ideally, they will facilitate researchers in forming more thorough, precise and testable hypotheses and theories.

V.4 Potential clinical implications

The findings of this thesis can potentially benefit populations with motor disorders, such as Parkinson's disease (PD) and developmental coordination disorder (DCD). Patients with PD feature deficits in gait or walking, neural degeneration in the basal ganglia, and neural timing deficits. Beyond motor deficits, they also have deficits in perceiving isochronous and beat-based rhythms (e.g., Grahn & Brett, 2009; Teki et al., 2012) and in forming rhythmic temporal predictions (Breska & Ivry, 2018). MEG studies further showed that patients with PD have reduced beta power entrainment to auditory rhythms (te Woerd et al., 2014, 2018). Given the covariation between motor and timing deficits of patients with PD, the basal ganglia are regarded as the common etiology of these deficits (Grahn, 2012). Regarding intervention, it has been shown that musical rhythms can facilitate the gait of patients with PD (Nombela et al., 2013), and even individualized rhythmic auditory-cueing has been developed for patients with PD to optimize rehabilitation effects (Dalla Bella et al., 2018). Our findings further imply that spectral regularity or predictability would be beneficial for the most effective rhythmic auditory-cueing, as this thesis showed that unpredictable spectral information disrupts beta power entrainment, the neural activity that is associated with motor control among patients with PD (te Woerd et al., 2014, 2018).

DCD is a congenital neurodevelopmental disorder with an approximately 5 to 15% prevalence rate among school-age children (American Psychiatric Association, 2013). Children with DCD are usually described as "clumsy", and they have deficits in fine and/or gross motor skills, such as motor planning, sequencing of movements, and motor timing that result in difficulties in writing, tying shoes, running, and catching a ball (Zwicker et al., 2012). Beyond motor deficits, children with DCD also have deficits in auditory duration and rhythmic time perception (Chang et al., submitted). Furthermore, it has been hypothesized that children with DCD will also have deficits in tracking auditory rhythm and that this might be associated with compromised neural entrainment activities (Trainor et al., 2018). Although the effect of rhythmic auditory-cueing on motor performance in children with DCD remains unknown, a study reporting six cases showed that intervention using auditory rhythm can be useful for improving motor performance (Leemrijse et al., 2000), suggesting that children with DCD might benefit from rhythmic auditory cueing similarly to patients with PD.

In general, it would be interesting to discover whether the neural entrainment activities reported in the current thesis are compromised in people with motor disorders. The results of such inquiries could greatly extend our understanding of both neural entrainment activities and motor disorders.

V.5 Conclusion

This thesis investigated the perceptual functions of high frequency power and low frequency phase neural oscillatory entrainment activities in tracking auditory temporal rhythmic regularity. The findings showed that beta power (high frequency) entrainment activity is associated with prediction in pitch, in addition to temporal prediction. Associating neural entrainment with behavioural performance, the findings suggest that low frequency phase entrainment might reflect temporal prediction, high frequency power entrainment might reflect attention, and their coupling might reflect auditory-motor cortical communication. Together, this thesis has advanced our understanding of how neural entrainment activities implement the perceptual functions needed for tracking auditory rhythms that are fundamental for speech and musical perception.

Bibliography

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Publisher.

Arnal, L. H. (2012). Predicting "when" using the motor system's beta-band oscillations. *Frontiers in Human Neuroscience* 6, 225.

Arnal, L. H., Doelling, K. B., and Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex* 25, 3077–3085.

Arnal, L. H. and Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences* 16, 390–398.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.

Bartolo, R. and Merchant, H. (2015). β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *Journal of Neuroscience* 35, 4635–4640.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1–48.

Bauer, A. K. R., Bleichner, M. G., Jaeger, M., Thorne, J. D., and Debener, S. (2018). Dynamic phase alignment of ongoing auditory cortex oscillations. *NeuroImage* 167, 396–407.

Bauer, A. K. R., Jaeger, M., Thorne, J. D., Bendixen, A., and Debener, S. (2015). The auditory dynamic attending theory revisited: a closer look at the pitch comparison task. *Brain Research* 1626, 198–210.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., and Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences of the USA* 106, 1672–1677.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 289–300.

Bernasconi, F., Manuel, A. L., Murray, M. M., and Spierer, L. (2011). Prestimulus beta oscillations within left posterior sylvian regions impact auditory temporal order judgment accuracy. *International Journal of Psychophysiology* 79, 244–248.

Bertrand, O., Bohorquez, J., and Pernier, J. (1994). Time-frequency digital filtering based on an invertible wavelet transform: An application to evoked potentials. *IEEE T Bio-Medical Engineering* 41, 77–88.

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., Emerson, R. G., and Ce., S. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *Journal of Neuroscience* 31, 3176–3185.

Bianco, R., Novembre, G., Keller, P. E., Villringer, A., and Sammler, D. (2018). Musical genre-dependent behavioural and EEG signatures of action planning: A comparison between classical and jazz pianists. *NeuroImage* 169, 383–394.

Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience* 8, 389–395.

Breska, A. and Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLOS Biology* 15, e2001665.

Buschman, T. J. and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862.

Buschman, T. J. and Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron* 63, 386–396.

Butler, R. A. (1968). Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. *Journal of the Acoustical Society of America* 44, 945–950.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

Calderone, D. J., Lakatos, P., Butler, P. D., and Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences* 18, 300–309.

Canolty, R. T. and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences* 14(11), 506–515.

Chang, A., Bosnyak, D. J., and Trainor, L. J. (2016). Unpredicted pitch modulates beta oscillatory power during rhythmic entrainment to a tone sequence. *Frontiers in Psychology* 7, 327.

Chang, A., Bosnyak, D. J., and Trainor, L. J. (2018). Beta oscillatory power modulation reflects the predictability of pitch change. *Cortex* 106, 248–260.

Chang, A., Bosnyak, D. J., and Trainor, L. J. (2019a). Rhythmicity facilitates pitch discrimination: Differential roles of low and high frequency neural oscillations. *NeuroImage* 198, 31–43.

Chang, A., Ide, J. S., Li, H. H., Chen, C. C., and Li, C. S. R. (2017a). Proactive control: Neural oscillatory correlates of conflict anticipation and response slowing. *eNeuro* 4, ENEURO.0061–0117.2017.

Chang, A., Kragness, H. E., Livingstone, S. R., Bosnyak, D. J., and Trainor, L. J. (2019b). Body sway reflects joint emotional expression in music ensemble performance. *Scientific Reports* 9, 205.

Chang, A., Li, Y. C., Chan, J. F., Dotov, D. G., Cairney, J., and Trainor, L. J. (submitted). Inferior auditory time perception in children with motor difficulties.

Chang, A., Livingstone, S. R., Bosnyak, D. J., and Trainor, L. J. (2017b). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences of the USA* 114, E4134–E4141.

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex* 18, 2844–2854.

Chen, J. L., Zatorre, R. J., and Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *NeuroImage* 32, 1771–1781.

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

Cirelli, L. K., Trehub, S. E., and Trainor, L. J. (2018). Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences* 1423, 66–72.

Cohen, M. X. (2014a). Analyzing Neural Time Series Data: Theory and Practice. *MIT press*.

Cohen, M. X. (2014b). Fluctuations in oscillation frequency control spike timing and coordinate neural networks. *Journal of Neuroscience* 34(27), 8988–8998.

Cole, S. R. and Voytek, B. (2017). Brain oscillations and the importance of waveform shape. *Trends in Cognitive Sciences* 21(2), 137–149.

Conroy, M. A. and Polich, J. (2007). Normative variation of P3a and P3b from a large sample: Gender, topography, and response time. *Journal of Psychophysiology* 21, 22–32.

Cope, T. E., Grube, M., and Griffiths, T. D. (2012). Temporal predictions based on a gradual change in tempo. *The Journal of the Acoustical Society of America* 131(5), 4013–4022.

Correa, Á., Cona, G., Arbula, S., Vallesi, A., and P., B. (2014). Neural dissociation of automatic and controlled temporal preparation by transcranial magnetic stimulation. *Neuropsychologia* 65, 131–136.

Costa-Faidella, J., Baldeweg, T., Grimm, S., and Escera, C. (2011). Interactions between "what" and "when" in the auditory system: temporal predictability enhances repetition suppression. *Journal of Neuroscience* 31, 18590–18597.

Coull, J. T. and Nobre, A. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Current Opinion in Neurobiology* 18, 137–144.

Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Quantitative Methods for Psychology* 1, 42–45.

Cravo, A. M., Rohenkohl, G., Wyart, V., and Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *Journal of Neuroscience* 33, 4002–4010.

Dalla Bella, S., Dotov, D., Bardy, B., and de Cock, V. C. (2018). Individualization of music-based rhythmic auditory cueing in Parkinson's disease. *Annals of the New York Academy of Sciences* 1423(1), 308–317.

den Ouden, H. E., Kok, P., and de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology* 3, 548.

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience* 19, 158–164.

Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., and Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences of the USA* 116, 10113–10121.

Doelling, K. B. and Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences of the USA* 112, E6233–E6242.

Dürschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., and Heinze, H. J. (2016). Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proceedings of the National Academy of Sciences of the USA* 113, 6755–6760.

El Karoui, I., King, J. R., Sitt, J., Meyniel, F., Van Gaal, S., Hasboun, D., Adam, C., Navarro, V., Baulac, M., Dehaene, S., Cohen, L., and Naccache, L. (2015). Event-related potential, time-frequency, and functional connectivity facets of local and global auditory novelty processing: An intracranial study in humans. *Cerebral Cortex* 25, 4203–4212.

Engel, A. K. and Fries, P. (2010). Beta-band oscillations - signalling the status quo? *Current Opinion in Neurobiology* 20, 156–165.

Faugeras, F., Rohaut, B., Weiss, N., Bekinschtein, T., Galanaud, D., Puybasset, L., Bolgert, F., Sergent, C., Cohen, L., Dehaene, S., and Naccache, L. (2012). Event related potentials elicited by violations of auditory regularities in patients with impaired consciousness. *Neuropsychologia* 50, 403–418.

Florin, E., Vuvan, D., Peretz, I., and Baillet, S. (2017). Pre-target neural oscillations predict variability in the detection of small pitch changes. *PLOS One* 12, e0177836.

Fox, J. and Weisberg, S. (2011). An R companion to applied regression.

Franken, M. K., Eisner, F., Acheson, D. J., McQueen, J. M., Hagoort, P., and Schoffelen, J. M. (2018). Self-monitoring in the cerebral cortex: neural responses to small pitch shifts in auditory feedback during speech production. *NeuroImage* 179, 326–336.

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

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

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences* 13, 293–301.

Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience* 6, 1216–1223.

Fujioka, T., Ross, B., and Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience* 35, 15187–15198.

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences* 1169, 89–92.

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience* 32, 1791–1802.

Fujioka, T., Trainor, L. J., and Ross, B. (2008). Simultaneous pitches are encoded separately in auditory cortex: an MMNm study. *Neuroreport* 19, 361–366.

Geerligs, L. and Akyürek, E. G. (2012). Temporal integration depends on increased prestimulus beta band power. *Psychophysiology* 49, 1632–1635.

Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., and Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Current Biology* 22(13), 1176–1182.

Giraud, A. L. and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience* 15, 511–517.

Godey, B., Schwartz, D., De Graaf, J. B., Chauvel, P., and Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clinical Neurophysiology* 112, 1850–1859.

Golumbic, Z., M., E., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., Poeppel, D., and Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron* 77(5), 980–991.

Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., and Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *Journal of Neuroscience* 31, 18556–18567.

Grahn, J. A. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. *Topics in Cognitive Science* 4, 585–606.

Grahn, J. A. and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience* 19, 893–906.

Grahn, J. A. and Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex* 45(1), 54–61.

Griffiths, T. D. and Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience* 5, 887–892.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., and Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLOS Biology* 11(12), e1001752.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., and Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the USA* 101, 13050–13055.

Grube, M., Cooper, F. E., Chinnery, P. F., and Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences of the USA* 107, 11597–11601.

Haegens, S. and Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: a critical review. *Neuroscience & Biobehavioral Reviews* 86, 150.

Haenschel, C., Baldeweg, T., Croft, R. J., Whittington, M., and Gruzelier, J. (2000). Gamma and beta frequency oscillations in response to novel auditory stimuli: A comparison of human electroencephalogram (EEG) data with in-itro models. *Proceedings of the National Academy of Sciences of the USA* 97, 7645–7650.

Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., and Voytek, B. (2018). Parameterizing neural power spectra. *bioRxiv* 11.

Hanslmayr, S. and Staudigl, T. (2014). How brain oscillations form memories - a processing based perspective on oscillatory subsequent memory effects. *NeuroImage* 85, 648–655.

Heideman, S. G., van Ede, F., and Nobre, A. C. (2018). Temporal alignment of anticipatory motor cortical beta lateralisation in hidden visual-motor sequences. *European Journal of Neuroscience* 48, 2684–2695.

Heilbron, M. and Chait, M. (2018). Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience* 389, 54–73.

Heinrich, H., Busch, K., Studer, P., Erbe, K., Moll, G. H., and Kratz, O. (2014). Refining the picture of reduced alerting responses in ADHD-A single-trial analysis of event-related potentials. *Neuroscience Letters* 582, 49–53.

Henry, M. J. and Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception* 2, 62–86.

Henry, M. J., Herrmann, B., and Obleser, J. (2014). Entrained neural oscillations in multiple frequency bands comodulate behavior. *Proceedings of the National Academy of Sciences of the USA* 111, 14935–14940.

Henry, M. J. and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences of the USA* 109, 20095–20100.

Herbst, S. K. and Obleser, J. (2017). Implicit variations of temporal predictability: shaping the neural oscillatory and behavioural response. *Neuropsychologia* 101, 141–152.

Herrmann, B., Henry, M. J., Fromboluti, E. K., McAuley, J. D., and Obleser, J. (2015). Statistical context shapes stimulus-specific adaptation in human auditory cortex. *Journal of Neurophysiology* 113, 2582–2591.

Herrmann, B., Henry, M. J., Haegens, S., and Obleser, J. (2016). Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *NeuroImage* 124, 487–497.

Herrmann, B., Henry, M. J., and Obleser, J. (2013). Frequency-specific adaptation in human auditory cortex depends on the spectral variance in the acoustic stimulation. *Journal of Neurophysiology* 109, 2086–2096.

Herrmann, B., Schlichting, N., and Obleser, J. (2014). Dynamic range adaptation to spectral stimulus statistics in human auditory cortex. *Journal of Neuroscience* 34, 327–331.

Herrmann, C. S., Munk, M. H., and Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences* 8, 347–355.

Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders* 45, 393–402.

Hickok, G., Farahbod, H., and Saberi, K. (2015). The rhythm of perception: Entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological Science* 26, 1006–1013.

Hickok, G. and Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393–402.

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology* 3, 96.

Hong, L. E., Buchanan, R. W., Thaker, G. K., Shepard, P. D., and Summerfelt, A. (2008). Beta (~16 Hz) frequency neural oscillations mediate auditory sensory gating in humans. *Psychophysiology.* 45, 197–204.

Hove, M. J., Marie, C., Bruce, I. C., and Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences of the USA* 111, 10383–10388.

Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., and Giraud, A. L. (2015a). Speech encoding by coupled cortical theta and gamma oscillations. *Elife* 4, e06213.

Hyafil, A., Giraud, A. L., Fontolan, L., and Gutkin, B. (2015b). Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends in Neurosciences* 38, 725–740.

Iversen, J. R. and Balasubramaniam, R. (2016). Synchronization and temporal processing. *Current Opinion in Behavioral Sciences* 8, 175–180.

Iversen, J. R., Repp, B. H., and Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences* 1169, 58–73.

Jenkinson, N. and Brown, P. (2011). New insights into the relationship between dopamine, beta oscillations and motor function. *Trends in Neurosciences* 34, 611–618.

Jones, M. R. (2010). Attending to sound patterns and the role of entrainment. In: *Attention and Time.* Ed. by A. Nobre and J. T. Coull. Oxford University Press. p, 317–330.

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

Jones, M. R., Johnston, H. M., and Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology* 53, 59–96.

Jones, M. R., Moynihan, H., MacKenzie, N., and Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological science* 13(4), 313–319.

Jones, S. R. (2016). When brain rhythms aren't "rhythmic": implication for their mechanisms and meaning. *Current Opinion in Neurobiology* 40, 72–80.

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

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178.

Kalcher, J. and Pfurtscheller, G. (1995). Discrimination between phase-locked and non-phase-locked event-related EEG activity. *Electroencephalography and Clinical Neurophysiology* 94, 381–384.

Kayser, S. J., McNair, S. W., and Kayser, C. (2016). Prestimulus influences on auditory perception from sensory representations and decision processes. *Proceedings of the National Academy of Sciences of the USA* 113, 4842–4847.

Keitel, A., Gross, J., and Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLOS Biolology* 16, e2004473.

Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., and Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Experimental Neurology* 245, 15–26.

King, A. J. and Nelken, I. (2009). Unraveling the principles of auditory cortical processing: Can we learn from the visual system? *Nature Neuroscience* 12, 698–701.

Kingdom, F. A. A. and Prins, N. (2010). *Psychophysics: A Practical Introduction*. Academic Press.

Kisley, M. A. and Cornwell, Z. (2006). Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clinical Neurophysiology* 117, 2549–2563.

Klein, J. M. and Jones, M. R. (1996). Effects of attentional set and rhythmic complexity on attending. *Perception & Psychophysics* 58(1), 34–46.

Kopell, N., Whittington, M. A., and Kramer, M. A. (2011). Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proceedings of the National Academy of Sciences of the USA* 108, 3779–3784.

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., and Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology* 28(18), 2867–2875.

Kotz, S. A., Ravignani, A., and Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences* 22(10), 896–910.

Kragness, H. E. and Trainor, L. J. (2016). Listeners lengthen phrase boundaries in self-paced music. *Journal of Experimental Psychology: Human Perception and Performance* 42(10), 1676.

Kragness, H. and Trainor, L. (2018). Young children pause on phrase boundaries in self-paced music listening: The role of harmonic cues. *Developmental Psychology* 54, 842–856.

Kranczioch, C., Debener, S., Maye, A., and Engel, A. K. (2007). Temporal dynamics of access to consciousness in the attentional blink. *NeuroImage* 37, 947–955.

Lakatos, P., Barczak, A., Neymotin, S. A., McGinnis, T., Ross, D., Javitt, D. C., and O'Connell, M. N. (2016). Global dynamics of selective attention and its lapses in primary auditory cortex. *Nature Neuroscience* 19, 1707–1717.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., and Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77, 750–761.

Lanting, C. P., Briley, P. M., Sumner, C. J., and Krumbholz, K. (2013). Mechanisms of adaptation in human auditory cortex. *Journal of Neurophysiology* 110, 973–983.

Large, E. W., Herrera, J. A., and Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience* 9, 159.

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

Lee, J. H., Whittington, M. A., and Kopell, N. J. (2013). Top-down beta rhythms support selective attention via interlaminar interaction: A model. *PLOS Computational Biology* 9.

Leemrijse, C., Meijer, O. G., Vermeer, A., Adèr, H. J., and Diemel, S. (2000). The efficacy of Le Bon Départ and Sensory Integration treatment for children with developmental coordination disorder: a randomized study with six single cases. *Clinical Rehabilitation* 14, 247–259.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America* 49, 467–477.

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., and Stephan, K. E. (2013). Modelling trial-by-trial changes in the mismatch negativity. *PLOS Computational Biology* 9.

Loehr, J. D., Large, E. W., and Palmer, C. (2011). Temporal coordination and adaptation to rate change in music performance. *Journal of Experimental Psychology: Human Perception & Performance* 37(4), 1292–1309.

Luce, R. (1986). *Response Times: Their role in inferring elementary mental organization*. New York (US): Oxford University Press.

Luck, S. J. and Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology* 54, 146–157.

Lüdecke, D. (2016). *sjPlot: Data visualization for statistics in social science*. R package version.

Manning, F. and Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin and Review* 20, 1133–1139.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* 164, 177–190.

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

Matsuda, A., Hara, K., Miyajima, M., Matsushima, E., Ohta, K., and Matsuura, M. (2013). Distinct pre-attentive responses to non-scale notes: An auditory mismatch negativity (MMN) study. *Clinical Neurophysiology* 124, 1115–1121.

Max, C., Widmann, A., Schröger, E., and Sussman, E. (2015). Effects of explicit knowledge and predictability on auditory distraction and target performance. *International Journal of Psychophysiology* 98, 174–181.

McAuley, J. D. and Fromboluti, E. K. (2014). Attentional entrainment and perceived event duration. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, 20130401.

McAuley, J. D., Henry, M. J., and Tkach, J. (2012). Tempo mediates the involvement of motor areas in beat perception. *Annals of the New York Academy of Sciences* 1252, 77–84.

McAuley, J. D. and Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance* 29(6), 1102–1125.

McDermott, J. H. (2018). Audition. In: *Sensation, Perception, and Attention*. Ed. by J. Serences. Vol. 2. Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience. Wiley.

Meijer, D., te Woerd, E., and Praamstra, P. (2016). Timing of beta oscillatory synchronization and temporal prediction of upcoming stimuli. *NeuroImage* 138, 233–241.

Merchant, H. and Bartolo, R. (2017). Primate beta oscillations and rhythmic behaviors. *Journal of Neural Transmission* 124, 1–10.

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

Morillon, B. and Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences of the USA* 114, E8913–E8921.

Morillon, B., Hackett, T. A., Kajikawa, Y., and Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology* 31, 230–238.

Morillon, B. and Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences* 1337, 26–31.

Morillon, B., Schroeder, C. E., and Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications* 5, 5255.

Morillon, B., Schroeder, C. E., Wyart, V., and Arnal, L. H. (2016). Temporal prediction in lieu of periodic stimulation. *Journal of Neuroscience* 36, 2342–2347.

Mouraux, A. and Iannetti, G. D. (2008). Across-trial averaging of event-related EEG responses and beyond. *Magnetic Resonance Imaging* 26, 1041–1054.

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

Näätänen, R. and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology* 24, 375–425.

Näätänen, R., Sams, M., Alho, K., Paavilainen, P., Reinikainen, K., and Sokolov, E. N. (1988). Frequency and location specificify of the human vertex N1 wave. *Electroencephalography and Clinical Neurophysiology* 69, 523–531.

Nakata, T. and Trainor, L. J. (2015). Perceptual and cognitive enhancement with an adaptive timing partner: Electrophysiological responses to pitch change. *Psychomusicology: Music, Mind and Brain* 25(4), 404–415.

Ng, B. S. W., Logothetis, N. K., and C., K. (2012a). EEG phase patterns reflect the selectivity of neural firing. *Cerebral Cortex* 23, 389–398.

Ng, B. S. W., Schroeder, T., and C., K. (2012b). A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *Journal of Neuroscience* 32, 12268–12276.

Nobre, A. C., Correa, A., and Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology* 17, 465–470.

Nobre, A. C. and van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Review Neuroscience* 19, 34–48.

Nombela, C., Hughes, L. E., Owen, A. M., and Grahn, J. A. (2013). Into the groove: can rhythm influence Parkinson's disease? *Neuroscience & Biobehavioral Reviews* 37(10), 2564–2570.

Novembre, G. and Iannetti, G. D. (2018). Tagging the musical beat: Neural entrainment or event-related potentials? *Proceedings of the National Academy of Sciences of the USA* 115(47), E11002–E11003.

Novembre, G., Knoblich, G., Dunne, L., and Keller, P. E. (2017). Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Social Cognitive and Affective Neuroscience* 12, 662–670.

Obleser, J., Henry, M. J., and Lakatos, P. (2017). What do we talk about when we talk about rhythm? *PLOS Biolology* 15, e2002794.

O'Brien, R. G. and Kaiser, M. K. (1985). Manova method for analyzing repeated measures designs: an extensive primer. *Psychological Bulletin* 97, 316–333.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience* 2011, 156869.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance* 15(2), 331–346.

Patel, A. D. (2010). *Music, language, and the brain*. Oxford: university press.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., and Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage.* 50, 302–313.

Peelle, J. E. and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology* 3, 320.

Pefkou, M., Arnal, L. H., Fontolan, L., and Giraud, A. L. (2017). θ-Band and β-band neural activity reflects independent syllable tracking and comprehension of time-compressed speech. *Journal of Neuroscience* 37, 7930–7938.

Peretz, I. (2016). Neurobiology of congenital amusia. *Trends in Cognitive Sciences* 20, 857–867.

Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Research* 41, 1257–1260.

Pfurtscheller, G. and Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology* 110, 1842–1857.

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

Polich, J. (2007). Updating P300:An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118, 2128–2148.

Prins, N. and Kingdom, F. A. A. (2009). *Palamedes: matlab routines for analyzing psychophysical data*. `http://www.palamedestoolbox.org`.

Rankin, S. K., Large, E. W., and Fink, P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Perception* 26(5), 401–413.

Regnault, P., Bigand, E., and Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience* 13, 241–255.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review* 12, 969–992.

Riecke, L., Formisano, E., Sorger, B., Başkent, D., and Gaudrain, E. (2018). Neural entrainment to speech modulates speech intelligibility. *Current Biology* 28(2), 161–169.

Rimmele, J. M., Morillon, B., Poeppel, D., and Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences* 22(10), 870–882.

Rinne, T., Särkkä, A., Degerman, A., Schröger, E., and Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research* 1077, 135–143.

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

Rohenkohl, G., Cravo, A. M., Wyart, V., and Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *Journal of Neuroscience* 32, 8424–8428.

Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., and Hatsopoulos, N. G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron.* 65, 461–471.

Savage, P. E., Brown, S., Sakai, E., and Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences of the USA* 112, 8987–8992.

Scherg, M. and Von Cramon, D. (1985). Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalography and Clinical Neurophysiology* 62, 32–44.

Schönwiesner, M. and Zatorre, R. J. (2009). Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. *Proceedings of the National Academy of Sciences of the USA* 106, 14611–14616.

Schroeder, C. E. and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences* 32, 9–18.

Schröger, E., Giard, M. H., and Wolff, C. (2000). Auditory distraction: Event-related potential and behavioral indices. *Clinical Neurophysiology* 111, 1450–1460.

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

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences* 11, 211–218.

Schubotz, R. I., Friederici, A. D., and Von Cramon, D. Y. (2000). Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *NeuroImage* 11, 1–12.

Schubotz, R. I. and Von Cramon, D. Y. (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cerebral Cortex* 11, 210–222.

Schubotz, R. I. and Von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *NeuroImage* 15, 787–796.

Schubotz, R. I., Von Cramon, D. Y., and Lohmann, G. (2003). Auditory what, where, and when: A sensory somatotopy in lateral premotor cortex. *NeuroImage* 20, 173–185.

Schwartze, M. and Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews* 37, 2587–2596.

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., Howard, M. A., and Griffiths, T. (2016). Neural signatures of perceptual inference. *Elife* 5.

Slotnick, S. D. (2004). Source localization of ERP generators. In: *Event-Related Potentials: A Methods Handbook*. Ed. by T. C. Todd. MIT Press, 149–166.

Snyder, J. S. and Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research* 24, 117–126.

Spencer, K. M. (2004). Averaging, detection, and classification of single-trial ERPs. In: *Event-Related Potentials: A Methods Handbook*. Ed. by T. C. Todd. MIT Press, 209–228.

Spitzer, B. and Haegens, S. (2017). Beyond the status quo: A role for beta oscillations in endogenous content (re) activation. *eNeuro* 170, ENEURO–0170.2017.

Sridharan, D., Levitin, D. J., Chafe, C., Berger, J., and Menon, V. (2007). Neural dynamics of event segmentation in music: Converging evidence for dissociable ventral and dorsal networks. *Neuron* 55, 521–532.

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., and Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience* 30(41), 13578–13585.

Sussman, E., Winkler, I., and Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin and Review* 10, 630–637.

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., and Zion Golumbic, E. (2017). Neural entrainment to the beat: the "missing-pulse" phenomenon. *Journal of Neuroscience* 37, 6331–6341.

te Woerd, E. S., Oostenveld, R., de Lange, F. P., and Praamstra, P. (2014). A shift from prospective to reactive modulation of beta-band oscillations in Parkinson's disease. *NeuroImage* 100, 507–519.

te Woerd, E., Oostenveld, R., de Lange, F., and Praamstra, P. (2018). Entrainment for attentional selection in Parkinson's disease. *Cortex*, 166–178.

Teki, S. and Griffiths, T. D. (2014). Working memory for time intervals in auditory rhythmic sequences. *Frontiers in Psychology* 5, 1329.

Teki, S. and Griffiths, T. D. (2016). Brain bases of working memory for time intervals in rhythmic sequences. *Frontiers in Neuroscience* 10, 239.

Teki, S., Grube, M., and Griffiths, T. D. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience* 5, 90.

Teki, S., Grube, M., Kumar, S., and Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *Journal of Neuroscience* 31, 3805–3812.

Teki, S. and Kononowicz, T. W. (2016). Commentary: Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Frontiers in Neuroscience* 10, 389.

ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., Mehta, A. D., Mégevand, P., Groppe, D. M., and Zion Golumbic, E. (2017). Low-frequency cortical oscillations entrain to subthreshold rhythmic auditory stimuli. *Journal of Neuroscience* 37, 4903–4912.

Todorovic, A., Schoffelen, J. M., van Ede, F., Maris, E., and de Lange, F. P. (2015). Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PLOS One* 10, e0120288.

Trainor, L. J., Chang, A., Cairney, J., and Li, Y. C. (2018). Is auditory perceptual timing a core deficit of developmental coordination disorder? *Annals of the Academy of Sciences* 1423, 30–39.

Triviño, M., Arnedo, M., J., L., Chirivella, J., and Correa, Á. (2011). Rhythms can overcome temporal orienting deficit after right frontal damage. *Neuropsychologia* 49, 3917–3930.

Tse, P. U., Intriligator, J., Rivest, J., and Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics* 66(7), 1171–1189.

van Driel, J., Cox, R., and Cohen, M. X. (2015). Phase-clustering bias in phase-amplitude cross-frequency coupling and its removal. *Journal of Neuroscience Methods* 254, 60–72.

van Ede, F., Jensen, O., and Maris, E. (2010). Tactile expectation modulates pre-stimulus β-band oscillations in human sensorimotor cortex. *NeuroImage* 51, 867–876.

van Ede, F., Quinn, A. J., Woolrich, M. W., and Nobre, A. C. (2018). Neural oscillations: sustained rhythms or transient burst-events? *Trends in Neurosciences* 41(7), 415–417.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., and Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences of the USA* 108, 20754–20759.

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews* 90, 1195–1268.

Warren, J. E., Wise, R. J., and Warren, J. D. (2005). Sounds do-able: Auditory-motor transformations and the posterior temporal plane. *Trends in Neurosciences* 28, 636–643.

Wetzel, N., Schröger, E., and Widmann, A. (2013). The dissociation between the P3a event-related potential and behavioral distraction. *Physiology* 50, 920–930.

Widmann, A., Schröger, E., and Maess, B. (2015). Digital filter design for electrophysiological data-a practical approach. *Journal of Neuroscience Methods* 250, 34–46.

Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences* 13, 532–540.

Wróbel, A. (2000). Beta activity: a carrier for visual attention. *Acta Neurobiologiae Experimentalis* 60, 247–260.

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences* 6, 37–46.

Zatorre, R. J. and Evans, A. C. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846.

Zendel, B. R., Lagrois, M. É., Robitaille, N., and Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: the curious case of amusia. *Journal of Neuroscience* 35, 3815–3824.

Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., Wei, C., and Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences of the USA* 102, 2293–2298.

Zoefel, B., Archer-Boyd, A., and Davis, M. H. (2018). Phase entrainment of brain oscillations causally modulates neural responses to intelligible speech. *Current Biology* 28(3), 401–408.

Zwicker, J. G., Missiuna, C., Harris, S. R., and Boyd, L. A. (2012). Developmental coordination disorder: a review and update. *European Journal of Paediatric Neurology* 16, 573–581.