AUDITORY-MOTOR SYNCHRONIZATION AND LISTENING EFFORT IN ISOCHRONOUS SOUND SEQUENCE

Ph.D. Thesis

By

YOGESH KUMAR SHIVHARE

DEPARTMENT OF BIOSCIENCES AND BIOMEDICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

AUGUST, 2021

AUDITORY-MOTOR SYNCHRONIZATION AND LISTENING EFFORT IN ISOCHRONOUS SOUND SEQUENCE

A THESIS

Submitted in partial fulfillment of the requirements for the award of the degree
of
DOCTOR OF PHILOSOPHY

by

YOGESH KUMAR SHIVHARE

DEPARTMENT OF BIOSCIENCES AND BIOMEDICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

AUGUST, 2021

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **AUDITORY-MOTOR SYNCHRONIZATION AND LISTENING EFFORT IN ISOCHRONOUS SOUND SEQUENCE,** the partial fulfillment of the requirements for the award of the degree of **DOCTOR OF PHILOSOPHY,** and submitted in the **DEPARTMENT OF BIOSCIENCES AND BIOMEDICAL ENGINEERING, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from July 2015 to August 2021 under the supervision of Dr. Sanjram Premjit Khanganba, Associate Professor, Indian Institute of Technology Indore.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

29/08/2021
(YOGESH KUMAR SHIVHARE)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of Thesis Supervisor
(Dr. SANJRAM PREMJIT KHANGANBA)

YOGESH KUMAR SHIVHARE has successfully given his Ph.D. Oral Examination held on **June 27, 2023.**

Signature of Thesis Supervisor
Date: 27/06/2023

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Thanks to almighty Lord Shiva, who's divine "DAMRU" inspired me to understand the cosmic energy movement through low-frequency sounds. I thank the god for bestowing upon me the utmost strength, patience, and endurance to complete this journey. This dissertation is an end result of a perfect working relationship with my thesis supervisor, Dr. Sanjram Premjit Khanganba, to whom I am eternally grateful. It would not be any exaggeration to say that Dr. Sanjram provided unreserved support during my PhD tenure and generously paved the way for my development as a research scientist. I thank Dr. Sanjram for finely carving the attitude of a true scientist in me. Perhaps most importantly, I appreciate him for believing in me and supporting me in terms of all the requirements that I needed for executing this research.

I wish to thank wholeheartedly those who have contributed to this dissertation as my committee members, Dr. Ram Bilas Pachori, Dr. Debashish Nayak, Dr. Vivek Kant, Dr. Laxmi Iyenger, for their invaluable support, timely advice, and critical feedback for improving the quality of my research.

I am also greatly indebted to my spiritual Guru, Dr. Santosh Kumar Bhadauria ji, without whom I could not complete this journey until the last moment. Writing this dissertation was not a lonely experience. It really would have been tough if I was not surrounded by some cheerful friends who would do anything to reduce my stress and boost my morale. I deeply acknowledge and appreciate the most precious and genuine support received from my sister, Ms. Radha Shivhare (Research Scholar NBRI), Ms. Juhee Singh Verma (Research Scholar IIT Indore), Mr. Anantha (Research Scholar IIT Indore), Ms. Vidushi Jaya (Research Scholar IIT Indore), Ms. Prachi Verma (Research Scholar IIT Indore), and Ms. Ritu Sharma (Research Scholar IIT Indore). They stood by me when I needed them the most.

To

My Parents

LIST OF ABBREVIATIONS

AMS	Auditory-motor synchronization
CV	Coefficient of Variation
fMRI	Function Magnetic Resonance Imaging
HF	High-Frequency
ISI	Inter-Stimulus Interval
ITI	Inter-tap Interval
LE	Listening Effort
LF	Low-Frequency
MV	Motor Variability
NMA	Negative Mean Asynchrony
PD	Parkinson’s Disease
PMA	Positive Mean Asynchrony
RAS	Rhythmic Auditory Stimulation
SE	Synchronization Error
SMS	Sensory-motor synchronization
STM	Short-Term Memory

ABSTRACT

This research examines auditory motor synchronization (AMS) in relation to tone frequency (perceptual property of auditory stimuli) and interval timing present in an isochronous sound sequence. Analyses of tapping behavior, synchronization errors, and listening effort (LE) are emphasized. The study was conducted in a controlled laboratory environment having acoustic treated facility. The role of frequency of tone and interval timing are investigated in affecting the AMS in its entirety, which includes tapping performance, listening effort, motor behavior, and variability. This study employs a synchronized tapping paradigm. This dissertation is comprised of four parts in terms of understanding — anticipatory tapping and reactive tapping under two different inter-stimulus intervals (ISI), motor response to the tone and listening effort involved in performing a continuous tapping task, tone frequency (perceptual property of auditory stimuli) on inter-tap interval (an index of motor process) under the peri-second and supra-second range, and tone frequency on motor variability (MV) studied through the coefficient of the variation of ITI (CV) under the peri-second and supra-second range.

Part-1 examines anticipatory tapping and reactive tapping under two different ISIs. The analysis reveals that the ISI plays an important role in sensory motor synchronization (SMS). The analysis of asynchrony revealed that two different types of tapping occurred under two different ISIs. Under short ISI (1000 ms), participants executed their responses before the tone (i.e., anticipatory tapping driven by feed-forward motor control). Under long ISI (2000 ms), participants executed their responses after the tone (i.e., reactive tapping driven by feed-back motor control mechanisms). In summary, participants showed anticipatory tapping in the absence of

top-down attention and reactive tapping with the involvement of top-down attention.

Part-2 examines the motor response of the participants to the tone and their listening effort involved in performing the continuous tapping task. The emphasis is on how the effect of tone frequency and ISI affect synchronization error and listening effort in a isochronous sound sequence. The analysis reveals that the frequency of tone plays a crucial role in tapping performance and listening effort. In summary, this study demonstrates that there is better temporal alignment with the low-frequency tones with lesser listening effort as compared to high-frequency tones.

Part-3 examines the effect of tone frequency on the inter-tap interval (an index of motor process) under the peri-second and supra-second range. The analysis reveals that perception-action coupling is influenced by cognition and more prominent in peri-second range where less cognitive resources involved unlike supra-second range. Under peri-second range, on average, participants over translated objective time (ITI > ISI) with low-frequency tone and under translated (ITI < ISI) with high-frequency tone in peri-second range. The results suggest that tone frequency has a greater influence on motor timing and perception-action coupling within peri-second range as compared to supra-second range. In conclusion, peri-second timing is mainly subject to local sensory processing, while supra-second timing is dependent on more centralized mechanisms. Tone frequency should therefore be a significant factor in the design of the sound sequence to be used for rhythmic auditory stimulation.

Part-4 focuses on examining the effect of tone frequency on MV studied through CV of ITI under peri-second and supra-second range. The analysis reveals that motor variability is susceptible to frequency influences in long ISI where more cognitive resources involved unlike peri -second range. Under long ISI, participants

showed higher MV with low-frequency tone as compared to high-frequency tone. The results suggest that low frequency tones are most suitable for auditory- motor learning within supra-second range.

Keywords: Sensory–motor synchronization; Auditory-motor synchronization (AMS); Frequency; Inter-Stimulus Interval; Synchronization Error; Motor Variability; Anticipatory Tapping; Reactive Tapping; Inter-Tap Interval; Subjective translation index (STI)

Chapter 1

Introduction

1.1. Background of the study

Music has a vital role in expressing feelings and emotions in people's daily lives. In human evolution, music has played a conspicuous role in influencing human cognition, interpersonal interaction, and cultural history. Besides this, music has the potential to trigger human movements. Due to this unique property, music has a significant role in movement-based therapies, especially in movement disorders, which are commonly seen in older people. One of the most common movement disorders seen in elderly people is Parkinson's disease (PD).

The basal ganglia play an important role in controlling movements through its close association with internal timing and rhythm perception. PD is a neurodegenerative pathological condition in which the loss of dopamine producing neurons in the substantia nigra of the midbrain, affects dopamine transmission from this area to the striatum of the basal ganglia (Jones & Jahanshahi, 2006). Due to such a dopamine deficiency, a loss of rhythmicity in timed movements, impediment in initiating movement, diminished length of steps and speed, freezing of strides due to impairment in alternating timed movements like walking, and other similar abnormalities in movement are observed in PD. Different mechanisms have been proposed to explain the therapeutic effects of music and rhythm on motor symptoms in patients with PD. One potential mechanism involves recruitment of alternative neural pathways that are relatively spared in PD. The movement impairment in PD may be atoned by the engagement of cerebellar networks, as they are spared in PD and depend upon external auditory or visual cues.

There have been many rehabilitative approaches for providing Gait training for various movement disorders. Auditory rhythmic signaling, auditory rhythmic stimulation, or acoustic rhythmic cueing is one of the most recent treatment approaches with greater proof of efficacy for its beneficial effect on functional walking. In this approach, metronomes or previously chosen music create auditory rhythms to signal the cadence while walking so that the patient synchronizes their footsteps with the stimulus provided by the rhythm.

Accomplishing such synchronization results in immediate and sustained effects that are passed on to different gait features or parameters, like:

1. Increase in the length of steps, cadence, and symmetry
2. Improvement in functional walking capability by generating quicker walking and longer strides, and perceptible by increased speed of gait.

To date, there have been no studies investigating RAS from the point of view of attention, where the magnitude of attention's involvement is different at two different intervals. It is not well understood, up to what interval-range RAS works efficiently, which interval range is most suitable for motor learning (as it requires conscious control), and what type of frequency of tone is required for RAS to work efficiently.

The concept of rhythmic auditory stimulation has acquired much research attention in recent decades, but there is still scope to explore issues concerning temporal and pitch information processing and their causal relationship with tapping behavior, motor timing and variability, listening effort (LE), and synchronization errors under different levels of attention. There is meager knowledge about these issues. This may have been partially so, because auditory-motor synchronization (AMS) lacks experimental paradigms. This dissertation reports a study emphasizing scientific investigation of issues related to motor timing and variability, tapping behavior and synchronization error.

1.2. Context of the problem

In order to work effectively in life, we need to align our behavior with the surrounding environment and other human beings. In our normal day to day activities, like speech and expression, and even in more complex endeavors like singing and dancing, auditory-motor synchronization plays a crucial role. An understanding of both normal and abnormal brain functions is crucial to comprehend the fundamental mechanisms that enable us to accomplish such complex coordination tasks. Several studies in the past have examined periodic motor activity to assess time perception and the capability to coordinate with external stimuli (Repp, 2005; Repp and Su, 2013). In terms of implications, such a body of knowledge is of vital importance in dealing with and treating movement disorders like PD.

There have been many rehabilitative approaches for providing Gait training for various movement disorders. Auditory rhythmic signaling, auditory rhythmic stimulation, or acoustic rhythmic cueing is one of the most recent treatment approaches with greater proof of efficacy for its beneficial effect on functional walking. In this approach, metronomes or previously chosen music create auditory rhythms to signal the cadence while walking, so that the patient synchronizes their footsteps with the stimulus provided by the rhythm.

In the context of rhythmic auditory stimulation, since the environment is dynamic, and the subjects need to process the temporal information from the isochronous sound sequence to synchronize their movement with the tone, Attention and WM are involved in attending and processing of information.

In order to transfer the temporal information to motor structures, the auditory system plays a vital role in processing the temporal information rapidly and accurately, thereby maintaining a synchrony between the rhythmic sound sequence and the motor response (e.g., the tapping response used in this study). This is because there is a high range of connectivity between the auditory and motor systems at the cortical, subcortical, and spinal levels (Thaut & Abiru, 2010). The efficacy of metronome or music-based RAS training (artificial RAS) is therefore very well known in the PD literature. Although the tempo of RAS is typically assessed on the basis of the patient's own cadence, the issue of determining the tempo and frequency of auditory stimuli irrespective of age, gender, or cognitive abilities has not been adequately addressed in the research literature. Moreover, RAS has not been studied in its entirety, including movement efficiency and listening effort.

1.3. Purpose of the study

Individuals, when engaged in music listening, start moving their bodies in automated mode. But what properties of music lead to this spontaneity and automaticity need to be investigated. In this process, the perception-action link plays an important role because perception translates into action, which leads to auditory motor synchronization. In AMS, participants perform two tasks simultaneously; one is listening and the other is tapping, and both consume the same pool of attentional resources. First the subject listens to the metronome, then the brain processes the temporal and spectral information present in it and processes it into motor structures, thus creating entrainment.

AMS has been investigated from the perception-action perspective, but how this coupling affects cognition, which differs in its level of magnitude under peri-second and supra-second inter-stimulus interval (ISI) timing, is the main purpose of this study. The peri-second range is the transition zone between the sub-second and supra-second ranges. In the context of neural correlates under sub-second and supra second time systems, there are two different neural networks for these time systems (Hayashi et al., 2014; Jahanshahi et al., 2006; Lewis and Miall, 2003b; Pouthas et al., 2005; Wiener et al., 2010). Supra-second timing stimulates the frontal and posterior parietal parts of the cortex, which are associated with working memory and attention. The other region stimulated in the sub-second and peri-second timing ranges is the motor and somatosensory systems, which incorporate the precentral and postcentral regions and the supplementary motor area (SMA).

In previous studies, AMS had been examined through a synchronized tapping framework, in which subjects were instructed to tap along with the auditory stimuli present in the metronome in a synchronized manner, and in order to do so, subjects did error and trail to achieve the synchronization, which led to tap-to-tap variability, also called motor variability. To achieve exact synchronization is an ideal condition that is difficult to achieve because of the nerve conduction delay, which includes the time from the ear to the brain and from the brain to the muscle. The time taken from the ear to the brain will depend upon the response time of auditory nerve fibers towards tone, which is different for low-frequency (LF) and high frequency (HF) tones respectively. In this way, frequency of tone and interval timing can affect the AMS, but in what way is a matter for this current investigation.

Till now, AMS to isochronous sound sequence has been investigated in terms of tapping performance analyzed by synchronization error (SE), but not in terms of the mental effort associated with it. This is because efficient AMS involves not only tapping performance but also requires mental effort, as two different individuals with different cognitive abilities and learning styles can demonstrate the same performance with their different efforts.

In synchronizing with an isochronous sound sequence, some attention and cognitive resources are needed to retrieve the temporal information present in the sound sequence in order to have AMS. This is called listening effort (LE), which may be understood as a kind of mental effort. *Mental effort* is an energy mobilising (Fairclough & Houston, 2004) neurocognitive mechanism that represents the regulated allocation of psychological information-processing resources in the performance of perception, cognition, and action tasks. In the face of increased task requirements and the presence of psychological stressors, the deployment of mental effort is a compensatory tactic to maintain performance (Hockey, 1997). To mitigate the negative effects of mental fatigue on real-world tapping efficiency, there is a realistic need to incorporate and assess mental effort.

Listening is multi-dimensional and auditory-motor synchronization likely requires auditory-cognitive interactions. According to the FUEL model, cognitive domains involved in listening include working memory, attention, and speed of processing. There is the possibility that the spectral content of auditory stimuli, which is the frequency, will affect the LE. People experience high-frequency sounds as sharp and unpleasant compared to low-frequency (LF) sounds, which are widely known to be bass sounds and high-frequency (HF) sounds are perceived as being louder than low-frequency sounds at the same intensity levels. Apart from the frequency of tone, which is part of stimulus characteristic, LE also depends upon the demands of the listening situation. Now that the demand of the listening situation is higher in the ISI range between 1800 and 3600 ms (due to the involvement of top-down attention), there are high chances that LE will be different in both ISI.

In this way, tone frequency and ISI are important aspects that need to be looked at from an efficient AMS point of view, which includes listening effort and AMS performance, which comprises motor behavior, motor variability, and synchronization error and their actual causal factors. Considering that establishing suitable procedures for experimentation is essential, this dissertation emphasizes the development and reporting of synchronized tapping experimental paradigms. The current research intends to gain a scientific understanding of the underlying causal relationships between tone frequency, interval timing, tapping performance, and tapping behavior in the context of AMS.

1.4. Significance and objectives of the study

Tapping in response to an isochronous sound sequence is a task undertaken in a highly dynamic environment and relies on temporal and non-temporal information present in the isochronous sound sequence, such as tone frequency and ISI. The subject needs to process the time and pitch related information from the sound sequence and project it to the motor structures of the brain.

However, in an effortful listening scenario, subjects often fail to devote sufficient attentional resources to tapping tasks, which affect the anticipatory time control process. On the whole, in this study, the current research investigates how pitch and temporal information processing under two different interval timings affect tapping performance, which includes synchronization accuracy and motor timing, motor variability, and listening effort. This study employs a synchronized tapping paradigm. The results of this research signify the importance of tone pitch and interval timing in relation to effortless motor learning, which comprises listening effort, synchronization, tapping behavior, and motor behavior. To achieve efficient synchronization, better synchronization must be achieved with less listening effort. Different neural mechanisms are involved under two different interval timings. One activates the brain regions responsible for involuntary and spontaneous movements, and the anticipation and other activates the region responsible for voluntary movement and important for motor learning. In such a way, pitch and interval-timing are important aspects that must be considered while designing any movement therapy that uses AMS for motor rehabilitation purposes.

Particularly, the research explores the role of tone frequency on AMS under two different interval timings that engage different levels of cognitive processes. In this scientific endeavor, a single comprehensive experiment was conducted with the following specific objectives.

- To investigate the role of ISI in tapping behavior.

- To examine the effect of tone frequency on synchronization error and listening effort.
- To examine the effect of ISI on synchronization error and listening effort.
- To examine the effect of tone frequency on ITI and STI.
- To examine the effect of ISI on ITI and STI.
- To examine the effect of tone frequency on MV and predictive tapping.
- To examine the effect of ISI on MV and predictive tapping.

1.5. Thesis outline

This chapter has highlighted the background and context of the study. It has also specified the purpose and objectives of the current study (in four parts). Chapter 2 discusses the available literature specifying the role of frequency and inter-stimulus interval (ISI) in auditory-motor synchronization, issues related to synchronization error, listening effort, and the changes in tapping behavior and motor timing and variability under two different ISI ranges. This includes a review of previous research on the role of sensory motor synchronization (SMS) and Rhythmic auditory stimulation (RAS) in affecting gait. Chapter 3 reports the methodology of the experiment. Chapter 4 reports Part-1, which investigated how the tapping behavior gets affected in different ways under two different interval timings. Chapter 5 reports Part-2, which investigated how the tone frequency affects listening effort and tapping performance under two different interval timings. Part-3 is reported in Chapter 6, which investigated how auditory perception translated to motor action under two different interval timing. Part-4 is reported in Chapter 7, which investigates how the frequency of tone affects motor variability in a differential manner under two different interval timings. Chapter 8 provides a comprehensive discussion based on the findings. Finally, chapter 9 presents a consolidated account of the findings, discusses the limitations and implications of this research, and describes the scope for future research. Chapter 10 includes the appendix section.

Chapter 2

Review of Literature

2.1. Music and Movement

Music is an enjoyable activity that helps people communicate their feelings and emotions and soothe their body and mind. Some early scholars like Spencer (1890) understood the origin of music as arising from the physical manifestation of feeling due to mental arousal, while others like Darwin (1896) expounded an evolutionary theory of origin of music, in response to sexual selection (Kleinman, 2015). Nevertheless, the origin of music is older than human history itself. Moreover, music has a crucial role in human evolution, by affecting human perception, their cognitive faculties and their relational behaviour. Apart from this, music has a causal relationship with movement in the human body. It has the ability to induce movements in human bodies such that we perceive music while moving (Jensenius, 2007). Due to this unique property, music plays a significant part in movement-based therapies, especially in movement disorders, which are commonly seen in older people. One of the most common movement disorders seen in elderly people is Parkinson' disease (PD).

The basal ganglia play an important role in controlling movements by close association with internal timing and rhythm perception. PD is a neurodegenerative pathological condition in which the loss of dopamine producing neurons in substantia nigra of midbrain, affects the dopamine transmission from this area to the striatum of basal ganglia (Jones & Jahanshahi, 2006). Due to such a dopamine deficiency, a loss of rhythmicity in timed movements, impediment in initiating movement, diminished length of steps and speed, freezing of strides due to impairment in alternating timed movement like walking and other similar abnormalities in movement are observed in PD. The movement impairment in PD may be atoned by engagement of cerebellar networks, as they are spared in PD and depend upon external auditory or visual cues.

Music entrainment is a widely practiced cultural activity with increasingly recognized pro-social and therapeutic benefits. Music encourages us to move with the beats, demonstrating humans' extraordinary ability to detect and create rhythmic impulses. The underlying functional mechanisms, meanwhile, are unknown. One school of thought, which dates back to Darwin, holds that the underlying pathways are primitive and rooted in the human brain evolutionary earliest subcortical regions.

Recent studies suggest that, even in the absence of visible locomotion or desire to move, rhythm perception is a complicated cognitive process involving temporally accurate communication between cortical sensory and motor regions.

2.2. Perception-Action coupling and its link to body movement

In a variety of cognitive contexts, interactions between perception and behavior have been studied. As a result, different responses to the issue of connection between the input and output sides of cognitive and behavioral activity are connected to each other should not be shocking. According to ideomotor theory, the relation between perception and behavior is based on shared representational resources or structures (Hommel et al., 2001; Prinz, 1997). In short, ideomotor theory, supposes that actions are illustrated by their perceptible effects. While the sensory and motor domains are initially assumed to be incompatible, they become compatible when motor entries are converted into sensory signals (Prinz, 1984).

As a result of representational overlap or similarity, action and perception may cause or obstruct each other (Kornblum et al., 1990; Prinz, 1990, 1997). The ideomotor approach has gained a lot of empirical support in the last two decades, which can be roughly clustered into three main issues (Hommel & Elsner, 2009): (1) how is action knowledge interpreted, (2) how is knowledge of action obtained, and (3) What role does this knowledge have in the selection of action and regulation. In this context, action awareness involves details about objectives, movements, and the linkage between movements and goals. Furthermore, declarative action knowledge, which is frequently regarded as clear and conscious, as well as technical knowledge for action, which is frequently assumed to be intuitive and unconscious, may be included in action knowledge (Prinz, 2014; Schack & Mechsner, 2006).

The ideomotor framework is based on the common coding principle which states that action information and perception are produced and retained in the common representational medium (Prinz, 1990, 1997). As a result of representational overlap or similarity, perception and behavior can cause or interfere with one another. A variety of studies back up this hypothesis. Perception, on the one hand, has been shown to influence actions. As demonstrated by research on ideomotor movements and emulation (e.g., Knuf et al., 2001; Meltzoff & Prinz, 2002), watching others' acts can either stimulate or hinder the display of expected actions (Brass et al., 2001).

In contrast, perception has been found to be influenced by action. This implies that, action preparation and development can improve (assimilation effect, see Craighero et al., 1999; Symes et al., 2008) or impede (contrast effect, see Müsseler & Hommel, 1997; Zwickel et al., 2008) the stimuli's perception that are close to action features. Furthermore, information of the connection between goals and gestures is collected in the form of a bi-directional association) of behavior and their associated perceptual consequences (e.g., Elsner & Hommel, 2001; Herwig & Waszak, 2012). Classical or operant conditioning principles tends to govern the learning of these action–effect relationships. Action–effect relationships, for example, may be applied to other occurrences with similar characteristics (e.g., Beckers et al., 2002; Hommel et al., 2003).

2.3. Sensory-Motor Synchronization (SMS)

Sensorimotor synchronization (SMS) can be understood as a kind of 'referential behavior' (Pressing, 1999) where an action is timed to a foreseeable extrinsic event known as the referent. In indirect way, SMS may be called as a temporal synchronization of internal motor rhythm with external rhythm which is essential for any musical event. In a musical band, musicians must try to align and synchronize their actions with the other partners' actions of the band. In addition, orchestral players must imitate the composer's gestures. Also, during practice sessions, to time their actions, classical musicians very frequently use a metronome.

Soldiers march to the sound of music, and artists perform with the beats. People create temporal supposition (a sort of an intuitive, intrinsic synchronization) when they hear music, and they might move in sync with the beat. Animals, unlike humans, do not instinctively show synchronized movement in response to auditory or visual stimuli, and no successful attempts have been made to teach them to do so (Patel et al., 2005).

The faculty to engage in SMS at a variety of tempi is also be peculiar to humans, and it may have driven the development of music and language in an evolutionary context (Merker, 2000). The task of tapping ones' finger in response to an auditory series of tones or clicks is widely used in laboratory studies of SMS. Different modes of movement (e.g., finger flexion or movement of limb with no contact compared tapping on a firm surface with contact), various stimulation modes, and different forms of coordination all lead to the variety of SMS tasks.

SMS research dates back to 1910 (Dunlap, 1910; Miyake, 1902; Stevens, 1886; Woodrow, 1932), but Paul Fraisse in the 1950s–1970s and John Michon (his dissertation of 1967) were the most significant developers of the field. In recent years, there has been a significant rise in research activity, making this analysis timely. The objective is to provide an understanding of theories and observations from the viewpoint of the one who is passionate about music performance and the skills that go with it. As a result, the finger-tapping task takes precedence over research on contact-free movement of the limb, since it is potentially more important to creation of musical sound, especially on keyboards and percussion instruments (the ones producing sound when struck). The information-processing theory, which usually caters to responses interpreted as a distinct time sequence, is one of the key theoretical approaches to SMS. In addition, while information-processing methods seek to explain hypothetical internal mechanisms underlying actions, theory, which is more commonly applied in engineering applications than in psychology science, lies in the middle of the spectrum (Jagacinski & Flach, 2003).

Since physical tapping causes distinct events, the researchers who use that model focus on information processing (and vice versa), while the researchers of dynamic systems prefer continuous movement activities. Movements ordered as a sequence of discrete contacts tend to necessitate more evident temporal regulation than continuous movements (Delignières et al., 2004; Zelaznik et al., 2002), and thus can engage various brain circuits, especially the cerebellum (Spencer et al., 2005). Paced or timed finger movements (without making contact with the surface) are frequently less symmetrical than finger movements which are not paced, making an oscillator model more challenging to handle (Balasubramaniam et al., 2004). There are several aspects of SMS in executing a finger tapping task to an isochronous sound sequence, such as synchronization error, tap to tap variability, tapping behavior, and motor timing.

In order to achieve synchronization with the tone, individuals try to adjust their internal time keeper with period of isochronous sound sequence and in an attempt to do so, they usually perform trial and error process which leads to the variability in inter-tap interval (ITI). This phenomenon occurs because, initially they undergo learning process and the variability helps in gaining this learning. Once they acquire learning, then after sometime variability decreases. In order to achieve synchronization, sometimes their tapping precedes the tone and at other times succeeds the tone. This occurs as it is impossible to achieve exact synchronization due to nerve conduction, and hence, tapping behavior also gets affected.

In every SMS Task, 3 stages are involved: perception, cognition and action. Sensation is a part of perception process. The coupling between perception and action will depend upon how much an individual exhaust their cognitive resources. Since listening is involved in AMS, so it is likely to interfere in the tapping task.

2.4. Rhythmic-Auditory Stimulation

Rhythmic auditory stimulation (RAS) is based on the oscillator-entrainment paradigm, which states that rhythmic processes in motor system synchronize with rhythmic time-keeper networks in the auditory system. Rhythmic stimuli, such as music or a metronome controls the time-keeper networks in the background. Among the first to explore these models were Safranek et al., 1982 (rhythmic cuing's modulatory effect on electromyogram (EMG) in arm movements) and Mandel et al, 1990 (EMG feedback versus rhythmic feedback in rehabilitation of stroke gait). Thaut et al. (1991) studied the impact of rhythmic cuing on arm movements through EMG analysis in much greater detail. Additionally, Thaut et al. (1992) also investigated stride characteristics and EMG patterns in healthy people's gaits and discovered improvements in stride symmetry and EMG patterns, in particular the variability in amplitude of muscle contractions during movement.

When individual stroke, cerebellar, and transverse myelitis patients were measured, the results were identical (Thaut et al., 1992). Based on their findings, they began exploring Rhythmic Auditory Stimulation (RAS) as a motor rehabilitation tool and standard RAS gait training method was developed, which are now used in neurologic music therapy (NMT) and other rehabilitation settings.

RAS extension to upper extremity and balance training has been successfully implemented and are now being used in clinical settings (Thaut, 2005). Following a CVA, RAS Gait Rehabilitation is needed. Five subjects suffered thrombotic or hemorrhagic episodes in the right hemisphere, usually in the middle cerebral artery distribution. To assess their baseline walking capacity, patients walked six meters at their own pace. During the second walk their baseline gait cadence was matched to external auditory cueing to see whether immediate entrainment effects can arise without training. The same trial was conducted three times, each one separated by a week. The majority of the participants in this study displayed a strong pattern of auditory-motor synchronization. On the paretic hand, stride time symmetry and stride length symmetry increased greatly, as did weight bearing time. Stride time variability was significantly reduced.

On EMG, participants' paretic and non-paretic limbs displayed a more balanced muscular activation pattern. On the paretic side, there was also a decrease in integrated amplitude variability of EMG. RAS also led in a significant decrease in lateral displacement and a significant rise in vertical centre of mass displacement, resulting in a smoother forward gait trajectory (Prassas et al., 1997). Following that, a six-week daily training research was done with 10 stroke patients in a RAS gait training group and ten stroke patients in a matched control group utilising traditional physical therapy gait training (Thaut et al., 1997).

In Thaut et al. (1997), patients were enrolled in the study after walking for five meters with hand-held assistance in an average of 15 days following a CVA. The RAS group showed a non-significant trend for improved stride symmetry and a slightly stronger increase in gait velocity and stride duration. The RAS group increased their gait velocity by 164 percent compared to 107 percent in the control group. RAS increased stride length by 88 percent compared to 34 percent for traditional physical therapy. It is worth noting that RAS's extensive gait pattern shifts in immediate entrainment, as well as in clinical training, lead us to believe that RAS plays a bigger role in motor control than a pure pacemaker. RAS affects and enhances positional and muscular regulation by augmenting timing as the primary coordinative control structure in the generation of complex movement sequences.

2.5. Role of auditory information on rhythmic auditory stimulation

In light of the aforementioned concerns, it is clear that RAS plays a crucial part in rehabilitation training, but what function does auditory information (available in metronome) such as pitch and tempo play is a relevant question. The main component of a metronome is a beat, with which individuals normally synchronize their movements, but other aspects of the music also influence the manner and intensity of their movements.

The bass drum is one of the most important musical elements in today's dance music. Van Dyck et al. (2010) investigated the impact of bass drum's dynamics on subjects during dancing. An environment like that of a club was created, and 100 participants (both adolescent and adults; and 50 males and 50 females) were invited for dancing in five-person clusters to a 10-minute and 30-second musical mix that contained six songs, three bass drum solitary pieces, and some transitional music. Each song was made up of a single part, which was repeated thrice, every time with a distinct bass drum dynamic level. Video capturing and motion detection technologies were used in the study. A motion tracking device was used to measure head movement while motion sensing acceleration data was collected from the hips. The study inferred that the dynamic shift of the bass drum has a considerable impact on the pace of the movement during dancing. It appears to be a normal phenomenon that dancers are unaware of. These findings suggest that the bass drum's increasing significance in contemporary dance music is more than just a stylistic choice; it also has a favorable impact on the quantity of movement produced by the dancers. External auditory cues can be coordinated with our motor movements. For this externalization, we can utilize a number of gestures, such as finger tapping, head banging, or a complicated combination of numerous movements, as seen in dancing. People may be urged to perform certain moves in synchronization with the music, but they can also do it on their own consciously or unconsciously. Also, it is a common observation that when listening to sounds and music, infants engage in greater rhythmic activity than when listening to words (Zentner & Eerola, 2010).

The fact that music motivates us to move is one of the foundations of the music cognition theory, which views body movement as an active personal involvement of a human perceiving music and considers the human body to be the primary component of musical perception and comprehension (Leman, 2008). The movement that accompanies music in most cases represents structural properties of the music, such as pitch, tempo, and timbre, while the periodicity of the movements is usually related to the meter. Other elements of music may have an effect on the form or volume of movement (Leman, 2008). There are currently components in stereo systems that are designed to enhance the lower tones in music. Excessive use of low frequency tones has become standard practice, particularly in mainstream dance music.

The human ear can detect tones as low as 20 Hz (Plack, 2005), and young males in particular seem to have a preference for low tones that can be heard both by body contact and through the ear (McCown, 1997). Experiments have been conducted to decide why so many people appreciate better bass in music. Hellbruck (1984) discovered that the male preference for bass enhanced music is due to the average larger size of the male ear canal, using music played through earphones. Female subjects consider noises to be louder than male subjects because loudness decreases with rising ear canal length. As a result, male subjects can withstand louder sounds than female subjects. The relationship between bass drum volume and corporeal articulations is less well understood. Despite its apparent connection to dance music, the bass drum can have a significant influence on human movement.

Hove et al. (2014) showed that subjects tapped in synchronization with a pacing sequence similar to that employed in EEG experiments, where the higher or lower tone occurred 50 ms early on occasion (deviants). Participants' taps were substantially altered in time considering a lower tone that was 50 ms too early as opposed to a higher tone that was 50 ms prior. According to these findings, the lower pitched stream has a significant influence on motor synchronization to a polyphonic auditory input. When people dance in groups of five to contemporary dance music, their bodily activity changes in response to the complex level of the bass drum, according to data analysis. The average amplitude of the participants' hip movements usually increases as the bass drum level is raised. This may be clarified by the fact that when the bass drum is more prominent, subjects focus more on the music's pulse. This results in quicker and shorter movements, which are most noticeable in hip movements. When the pulse is less pronounced, however, subjects make larger and slower movements.

In the light of above-mentioned issues, it is obvious that, there is a role of auditory information such as pitch in affecting auditory-motor synchronization (AMS) and music strongly compels us to move to its rhythm. There is temporally precise communication between cortical sensory & motor regions but none of the study is tried to see the AMS from effort point of view which is involved in it under different interval timing and time window. Till now AMS to isochronous sound sequence has been investigated in terms of tapping performance analyzed by synchronization error (SE) but not in terms of mental effort associated with it. This is because efficient AMS involves not only tapping performance, but also requires mental effort as two different individuals having different cognitive abilities and learning can demonstrate the same performance with their differential effort. For example, people with lower cognitive capacities will put in more effort to initiate cognitive activity and achieve comparable objective performance levels as the ones with better cognitive talents. (Oberauer, 2005). But efficient synchronization would be the one that can be achieved with less effort (Sparrow et al., 1999; Sparrow et al., 1998).

In order to have synchronization with the tone, it is needed for an individual to listen the isochronous sound sequence actively, and then he/she will have to comprehend and retrieve the temporal information sequence and translate it into inter-tap interval (ITI). So in this process, listening and comprehension are the most important aspects of AMS.

Listening is a multidimensional activity, and it is possible that auditory-motor synchronization necessitates auditory-cognitive connections. Working memory, attention, and processing speed are all cognitive domains involved in listening, according to the FUEL model (Phillips, 2016). An effort is necessary to accomplish a task that demands active listening, such as synchronized tapping in response to an auditory input, which is frequently characterized as the attention and cognitive resources required to comprehend (hear, attend, and understand) the speech (McGarrigle et al., 2014). Effort is distinct from attention in that it is a response to fulfill the attentional demands necessary to complete any work at hand (Strauss & Francis, 2017). An increase in effort, according to Sarter et al. (2006), is defined as the conscious and motivated activation of attentional resources to overcome impediments (such as spectral deterioration of signal, stimulus features) to task performance. Since listening is involved in the tapping task so listening effort will increase demand on cognitive resources and would be reflected in tapping performance in terms of compromise in it.

Unfortunately, there has been relatively less research explicitly examining the role of frequency in affecting AMS in its entirety (which includes listening effort) under two different interval timing. Miyake et al. (2002) in his study, found that temporal information processing in the ISI range between 1800 to 3600 ms is influenced by top-down attention, but it is unaffected by top-down attention in the 450 to 1500 ms range. Different temporal information timeframes are linked to various behavioral processes and interpreted by various neural systems (Buhusi & Meck, 2005). The motor system is regulated by two timing systems, one separated by 1-second range (Lewis &

Miall, 2006) and the other by milliseconds (Merchant & Georgopoulos, 2006). Weber percentage of perceived duration changes at the 1-second border (Grondin, 2014), and cognitively influences sub-second and supra-second timing output differently (Rammsayer & Lima, 1991). The neural implementations of sub and supra-second time vary, according to neuroimaging investigations (Lewis & Miall, 2003b; Wiener et al., 2010). Many regions in the motor system, such as the primary motor area, supplementary motor area, and primary somatosensory area, are activated and govern motor control in the sub-second timing system. Working memory and attention are aided by the supra-second time system (Lewis & Miall, 2006), which includes the posterior parietal, prefrontal cortex, and basal ganglia.

Despite recent neuroimaging research highlighting the distinctions between the sub- and supra-second timing systems (Pouthas et al., 2005; Jahanshahi et al., 2006; Lewis & Miall, 2006; Wiener et al., 2010; Hayashi et al., 2014), the prevalence of concordance between both the systems is unclear. If the goal time is sub-second, supra-second, or peri-second (i.e., about 1 second), we can undertake psychological timing activities. However, it is impossible to provide seamless timing across various temporal ranges in absence of a transit zone between sub-second and supra-second range. This put-up many question with regard to peri-second processing. Weiner et al. (2010) found the existence of certain areas triggered in supra-second and sub-second range. Neuroimaging studies such as fMRI showed that different areas of the brain are associated with event duration on timescales spanning from milliseconds to seconds. These finding show that mechanism involved in sub-second and supra-second have some links but no fMRI research looked at how these two different neural timing mechanism function for peri-second range.

A monotonic shift in BOLD behavior throughout sub- and supra-second time intervals was hypothesized in these fMRI studies, and findings reveal some connections between the sub-second and supra-second timing systems. Nevertheless, no neuroimaging researchers have examined into how these two cerebral timing systems work during peri-second time intervals.

Since the peri-second and supra-second time windows recruit two separate cognitive timing mechanisms, there's a good chance that tapping behavior, tap-to-tap variability, and perception-action coupling (all of which are different aspects of AMS) can vary depending on the time frame. Different levels of top-down focus are needed in these two time windows. The sub-second time window requires less top-down attention and is generally automatic, whereas the supra-second time window requires more cognitive energy and is controlled by a centralized process. Since effort is needed to satisfy attentional demand for completing a task, the results are likely to vary when the intervals are timed differently. Since synchronization, motor timing, motor variability, and tapping actions are all related and interconnected, they all rely on temporal information processing, which differs depending on interval timing. As a result, it is most likely that the aforementioned AMS parameters would be influenced by tone frequency under different interval timing, but in what way is the subject of this study.

2.6. Listening Effort and Pupil Dilation

Routine listening task is frequently marred by acoustic issues that degrade the auditory signal (Mattys et al. 2012). External reasons of acoustic difficulty include background noise, competing voices, and speakers with a different accent (Van Engen &Peelle 2014). Even if the external auditory input is perfectly clear, hearing loss reduces the quality of information reaching the listener's perceptual system. Contrary to cognitive demand, listening effort includes the resources and the energy spent by a listener in order to meet cognitive needs. A recent consensus paper established the Framework for Understanding Effortful Listening, which tackles many of the intricacies that define verbal communication and listening effort (Pichora-Fuller et al. 2016).

Listening effort has been defined by Pichora-Fuller et al. (2016) as "the deliberate allocation of mental energy to resolve challenges in target pursuit while carrying out a listening task. Researchers suggest pupil dilation as an appropriate test for LE because the term 'pupil dilation' appertains to a decrease in the diameter of the eye's pupil caused by factors other than light, most common factors being emotional or cognitive. The pupil dilation response has long been correlated with cognitive effort tasks (Beatty, 1982; Kahneman, 1973), such as listening effort (Koelewijn, Shinn Cunningham, Zekveld, & Kramer, 2014; Winn, 2016). Since a long time, the pupil dilation response has been linked to cognitive exertion tasks (Beatty, 1982; Kahneman, 1973) like listening effort (Koelewin et al., 2014; Winn, 2016).

Pupil dilation has been associated with the involvement of selective attention (Wierda et al., 2012) and working memory (Goldinger & Papesh, 2012), and also to intermediary concepts like generic arousal connected to task demand (Kahneman, 1973) and engagement (Gilzenrat et al., 2010; Goldinger & Papesh, 2012). Both sympathetic and parasympathetic arousal control pupil dilation - constriction being induced by parasympathetic arousal and dilation by sympathetic arousal (Steinhauer et al., 2004).

2.7. Synchronization Error and Tapping Performance

Human behavior is quite susceptible to errors in life which is most unwelcoming part of human existence and it is usually seen in every aspect of life such as medical procedures, battlefields, transportation and vehicle operation. This happens because of impairment in motor timing skills due to which humans usually find themselves unable to synchronize their timing with external events. Error occurs when response does not meet target. Earlier SMS studies suggest that while syncing with a series of tones, individuals' tap appear to be a few milliseconds ahead of tones rather than symmetrically dispersed across the tone initiation (Miyake, 1902; Woodrow, 1932). This phenomenon is known as anticipation phenomena or negative mean asynchrony (NMA). NMA has been observed in many studies and has paved the way for future studies (see Aschersleben, 2002). Despite this, the causes of the NMA are still unknown. For NMA, several explanations have been proposed but none is comprehensive but each of them has some merit. Differences in nerve transmission periods from the finger to the brain versus the ear to the brain give rise to NMA (Fraisse, 1980).

The fact that the NMA for foot tapping is greater than for hand tapping offered early support for this concept (Aschersleben & Prinz, 1995; Aschersleben et al., 2002; Fraisse, 1980). This disparity could be related to different kinematics for foot and finger movement, which has to be investigated further.NMA appears to be minimal and non-existent in musical contexts and it is seen that participants who are musically trained shows smaller NMA as compared to untrained participants and also shows zero NMA in some cases (Aschersleben, 2002; Repp, 2004b). Furthermore, auditory feedback, rhythmic complexity and subdivision help in minimizing and removing the NMA (Aschersleben & Prinz, 1995; Wohlschläger & Koch, 2000). Thus, NMA occurs only when non-musicians tap in time with a simple metronome.

2.8. Motor Control and Tapping Modes

Individuals must predict the next stimuli from the current state when attempting to synchronize their tapping with the sound sequence. Often in anticipation of the next event, individuals typically tap before the incoming event referred to as negative asynchrony and this form of tapping mode is called anticipatory tapping which involves feedforward motor control. The interdependence of the effectors is preplanned and visible using feed-forward control before sensory feedback from the movement can be used. Positive asynchrony, on the other hand, occurs when top-down attention is engaged because it activates the feedback motor control system, which is a reflexive motor action that results in reactive tapping.

With feedback control, the motor system can respond to sensory data that signals departure from anticipated or normal movement. The distinguishing feature of coordinative feedback control, in contrast to local feedback control, is that motor commands to one effector are dependent on sensory feedback from another effector. After very short latencies (70 ms), coordinated feedback responses can be seen, which are most likely the product of motor cortical control circuits (Pruszynski et al., 2011). The transitions from anticipatory to reactive tapping could indicate a universal feature of human coordination, in which the brain adapts to changes in the environment, such as inter-stimulus intervals, by switching from one mode of coordination to the other.

2.9. Motor Timing and Time Perception

To maintain synchronization with external events, the subject begins to make judgments based on the length of events (tone duration) and the time gap between them, which can be of following two types.; beat-based judgments, and, interval based judgments (Keele et al., 1989).When the subject touches the surface, they do not immediately release their finger but also keep the finger in contact with the surface for very short time intervals. It occurs because the subject attempts to estimate the length of the sound, which is a kind of beat-based time function, perceptually throughout this time period timer, which involves synchronizing events. Following finger release, tapping response occurs with the next tone again. The period between one tap and the next tap (including the period when the finger stays in contact with the surface) is called an inter-tap interval, an index of motor timing. The time interval between taps change as subjects make a series of uniformly spaced taps. Individuals assess the passing of time depending on their personal perceptions of time.

At least two key origins of time perception are identified, namely, judgment of the passing of time and evaluation of interval length. It's worth noting that variables other than physical time influence the subjective perception of interval length. A time period, in particular, can be viewed differently by different people depending on their psychological states or personality traits. Timing abilities are aided by a dispersed network of frontal regions, including prefrontal and premotor equivalents, as well as non-frontal regions including the cerebellum, basal ganglia, and parietal areas, according to previous neuropsychological and neuroimaging studies (Buhusi & Meck, 2005; Lewis & Miall, 2006).

However, as a result of the current limits, the brain areas that process timing-related information are more flexible (Witt et al., 2008) . As a result, internal and external factors may cause resource processing to overlap, resulting in the coupling of timing skills. In this sense, motor timing and time perception might interact in such a way that repeating acts can help with interval timing calibration in a dual-tasking setting (Carlini & French, 2014).

2.10. Motor Variability and Learning

In order to achieve synchronization with the tone, individuals try to adjust their internal time keeper with period of isochronous sound sequence and in the attempt to do so, they usually perform error and trail processes leading to the variability in inter-tap interval (ITI). It's because they're going through a learning process at first, and the variation in ITI aids in learning. After they acquire learning, their variability reduces over time. Motor variability (MV), in this perspective, aids skill learning in the way that genetic variation has aided evolution: an integral element of a mechanism that molds adaptive behaviors through selection by outcomes (Skinner, 1981).

Movements are the result of a network of hierarchically organized motor controllers generating tightly choreographed patterns of muscle activity (Lemon, 2008). From variability in central circuit movement planning to noise in muscle force generation, variability in the motor circuits can occur at any stage. If the motor system can recreate neural activity patterns that produce successful outcomes, it may be able to take advantage of such variability. Variability in the motor systems can occur at any level. In this work, variability in central planning and control circuits termed as 'planning noise' is differentiated from variability in the motor peripheral, which is called as 'execution noise'.

Central circuits, receiving performance-related feedbacks, are somewhat more inclined to provide learning-friendly variability as compared to peripheral circuits, which might be more challenging to replicate and reinforce. At the motor periphery, the variability of the motor system has perhaps been explored the most. Studies have revealed that signal dependent noise in force generation, i.e. trial-to-trial variations that scale proportionally with mean force, reflect a fundamental characteristic of muscle activity. Peripherally derived variability, also known as "execution noise" (Van Beers et al., 2004), may not be ideal for learning-related motor inquiry due to its uncontrollable nature. On the contrary, learning-related motor exploration may be best served by variability in central planning loops. These circuits have higher experience-dependent plasticity and have better access to reinforcement signals (Björklund & Dunnett, 2007; Doyon & Benali, 2005; Nudo et al., 1996; Sanes & Donoghue, 2000; Schultz, 1998).

Chapter 3

Methodology

The study examines the effect of tone frequency and ISI on synchronization error and listening effort. Synchronization error (SE) is measured as the absolute temporal difference (i.e., irrespective of ± asynchrony) between the tone onset and the tap onset and ISI is defined as an interval between the offset of the previous tone and onset of the next tone.

The analyses reported in this dissertation are restricted to data collected prior to Covid-19 Pandemic. The study was conducted abiding by the required precautions of involving human subjects. When the pandemic crisis began, further data collection was suspended and at the time of submitting this dissertation, the prevailing pandemic crisis remains a safety concern.

3.1. Experimental room set-up

3.1.1. Eye-tracking system

The sound sequence clubbed with the video was presented using the Tobii TX30 Eye-tracking workstation. Investigator connected Tobii TX30 to the PC through USB in order to collect data pertaining to pupillary dilation (pupil diameter) for both the eyes at the rate of 30 samples per second. Calibration includes few dots on the screen, and it takes 2-5 minutes. The stimulus is displayed on the screen through the interface of Tobii studio.

3.1.2. Arduino Tapping Board Set-up

Piezoelectric sensor was connected to pin 1 of an analog port of Arduino microcontroller for measuring tapping data. Light Dependent Resistor (LDR) sensor's output was fed to pin 0 of an analog port of arduino microcontroller for synchronization with PC. Investigators placed LDR sensor over the lower right corner of the PC screen. In this arduino tapping board set-up, audio metronome starts with the start of the black screen which is detected by LDR. This generates start signal which received through analog pin 0 and by receiving this signal, it starts taking tapping data through analog pin1.

3.1.3. Participants

Thirty healthy participants (*M* = 24.6 years; *SD* = 4.9; range = 18-35 years) voluntarily participated in the study. They received a T-shirt for their participation in this study. There were equal number of participants assigned in randomized manner to short ISI group (1000 ms) and long ISI (2000 ms) group.

3.1.4. Design

The Design of the study is a 2 (Frequency: LF vs. HF) ×2 (Inter-stimulus interval: Short ISI vs. Long ISI) with the second factor as a between-groups variable.

In this experiment, LF is 100 Hz comes under low-frequency range, and HF is 1000 Hz comes under the high-frequency range. The inter-stimulus interval was operationalized having two categories, i.e., (i) Short ISI (1000 ms), (ii) Long ISI (2000 ms).

3.1.5. Experimental Task

Participants performed the experimental task that required them to tap on tapping board with their dominant index fingers in response to the tone present in isochronous sound sequence. The tone occurred 72 (i.e., 2 *initiation tones* + 70 *execution tones*) times in both ISI levels out of which first two tones were initiation tones whereas the remaining were execution tones.

Investigators created two isochronous sound sequence with the help of Audacity software (Mazzoni & Dannenberg, 2000), under two different ISI, i.e., (i) Short ISI (1000 ms), (ii) Long ISI (2000 ms).

The duration of isochronous sound sequence was 2.15 minutes under short ISI and was 4.30 minutes under long ISI (i.e., double the duration of short ISI). This duration does not include initiation period in which 2 initiation tones were presented i.e., 1 second and 2 second in short ISI and long ISI respectively.

As stated above, there were 70 occasions of presenting the execution tone to which participants were required to tap in a synchronized manner (i.e., tapping needs to be coincide with the presentation of the execution tone). These 70 execution tones were divided into two blocks. There were a thirty-five low-frequency tone in one block and thirty-five high-frequency tones in second (see Fig.2) . For loudness equalization of high and low-frequency tones, the investigators used a replay gain algorithm, an inbuilt plugin in audacity (Wolters et al., 2010).

Participants were randomly assigned to isochronous sound sequence (i.e., short ISI vs. long ISI). Further, under each ISI, there were two blocks of tone frequency (i.e., LF vs. HF) that were counterbalanced in the experiment. In a single session, each participant went through two blocks of an isochronous sound sequence consisting of 35 tones, each having low-frequency and high frequency and were separated by a fixed ISI. Initial six tones among 35 tones were for initiation trails and rests were for the experimental trials. Among 30 participants, 15 participants had gone through short ISI, and the other 15 participants went through long ISI.

In order to maintain steady intensity, participants used noise-cancellation headphones (68 SPL) during the experimental task. They also used Disposable Head Cap considering hygiene requirement. For measuring the tapping response, the investigators used Arduino-based customized tapping board and for measuring the pupil diameter, Tobii tx-30 eye tracker having a sample rate of 30 Hz was used.

This Arduino-based tapping board was customized (developed for this study) to synchronize with the PC via a light-dependent resistor, which was placed over the lower right corner of the PC screen to ensure synchronization of tapping data with the execution tones.

Tapping data were recorded using Arduino based customized tapping board connected to a Dell PC running Arduino IDE software; light-dependent resistor was placed on the lower right corner of the PC screen to achieve synchronization between tapping board and PC.

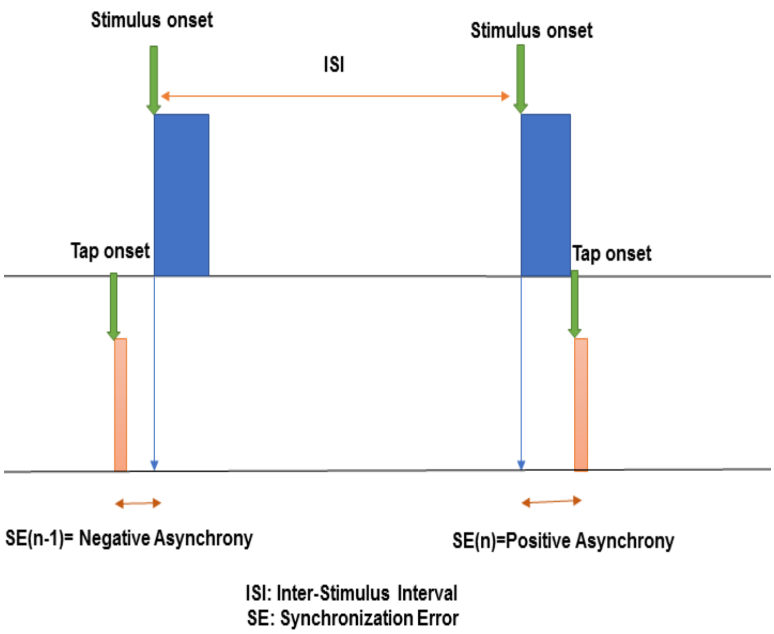


Figure.1. Temporal relationship with respect to onset of the stimulus and onset of tap. Information assessed during thi experiment is stimulus onset and tap onset. Synchronization error (SE) is calculated as absolute temporal lag between the onset o stimulus and the onset of tap the time difference was described as the Inter-stimulus Interval (ISI) between two successiv stimulus onsets. The tone duration is 200 ms.

For measuring the synchronization error, the investigator used Arduino based customized tapping board, and for measuring the pupil diameter, Tobii tx-30 eye tracker having a sample rate of 30 Hz was used.

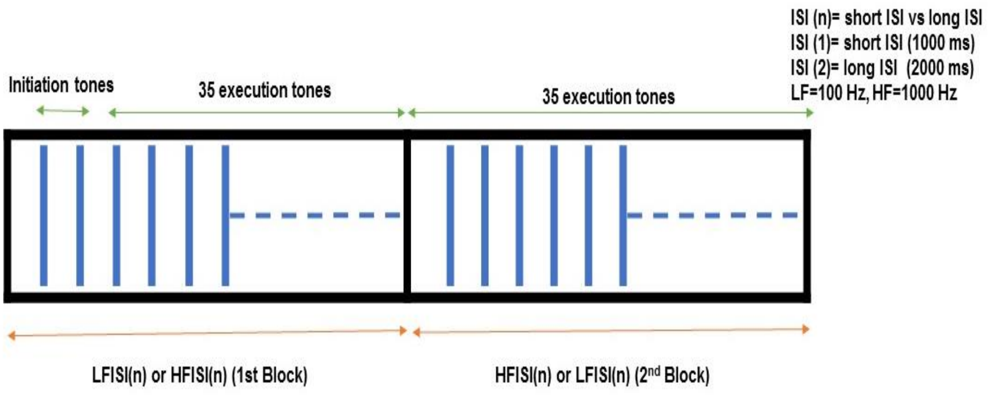


Figure 2: Presentation of Stimulus

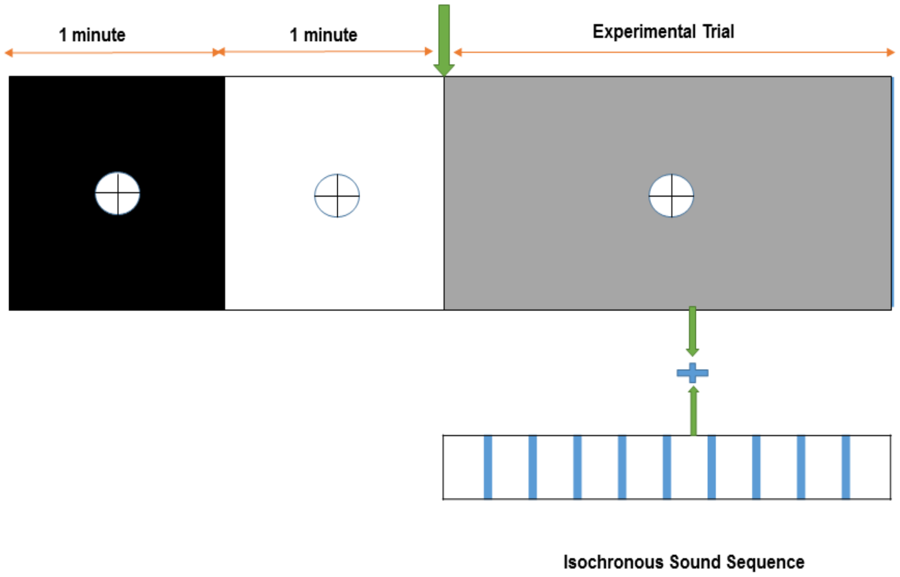


Figure. 3 Measuring the physiological pupillary response. Participants were presented a black screen following by the white screen for one minute each for the purpose of eliciting maximum pupil dilation and maximum pupil constriction. A gray screen is used for the rest of the experimental trials.

As the pupil diameter is very sensitive to the light so to ensure the sensitivity of pupil diameter to the listening effort, the investigator created a video stimulus consist of three frame in Filmora software in such a way that for the initial first minute, frame with black color was displayed so that participant got maximum pupil dilation, then a white image frame was displayed for the next 1 minute, to create maximum pupil constriction then gray image were meant to elicit inter-mediate pupil size for rest of the experimental trial period. Then isochronous sound sequence (created via audacity) has been merged with the frame of gray image in filmora software to create final stimuli and presented in front of the participant via Tobii studio software. A fixation cross was presented in the center of the computer screen to keep participant's gaze fixed throughout the session. Arduino based customized tapping board (developed for this study) was made to synchronize with the eye tracker via a light-dependent resistor, which was placed over the lower right corner of the PC screen to ensure low latency between recording time of tapping data and playback of sounds.

The Subjects were involved in two different tasks concurrently, the tapping task and the listening task. Participants were asked to sit in front of the table, and to tap their dominant index finger on the tapping board along with the tone of the isochronous sound sequence in a synchronized manner without compromising his/her gaze towards the fixed cross.

3.1.6. Dependent Variables

Anticipatory tapping and reactive tapping: Anticipatory tapping and reactive, tapping is examined in terms of asynchrony of tapping responses (i.e., negative- and positive- asynchrony). Negative asynchrony and positive asynchrony are differentiated in terms of the nature of taps with reference to deviation from exact coinciding point. In other words, the sign of asynchrony is determined based on whether a tap is executed *before* or *after* the occurrence of the tone. For instance, those taps which occur before the tone will be considered as negative asynchrony in contrast to positive asynchrony i.e., taps which occur after the tone.

Synchronization error (SE) is measured as the absolute temporal difference between the stimulus onset and the tap onset and inter-stimulus interval defined as an interval between the offset of the previous tone and onset of the next tone in isochronous-sound sequence

LE is described as the purposeful distribution of mental resources to overcome obstacles to the objective concern while performing a listening task (Pichora-Fuller et al., 2016).

Inter-Tap Interval (ITI) is calculated by taking temporal difference between successive taps and used as an index of motor timing in context of AMS.

To investigate how the subjects translated the ISI into action, subjective translation index (STI) was computed by dividing each inter-tap interval (ITI) by ISI. Further, the mean STI was calculated for each participant.

STI> 1 indicates, subjects over translated the ISI.

STI<1indicates, subjects under translated the ISI.

Motor variability (MV) studied through the coefficient of variation of ITI (CV) under the peri-second and supra-second range. CV was calculated by dividing the standard deviation of ITI by mean (ITI) for each participant.

Hypothesis and objective	**Dependent variable as referred to as hypothesis**	**Measure of dependent variables**	**How are measures related to measures of dependent variables?**
To investigate the effect of tone frequency and Inter-stimulus interval on Tapping Behaviour	Anticipatory tapping and Reactive tapping	Anticipatory and Reactive tapping were measured as number of taps which occurred before and after the tone	Taps with negative sign is related to anticipatory tapping and positive sign is related to reactive tapping
To investigate the effect of frequency of tone on listening effort and synchronization error under two different inter-stimulus intervals	Tapping Performance Listening Effort	Pupillary dilation is taken as measure of listening effort Synchronization Error (Absolute measure of asynchrony) is taken as a measure of tapping performance	Increase or decrease in pupillary dilation is related to higher and lesser listening effort. Lesser and higher synchronization error is related to increased or decreased tapping performance.

To investigate the effect of tone frequency on perception-action coupling and motor timing under two different interval timing	Motor timing , Subjective Translational Index (STI)	Inter-Tap Interval is taken as a measure of Motor timing and STI	STI> 1 indicates, subjects over translated the ISI. STI<1in dicates, subjects under translated the ISI.
To investigate the effect of tone frequency on motor variability and predictive tapping under two different interval timing	Motor Variability (MV), Predictive tapping	Coefficient of Variation is taken as a measure of Motor Variability	Larger variation of inter-taps is related to high motor variability and vice-versa

3.1.7. Independent Variables:

Two independent variables were manipulated in testing these hypotheses.

Frequency is measured as the number of wave cycles repeated in one-second and its unit is Hertz (Hz). It was being manipulated based on the studies which says that processing of temporal information in the brain and its coding in the central nervous system (CNS) is unique for low and high different frequency ranges, as an auditory neuron can maintain the temporal fine structure of a low-frequency tone by phase-locking its response to the stimulus (Reichenbach and Hudspeth, 2012) unlike a high frequency tone. For high frequencies, the temporal fine structure is not preserved in neural responses, below 300 Hz. auditory fibers fire almost in phase lock with stimulus frequency, and beyond 800 Hz, this phase-locking decline (Reichenbach& Hudspeth, 2012).

Inter-stimulus Interval (ISI) is described as the temporal difference between two successive stimulus onsets. It was being operationalize and manipulated based on the Miyake et al.(2002) which says that ISI is expected to affect the temporal information processing in its two ranges, namely short ISI (450 ms<ISI< 1500 ms) and long ISI (1800 ms<ISI <3600 ms), respectively. It is because that, one time scale lies in between 450 ms to 1500 ms, which does not involve subjects' top-down attention and other one lies in between 1800 ms to 3600 ms which involve subjects' top-down attention.

3.1.8. Procedure

Prior to engaging the participants, the investigators secured the approval of the experimental protocol from the Institute Human Ethics committee, Discipline of Biosciences and Biomedical Engineering which monitors the ethical compliance issues. Participants were explained about the task that they had to perform in the experiment so that they could take an informed decision regarding their consent. After this, a vital information sheet for recording personal details (such as, name, date of birth, interest in music etc.) was used by the experimenter.

Practice session preceded experimental task for each participant in order to ensure that they had understood the task. Each participant was engaged individually for the experimental session. The experimenter sat beside the participant (Figure. 4) for empirical observation through a parallel headset connected to the audio splitter. After data collection, participants went through a debriefing session in which each participant was asked some probing questions about his/her experience during the experiment.

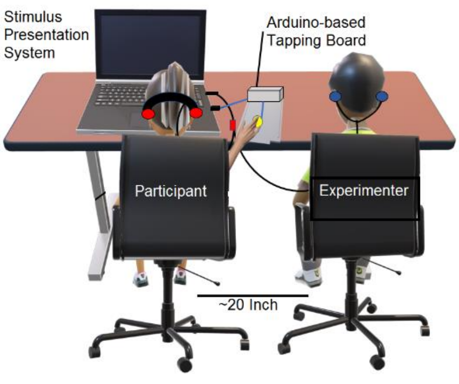


Figure.4. Experimental Setuo

Chapter 4

Tapping Behavior and Inter-Stimulus Interval in Isochronous Sound Sequence

4.1. Introduction

Sensory-motor synchronization (SMS) refers to the rhythmic coordination between perception and action (Repp, 2005). For instance, people move their body parts in response to musical beats in which actions are coordinated in the prediction of an external event (Pressing, 1999). It is important to have situation where both action and referent are periodic for SMS to occur (Rajendran & Teki, 2016). This is because periodicity of the referent increases its predictability where referent refers to auditory stimuli in a sequence comprises of series of auditory stimuli. Therefore, SMS requires temporal coordination between motor rhythm and external auditory rhythm (Repp & Su, 2013). In the context of Rhythmic auditory stimulation (RAS), Thaut (2010) have shown that auditory rhythm affects the motor system because both systems share crucial connectivity at cortical, subcortical, and spinal level. Auditory system plays an important role in processing temporal information quickly and precisely in order to pass it to motor structures in the human brain, thus creating synchrony between rhythmic sound sequence and motor response (such as tapping in this study) (Lezame & Hernandez, 2020).

This study highlights two aspects of asynchrony with respect to execution of tapping response. When individuals try to synchronize their tapping with the sound sequence, it is necessary to predict the next stimuli from the current state. Often in anticipation of the next event, individuals usually tap before the incoming event referred to as negative asynchrony. In contrast, positive asynchrony occurs with the involvement of top-down attention because it activates the feedback motor control system which is a reflexive motor action and it leads to reactive tapping.

The primary goal of the current study was to investigate the anticipatory tapping and reactive tapping in isochronous sound sequence under short ISI and long ISI levels. An important concern with respect to the role of temporal aspect is that of understanding how ISI affects SMS. Miyake et al. (2002) reported that there is different levels of involvement of attentional demand under ISI range of 1800-3600 ms and ISI range of 450-1500 ms such that the former is found to have more attentional demand.

As the ISI range of 450-1500 ms demands lesser attentional resources, this ISI range will switch motor control to feed-forward mode because it doesn't require voluntary control (Klein, 1976; Seidler, 2004). In the feed-forward model of motor control, cerebellum receives an efferent copy (internal copy of an efferent movement inducing signal) of a motor command from the primary motor cortex for a motor-to-somatosensory prediction (Malapani et al., 1998). This, in turn, induces the cerebellum to make predictions about the sensory consequences of these motor instructions, which prepares the muscle-skeletal system for movement execution. Incoming sensations are then contrasted with the predicted sensations. A constructive match between incoming and predicted sensation also maintains the same pattern for the succeeding motion. This whole process leads to anticipation. The forward model of motor control presumes that the cerebellum receives an efferent copy (internal copy of an efferent movement inducing signal) of a motor command from the primary motor cortex for a motor-to-somatosensory prediction. This, in turn, induces the cerebellum to make predictions about the sensory consequences of these motor instructions, which prepares the muscle-skeletal system for movement execution. Incoming sensations are then contrasted with the predicted sensations. A positive match maintains the same pattern for the succeeding movement also. But if there is a mismatch between the incoming and predicted sensations, then an alert signal is sent back to motor cortical and sub-cortical areas for movement corrections and calibration of forward model.

In the long ISI range of 1800-3600 ms which is also called the supra-second range involves more attentional demand which activates feedback motor control system (Gooch et al., 2010; Koch et al., 2006; Teki et al., 2012). Seidler et al.(2004) had a study in which people had to hit different size objects on the LCD screen with a moving joystick, and found that movement made to larger targets relied more on feed forward-motor control whereas movement made to smaller targets relied on feedback motor control. Smaller targets require more attention than larger targets due to increased precision requirements for aiming.

The cerebellum plays an important role in the sub-second interval timing while on the other hand, basal ganglia work in both the timing range (i.e., sub-second timing and supra-second timing). Studies have shown that in both the timing ranges, different motor networks get activated during temporal processing.

The thesis of Brown that memory traces decay over a brief period until some threshold is reached, and the memory becomes unreliable, is also in agreement with other studies. According to memory-prediction theory, prediction depends on memory so it is obvious that anticipation will be affected after that threshold and then the subject will rely only on the sensory feedback that will activate the feedback motor control system. Investigators envisage that, if the ISI increases beyond the sub-second timing, then memory trace will decay very rapidly and become unreliable after crossing the threshold. This will hamper the anticipatory timing control, and response will only depend on the present instead of the past. So there will be reactive response instead of being anticipatory.

In the light of the current discussion, it is conceivable that temporal aspect of ISI will play a role in SMS (tapping performance). The point is that if the involvement of attention is different in two different ISI namely, short ISI and long ISI they are expected to have significant difference in terms of SMS. As anticipation happen when there is lesser involvement of attention. So it is expected that anticipatory tapping will happen under short ISI whereas reactive tapping will happen under long ISI.

4.2. Results

Anticipatory tapping and reactive tapping: Anticipatory tapping and reactive, tapping are examined in terms of asynchrony of tapping responses (i.e., negative- and positive- asynchrony). Negative asynchrony and positive asynchrony are differentiated in terms of the nature of taps with reference to deviation from exact coinciding point. In other words, the sign of asynchrony is determined based on whether a tap is executed *before* or *after* the occurrence of the tone. For instance, those taps which occur before the tone will be considered as negative asynchrony in contrast to positive asynchrony i.e., taps which occur after the tone.

A paired-sample *t*-test was performed to compare anticipatory tapping and reactive tapping under the short ISI (1000 ms) in terms of their occurrences. The analysis revealed a significant difference between anticipatory tapping (M = 59.5, SD = 24.27) and reactive tapping (M =12.4, SD = 24.27); $t(14) = 3.75$, $p = 0.002$, two-tailed (Figure. 6). This suggests that there were more cases of anticipatory tapping as compared to reactive tapping under the short ISI. To the contrary, this trend is absent in case of long ISI as there was no significant difference between anticipatory tapping (M = 22.6) and reactive tapping (M = 49.3). An independent-sample *t*-test was performed to compare anticipatory tapping between short ISI and long ISI. There was a significant difference with respect to anticipatory tapping between short ISI (M = 59.5, SD = 24.27) and long ISI (M = 22.66, SD = 25.53); $t(28) = 2.537$, $p = 0.017<0.05$, two-tailed. Further, an independent-sample *t*-test revealed a significant difference in reactive tapping between short ISI (M = 12.46, SD = 24.27) and long ISI (M = 49.33, SD = 25.53); $t(28) = -2.35$, $p = 0.017<0.05$, two-tailed.

Figure. 6 show anticipatory tapping and reactive tapping under both ISI levels. On the whole, the results provide statistical evidence that there is higher occurrence of anticipatory tapping under short ISI as compared to long ISI. In contrast, higher occurrence of reactive tapping is observed under long ISI than short ISI. Figure. 5a and Figure. 5b show the asynchrony variation of participants for all the 70 occasions under short ISI and long ISI respectively. In the figures 5a and 5b, line plots below 0 (which is ref line) represents negative asynchrony and line plots above 0 represents positive asynchrony. Thirteen participants showed anticipatory tapping in terms of negative asynchrony and only two participants showed reactive tapping in terms of positive asynchrony under short ISI.

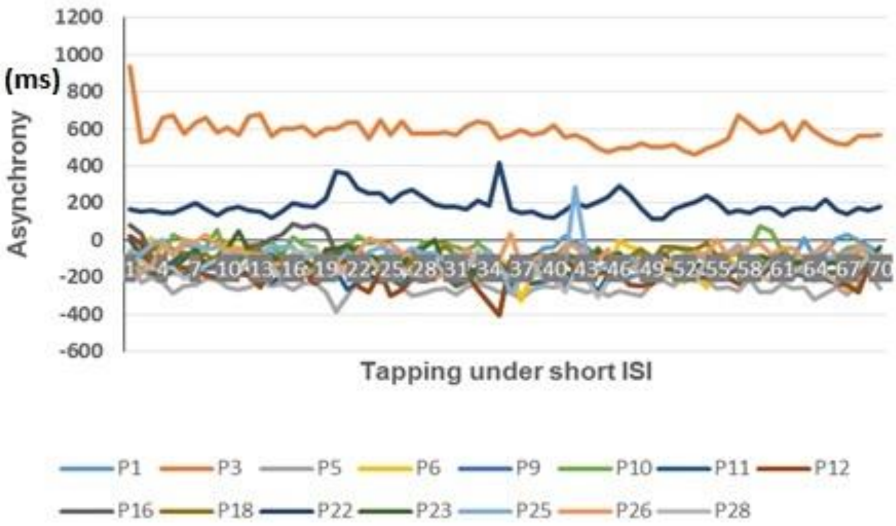


Figure. 5a. Asynchrony variation under short ISI (execution tones)

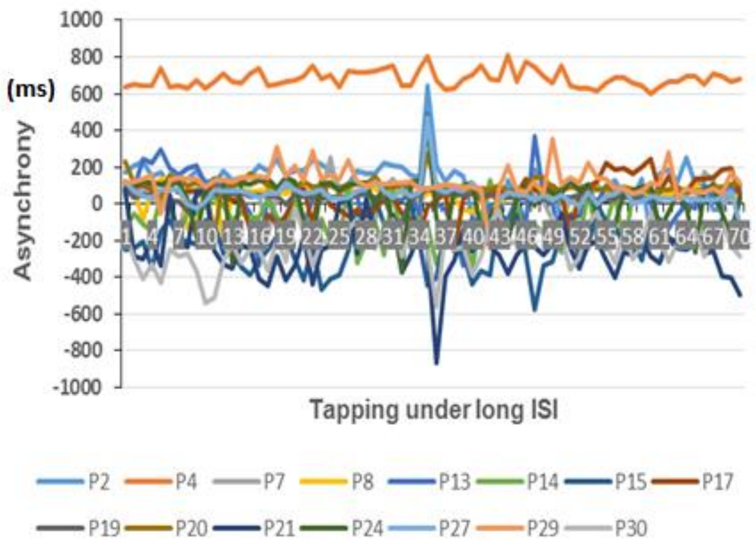


Figure. 5b Asynchrony variation under long ISI (execution tones)

It implies that anticipatory tapping is more prevalent in short ISI and Figure. 5b shows that participants showed both reactive and anticipatory tapping because none of these two was prevalent in long ISI. But there were more number of cases for reactive tapping in long ISI as compared to short ISI.

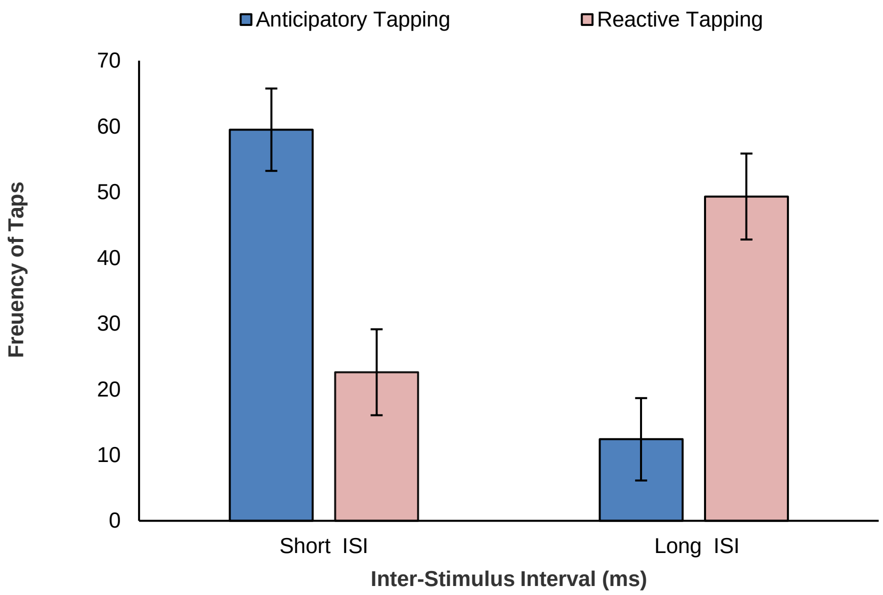


Figure. 6.Anticipatory Tapping and Reactive Taping under Short and Long ISI

4.3. Discussion

This study emphasizes the issues pertaining to SMS in isochronous sound sequence. ISI is an important aspect concerning response asynchrony.

According to the results, anticipatory tapping outperforms reactive tapping under short ISI. Under short ISI, this refers to the predictability of a synchronised motor movement with an anticipated stimulus (Repp, 2005; Ascherzleben & Prinz, 1995; Ascherzleben, 2002). Conversely, a reflex action is indicated by a larger occurrence of reactive tapping under long ISI compared to short ISI.This occurs in reaction to a stimulus associated with feedback motor control relying on sensory feedback for coordinating movements (Iosa et al., 2015).

Individuals under short ISI tend to tap ahead of the tone's beginning by anticipating it, rather than tapping in a symmetric uniform distribution around initial (irregular) inputs (Repp, 2005; Pressing, 1999;Repp, 2013; Fraisse, 1980; Wohlseqer, 1999).As a result, the researchers suggest that anticipation is crucial in the short ISI because of its ability to predict the next event while synchronising with the external auditory stimuli and it occurs in less absence of attentional resources because it adheres to the belief that in the short ISI, less attentional demand is required (Miyake et al., 2002).

In this context, one important question to consider is how anticipation occurs when top-down attention is less involved. The argument is that anticipation occurs when activities rely on future projections, expectations, or beliefs in addition to prior and current stimuli (Schuboltz, 2007). As a result of the current study, the role of sequential decision making in auditory-motor synchronisation becomes very significant, and anticipation is a key aspect of this process.

When individuals make decisions in in sequential decision making (Deco, 2011) based on anticipation involving memory (i.e., retrieval of information by attractor dynamics if a given input is sufficiently similar to a stored attractor state) then first stimulus is held in memory and then compared with the subsequent similar stimuli. Attractor neural activity of second stimuli depends not only on the present stimuli but also on the sensory memory of the first sound event (Deco et al., 2010). Anticipation is difficult without memory, which plays a key part in prediction (Hawler et al., 2009) and provides the individual with previous experience to anticipate the upcoming stimulus from its current state. Auditory sensory memory is the shortest element of memory and the critical first stage in auditory perception (Alain et al., 1998). It allows the cognitive system to retain sensory information representations after the prior auditory stimuli have passed. It works as a buffer for the stimuli received through the auditory senses, which are only stored for a short time (but accurately).

In case of short ISI, next stimulus occurred before the decaying of sensory memory, so this past sensory memory helped to predict the next event via anticipation mechanism and the'prediction' is a preplanned behavior that activates the feedforward motor control mechanism and it is unaffected by peripheral feedback (Pisotta & Molinari, 2019). Feed-Forward motor control is an open-loop motor control system that regulates fast, ballistic movements before any sensory information is processed (Seidler et al., 2004; Iosa et al., 2015). It necessitates understanding of the incoming stimuli in advance. Deafferentation studies on monkeys and cats with sensory nerves severed from their spinal cords have been performed for this type of control (Seidler et al., 2004; Iosa et al., 2015). It requires prior knowledge of the upcoming stimulus. For this type of control, deafferentation studies have been done on monkeys and cats whose sensory nerves were cut off from their spinal cords (Taub et al., 1966). After recovering from the deafferentation procedure, monkeys started normal behavior after losing sensory information from their arms. That is why this motor control mechanism is called 'automatic'.

In this automatic mechanism, some reflex loops require lower brain areas and are independent of perceptual processing active control but can be affected by prior commands; while the other loops are directed by the spinal cord with no active control of the brain.

In case of long ISI, participants showed reactive tapping behavior. Reactive response is a reflexive response which is not based on anticipation. It is because of memory traces decline (i.e., decaying) over a short span of time until a certain limit is reached, and that memory becomes unreliable (Lu & Sperling, 2003; Brower, 1993). Under such circumstances, it becomes difficult for an individual to retain representation of auditory information for a longer duration because memory decays very quickly, in the region between 200-500 milliseconds (less than a second). Moreover, in the absence of anticipation mechanism, top-down attention plays a critical role which switches an individual's motor action to feedback mode (in contrast feed-forward mode) for error compensation. In other words, feedback motor control relies on sensory feedback for coordinating movements. Reactive tapping is a reflex activity which occurs in reaction to a stimulus due to feedback motor control. As reactive tapping in long ISI (2000 ms) occurs under supra-second range and anticipatory tapping in short ISI (1000 ms) occurs in 1-second range, it indicates that as the ISI increases from 1-second to supra-second range, there is a transition in tapping from anticipatory to reactive. In future studies, it will be important to examine issues concerning sensory-automatic processing in 1-second range which is a transit zone (Rammasyer & Ulrich, 2011; Repp & Su, 2013) between sub-second range and the supra-second range.

4.4. Conclusion

In summary, this study demonstrates significant occurrence of anticipatory tapping and reactive tapping under short ISI (1000 ms) and long ISI (2000 ms) respectively.

In Anticipatory tapping, the feed-forward motor control mechanism plays a critical role (unlike the case of reactive tapping in which the motor control is dependent on the feed-back motor control mechanism). In terms of implications, the findings of this study indicate that ISI in isochronous sound sequence should be given due consideration in designing movement studies and incorporate use of the same for therapeutic purposes such as rhythmic auditory stimulation (RAS).

Chapter 5

Listening Effort and Tapping Performance in Auditory-Motor Synchronization

5.1. Introduction

In general, music is known to stimulate movement in humans. While listening to music, the human body moves spontaneously. The rhythmic nature of the music induces body movement (McGarrigle et al., 2014; Keller & Rieger, 2009). Janata et al. (2012) found that when participants were made to tap in response to music, they moved their body parts. Movement execution through music involves the coupling of the sensory system and motor system, which leads to sensory-motor synchronization. When the body parts are in motion, muscles convey the information to the brain, and the brain constantly revises the instructions and sends them back to the muscles.

Sensory-motor synchronization (SMS) is defined as the rhythmic coordination between perception and action and is a form of referential behavior in which actions are coordinated in the prediction of an external event (Repp & Su, 2013). SMS occurs in many contexts but especially in music and dance. When SMS occurs in response to a sequence of auditory stimuli, then it is referred to as auditory-motor synchronization (AMS). It has been studied in the form of finger tapping to a sequence of auditory stimuli. In the context of rhythmic auditory stimulation (RAS), Thaut et al. (2010) have shown that auditory rhythm affects the motor system because both systems share crucial connectivity at cortical, subcortical, and spinal level. The auditory system plays an important role in processing temporal information quickly and precisely in order to pass it to motor structures in the human brain, thus creating synchrony between rhythmic sound sequence and motor response.

In order to achieve synchronization in a spontaneous manner (in response to rhythmic sound sequence), predictability of external events (precisely the beat and the rhythm patterns) is important. It is thought that predictability leads to effortless actions, but this may not be true in all circumstances. Predictability might also occur without either smoothness or effortlessness (Montero, 2011). Effortless bodily movements involve reduced effort or an appearance of reduced effort, and such an effort can be physical or mental (Montero, 2016). In synchronizing with the isochronous sound sequence, some attention and cognitive resources are needed to retrieve temporal information present in the sound sequence in order to have AMS. This is called listening effort (LE) which is a type of mental effort (McGarrigle et al., 2014).

Till now AMS to isochronous sound sequence has been investigated in terms of tapping performance analyzed by synchronization error (SE) but not in terms of mental effort associated with it. This is because efficient AMS involves not only tapping performance, but also requires mental effort as two different individuals having different cognitive abilities and learning can demonstrate the same performance with their differential effort. For e.g., individuals having less cognitive ability will expend more effort to initiate cognitive activity and achieve similar objective levels of performance than those of higher cognitive abilities (Oberauer, 2005; Hess, 2011).But efficient synchronization would be the one that can be achieved with less effort (Sparrow et al., 1999;Sparrow et al., 1998)

Tapping performance and listening effort of an individual will depend on how efficiently he/she can process the temporal information present in an isochronous sound sequence. In the present context, temporal information processing can be modulated by two factors; one is the tone frequency present in the isochronous sound sequence. The frequency of sound is the speed of vibration and determines the pitch of the sound. Frequency is measured as the number of wave cycles repeated in one-second, and its unit is Hertz (Hz). The processing of temporal information in the brain and its coding in the central nervous system (CNS) is unique for low frequency and high-frequency spectrum. For low-frequency tone below 300 Hz. auditory fibers fire almost in phase lock with stimulus frequency, and temporal fine structure is preserved in neural responses. But for high frequencies beyond 800 Hz, this phase-locking declines (Reichenbach & Hudspeth, 2012). Another factor is the inter-stimulus interval (ISI), which is an interval between the onsets of successive auditory stimulus thought to require a different level of top-down attention in its two different ranges. Miyake et al.(2002) in his study, found that temporal information processing in the ISI range between 1800 to 3600 ms is influenced by top-down attention, but is unaffected by top-down attention in the 450 to 1500 ms range.

The primary objective of the current study is to examine the synchronization error and listening effort in relation to the inter-stimulus interval (ISI) and tone frequency (pitch of an auditory stimulus) in an isochronous sound sequence.

Listening is multidimensional, and AMS likely requires auditory-cognitive interactions. According to the FUEL model (Phillips, 2016), cognitive domains involved in listening include working memory, attention, and speed of processing. In performing a task that requires active listening such as synchronized tapping in response to an auditory stimulus, an effort is required, which is often defined as the attention and cognitive resources needed to comprehend (to hear, attend and understand) the speech (McGarrigle et al., 2014). Effort is different from attention in such a way that it is a response to meet attentional demands required to perform any task given at the moment (Strauss& Francis, 2017). Sarter et al. (2006) stated that an increase in effort is defined as deliberated and motivated activation of attentional resources to overcome obstacles (such as spectral degradation of stimulus, stimulus characteristics) for performing task.

There is the possibility that the spectral content of auditory stimuli, which is the frequency will affect the LE. People experience high-frequency sounds as sharp and unpleasant compared to low-frequency sounds, which is widely known to be as bass sound

(Kurakata et al., 2013; Ramachandran, 1996), and high-frequency sounds are perceived to be louder than low-frequency sound at the same intensity levels (Moore, 2012). To retrieve and process the temporal information from unpleasant high-frequency sounds might be more effortful, which is likely to increase the listening effort (LE).

Apart from the tone frequency, which is part of stimulus characteristics, LE also depends upon the demand of the listening situation (Phillips, 2016). Now since the demand for the listening situation is higher in the ISI range between 1800 to 3600 ms (due to involvement of top-down attention), so there are high chances that LE will be higher in this range as compared to ISI range less than 1500 ms. For the measurement of LE, pupillometric methods have been widely used to show changes as a function of listening task demand in clinical (Just et al., 2003) and non-clinical trials. Since cognitive processes are involved in listening and have a relationship with the pupillary response, hence, task evoked pupillary response quantifies listening effort in auditory task (Beatty, 1982; Hyönä et al., 1995).

Apart from LE, there is synchronization error (SE), which is to be seen in a broader perspective in the context of AMS. Synchronization error is a measurement of tapping performance, which manifests in the form of negative and positive asynchrony. Negative asynchrony occurs when individuals attempt to synchronize their finger taps with the tone by tapping a few milliseconds before the onset of the tone (Aschersleben, 2002). According to Fraisse (1980), the occurrence of negative asynchrony is due to differential nerve transmission times from finger to the brain and from the ear to the brain. On the other hand, when the taps succeed the tone, it is known as positive asynchrony.

Individuals need to retrieve the temporal information between successive tones, which will pass to the motor structures in order to create synchronization with the tones. Therefore, temporal information processing becomes a crucial factor in maintaining synchronization. As mentioned before, it will depend on the response of auditory nerve fibers towards the tone frequency and the range of ISI. For low-frequency tone below 300 Hz, auditory fibers fire almost in phase lock with stimulus frequency and temporal fine structure is preserved in neural responses, but for high frequencies beyond 800 Hz, this phase-locking decline (Reichenbach & Hudspeth, 2012). So it is most likely that individuals will show lesser synchronization error and better temporal alignment with the low-frequency tone than high-frequency tones. As far as ISI is concerned, temporal information processing in long ISI range is influenced by attention. Hence, in this range, tapping and listening compete with each other in the consumption of attentional resources, unlike short ISI, which is not influenced by attention. So more effort in listening will reduce the cognitive resources required for tapping task and will hamper the ability of a person to synchronize with tone. So, there is a possibility that SE will be larger in long ISI as compared to short ISI.

On the whole, the study emphasizes the effect of tone frequency and ISI on synchronization error and listening effort.

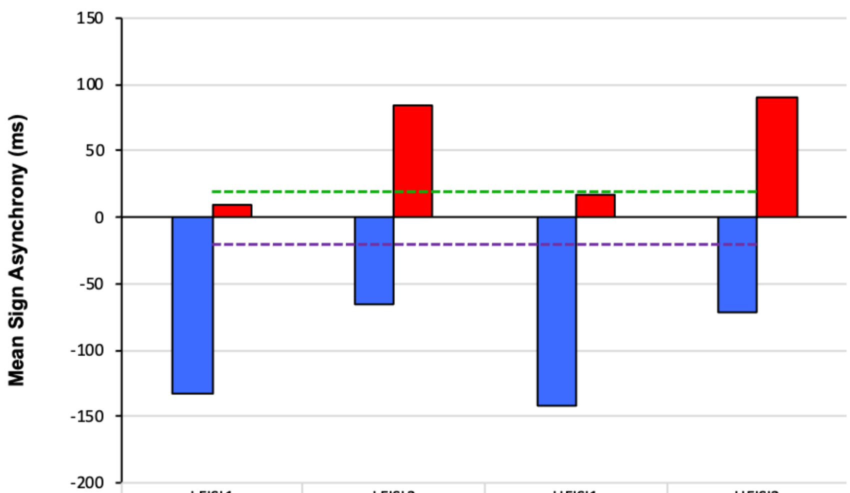
In the current study, ISI was operationalized by creating two isochronous sequences of different inter-stimulus intervals, one is a short ISI of 1000 ms (within the range of 450 ms to 1500 ms), and another one is a long ISI of 2000 ms (within the range of 1800 ms to 3600 ms). We divided each sequence into two blocks, each block having 35 identical tones, and the tone frequency was different in both the blocks. Based on previous explanations, investigators selected a low frequency (LF) tone of 100 Hz (< 300 Hz) and a high frequency (HF) tone of 1000 Hz (> 800 Hz). Synchronization error and Listening effort were measured as dependent variables and the following hypotheses were tested in this study:

- There is lesser synchronization error with LF tones than HF tones
- There is higher synchronization error involved in long ISI as compared to short ISI
- There is lesser listening effort involved with LF tones than HF tones

- There is higher listening effort involved in long ISI as compared to short ISI

5.2. Data Processing and Statistical Analyses

Participants deviated from ideal benchmark line (0 ms) of synchronization that leaded to asynchrony. Asynchrony (As) was calculated by subtracting the tap onset time from the tone onset time which resulted in the form of negative asynchrony -As and positive asynchrony + As which indicates that participant executed the tap before and after the tone (see Figure.7) under each condition. Then absolute measure of asynchrony (|As|) was taken to estimated overall error without regard to its sign. Since, it was difficult for each participant to get exactly synchronized because of nerve conduction delay, so threshold line of ± 20 ms was set based on previous studies (see Figure.7) which says that tap precedes or succeed the click by about 20 to 80 ms. Investigators selected a minimum value between 20 and 80 ms and set a threshold of 20 ms. With the set threshold, synchronization error (SE) was calculated (i.e., subtracting threshold of 20 ms from |As|) under each condition and this was submitted to the statistical analyses.



	LFISI1	LFISI2	HFISI1	HFISI2
NMA	-132.192	-64.6961	-141.437	-71.452
PMA	8.984	83.81	17.2026	90.406
benchmark-	-20	-20	-20	-20
benchmark+	20	20	20	20

Figure.7: Mean signed asynchrony under different conditions. NMA and PMA refers to Negative and Posititve Mean Asynchrony indicates each participant executed tap before the tone and after the tone under each conditions. Green and Violet dotted line refers to benchmark Line.

Listening effort was analyzed by measuring pupillary dilation. Mean pupil diameter was calculated by averaging the pupil responses for each trial recorded over period of time for LF and HF conditions separately under both ISI

5.3. Results

An alpha value of 0.05 was used to test for statistical significance in all the analyses.

5.3.1. Tapping performance

The analysis revealed a significant effect of frequency (LF, *M* = 123.62 ms, and HF, *M* = 141.39 ms, see Figure. 8), $F(1,28) = 8.187$, $p = 0.008$, partial $\eta^2 = 0.226$. But the main effect of ISI was not significant $F\ (1, 28)=0.899$, $p>0.05$, partial $\eta^2 = 0.001$. Thus, there was no overall difference in synchronization error between short ISI (*M* =129.782 ms) and long ISI (*M* = 135.239 ms) (see Figure.8). Levene's test indicated equal variances across groups under both LF (F =2.305,p =0.140) and HF (F =1.930, p =0.176) conditions.

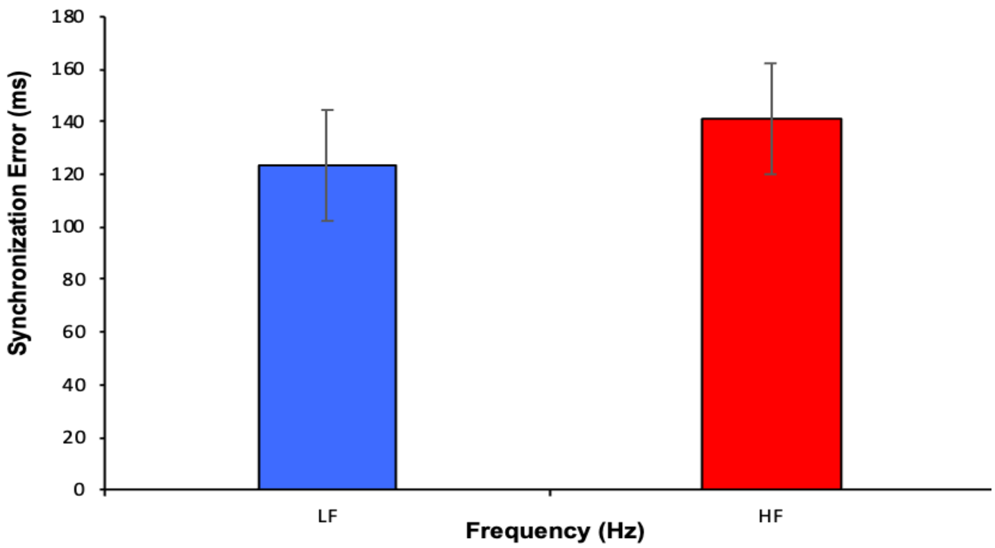


Figure.8: Synchronization Error. Error bars denote Standard Error.

5.3.2. Listening Effort

The analysis revealed a significant effect of frequency (LF, *M* = 4.3055 mm and HF, *M* = 4.4349 mm, see Figure. 8), $F(1,28) = 5.037$, $p = 0.032$, partial $\eta^2 = 0.154$. These results revealed that less listening effort was observed with low-frequency tones than high frequency tones. But the main effect of ISI was not significant $F(1, 28) = 0.177$, $p > 0.05$, partial $\eta^2 = 0.006$. Thus, there was no overall difference in listening effort between short ISI (*M* = 4.316 mm) and long ISI (*M* = 4.429 mm) (see Fig.9). Levene's test indicated equal variances across groups under both LF ($F = 0.035$, $p = 0.852$) and HF ($F = 0.885$, $p = 0.352$) conditions.

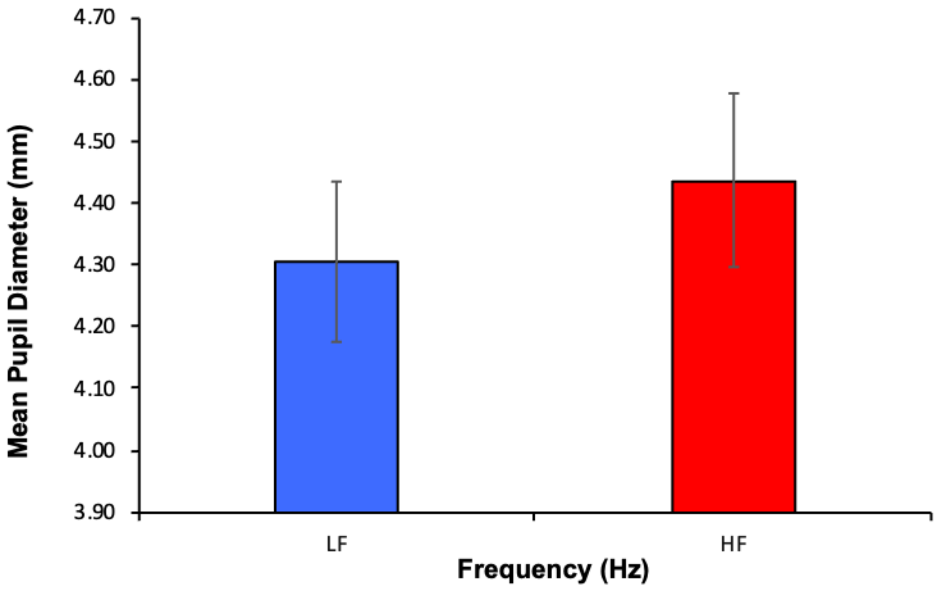


Figure.9 Pupil dilation with respect to listening effort. Error bars denote Standard Error.

5.4. Discussion

This study focused on issues pertaining to AMS in isochronous sound sequences. Tone frequency and ISI are important aspects concerning synchronization error and listening effort. As per the results, participants showed more temporal alignment with low-frequency tones as compared to high-frequency tones under short ISI and long ISI. It implies that low-frequency tones demand less attentional resources for efficient temporal information processing as compared to high-frequency tones. In addition, lower listening effort was observed with low-frequency tones as compared to high-frequency tones under both ISIs.

During the experiment, it was observed that participants were unable to catch the first beat but they were able to anticipate the next beat under both conditions of ISI. While debriefing, no variation in tapping was reported because of the fixed ISI in the isochronous sound sequence. In terms of the synchronization error and listening effort, no significant effect of ISI was found which rejects our hypothesis that was based on the fact that short ISI (which comes under 1-second range) and long ISI (which comes under supra-second range) has different involvement of attention. This implies that both level of ISI engages same cognitive process and this corroborates with Rammsayer's finding which states that 1-second range involves both cognitive and perceptual process (Rammsayer& Troche, 2014).

We observed a significant effect of tone frequency on synchronization error and found that it was significantly higher with HF tone (1000 Hz) than the LF tone (100 Hz), which comes under bass range. It implies that subjects' tapping were comparably more aligned with the low-frequency sound than high-frequency tone. This finding corroborates with the previous research which asserts that when people are presented with music, they have a strong tendency to align their body movement with the low-frequency sounds (Burger et al.,2013) and increased body movements are associated with an increase in loudness of low-frequency sounds (Van Dyck et al., 2013), and there is one transcranial magnetic stimulation (TMS) study that also substantiates the findings. That study shows that songs which have greater spectral flux in the low-frequency region activate the motor system to a greater extent (Stupacher et al., 2013).It is evident that sounds with low-frequency aim maximal actuation of the basilar membrane close to the apex of the cochlea whereas, high-frequency sounds initiate the basilar membrane close to the base of the cochlea, and intermediate frequency sounds actuate the membrane film at halfway split between these two boundaries. Recording of signals in the auditory tracts of the brainstem and in the auditory receptive fields of the cerebral cortex demonstrates that explicit brain neurons are actuated by explicit sound frequencies. In this way, the significant technique utilized by the sensory system to identify distinctive sound frequencies is to decide the positions along the basilar membrane that are most actuated.

Following frequency rules, low-frequencies are segregated by the volleys. These low-frequency sounds in the range of 20 to 1500 to cycles for each sound can cause volleys of impulses resonating at same frequencies, and cochlear nerve transmits these volleys into the brain's cochlear nuclei. It is assumed that the brain's cochlear nuclei can recognize the different frequencies of the volleys (Bamne et al. (2015)).

Such findings expand previous ones in relation to the synchronization activity relationship between low-frequency spectral flux and general music-induced movement (Burger et al., 2013). In the debriefing session, the majority of participants reported significant discomfort in listening to the high-frequency tones than to the low-frequency tones. This implies that the listening task was more demanding with the high-frequency tones than the low-frequency tones.

The finding of the current study showed that there was a significant effect of tone frequency on participants' listening effort. Listening effort was significantly higher in the case of HF as compared to LF which indicates that participants felt less listening discomfort with the low-frequency tones than with high-frequency tones, which was in an agreement with participant's subjective reports of listening effort. This is because low-frequency tones are perceived to be more pleasant and easier to process than high-frequency tones at the same intensity. Hence low frequency tones require less attentional demand for efficient temporal information processing leaving more attentional resources for the tapping task as compared to high-frequency tones. Tapping task with LF tones is more engaging because it requires less effort. With HF tones, the tapping task will be less engaging and people tend to avoid such a task which demands more effort (Song, 2019)

With LF tones, higher performance can be achieved with the less mental effort required for listening. In other words, performance under low-frequency tones is more efficient than under high frequency tones; the link between movement efficiency and tone frequency could be explored in future studies.

5.5. Conclusion

In summary, we conclude that a human's motor response is more easily synchronized with low-frequency tones as compared to high-frequency tones with less listening effort while responding to isochronous sound sequence. In terms of implications, this study suggests that frequency of tone in isochronous sound sequences should be given due consideration in designing movement therapies such as rhythmic auditory stimulation (RAS).

Chapter 6

Perception-Action Coupling and Inter-stimulus Interval in Auditory-Motor Synchronization

6.1. Introduction

Sensorimotor synchronization (SMS) is the ability to coordinate rhythmic movement with an external rhythm or beat, which is of fundamental importance to musical activities (Repp, 2005). The beat is not a physical characteristic of the music itself; rather, it is a perceptual phenomenon that results from the music through beat induction (Rajendran et al., 2018). Beat induction is extracting a periodic pulse from music and synchronizing the body's tempo (Ravignani et al., 2017). Therefore, beat induction reflects the involvement of SMS in environmental rhythmic sensory cues and everyday activities like walking, dancing, or clapping (Blais et al., 2015; Repp, 2005), which is often regarded as a cognitive skill (Honing, 2012). SMS is also a referential behavior(Pressing, 1999) in which an action is temporally coordinated with a predictable external event, the referent, and both the action and the referent are periodic so that the referent's predictability is derived from its regular recurrence.

People typically move to the beat of the music in real-world situations, and their actions are synchronized in advance of a beat. In laboratory research, finger tapping in synchrony with external (usually computer-controlled) stimuli, often an isochronous auditory or/and visual metronome, remains a popular paradigm because of its simplicity and long history (Blais et al., 2015; Jin et al., 2019). Synchrony is ascertained by calculating the mean time delay between each stimulus and its associated tap and the variability of the mean delay over time (Chen et al., 2002). Previous studies on SMS reflect the negative mean asynchrony in adults who are not musicians, indicating that people touch more frequently before the stimuli onset (Repp, 2005).

SMS is known as auditory-motor synchronization (AMS) when it happens in response to a series of auditory stimuli. In rhythmic auditory stimulation (RAS), auditory rhythm impacts the motor system as they share a critical connection at the cortical, subcortical, and spinal levels (Thaut & Abiru, 2010). In order to synchronize rhythmic sound sequence and motor response, the auditory system plays a crucial role in processing temporal information fast and correctly and transmitting it to motor structures in the human brain. The desired motor response results from an overall synchronization of the individual, task, and environment. This synchronization must occur not just within the body but also between the body and the environment through the dynamic coupling of perceptual information into the proper action units (e.g., perceptual-action coupling; Sternad, 2000).

Perception-action processing needs accurate time information for seamless execution. Bueti et al. (2008) emphasized the role of temporal information in perception, the corresponding action, and the necessity of accurate production of temporal intervals in daily life. The ability of a person to estimate time is frequently attributed to a central time-keeping device (Spapé et al., 2022), the paradigm—scalar expectancy theory (Gibbon, 1977; Wearden & McShane, 1988), which has two parts: a pacemaker and an accumulator. The pacemaker "ticks" at an undetermined rate, and the accumulator polls the pacemaker if attention is directed toward making a temporal judgment. The hypothesis uses attentional resource allocation to the timing task or a change in pacemaker rate as its primary mechanism for explaining systematic timing errors. This asynchrony is important in understanding perception-action coupling, where shared pacemakers need to recognize durations to function promptly (Bueti et al., 2008). For instance, the inter-stimulus interval (ISI) is systematically increased or decreased by the rapid periodic presentation of auditory stimuli (Treisman et al., 1990, 1992). The generated ISI will be shown as a motor response or inter-tap interval (ITI). In contrast, the perceived ISI will take place at the sensory level. However, the information does not move immediately from sensory to motor levels, indicating a delay.

People make judgments based on the length of events (the tone's duration) and the interval between them to have synchronization with external events. These judgments can be either beat-based or interval-based (Keele et al., 1989). In a beat-based judgment, an initial series of periodic events establishes an internal beat for an individual that persists after the initial events, which helps predict the following beat. In contrast, individuals make interval-based judgments based on the duration of an interval between events. In this process, temporal memory is compared to other intervals to judge whether they are the same. The relationship between perceptual and motor timing is so intertwined that changes to one affect the other (O'Regan et al., 2017), which occurs due to communication between basal ganglia, cerebellum, and parietal areas (i.e., non-frontal regions) and premotor and prefrontal cortical areas (i.e., frontal regions). Therefore, internal and external stimuli may cause overlapping processing resources and coupling timing abilities (Bueti et al., 2008). Concerning perceptual timing and motor timing, the fundamental question is whether there is a consistent relationship between motor timing and perceptual timing across all temporal ranges or whether variations depend on the temporal range. Top-down attention affects how quickly temporal information is processed, with different amounts of engagement occurring in the sub-second, peri-second, and supra-second time frames (Miyake et al., 2004). Therefore, it is important to examine the role of frequency in perceived ISI, which will, in turn, influence motor timing.

The ISI range is classified into three types based on the diverse attentional resource involvement (Lewis & Miall, 2003). Firstly, the sub-second range (<1000 ms) is brief and can be categorized as sensory-automatic, and they are involved in motor control (Merchant

& Georgopoulos, 2006) and speech generation (Schirmer, 2004). Second is the supra-second range (>1000 ms), which is involved in foraging (Kacelnik & Brunner, 2002) and decision-making (Sohn & Carlson, 2003). According to earlier research, the Weber fraction of the perceived duration or the coefficient of variation varies at the border at about 1000 milliseconds, and cognitive load has a varied impact on timing performance at the sub- and supra-second levels (Grondin, 2014; Rammsayer & Lima, 1991). There has been a report of a third transit zone between these two, the peri-second range (1000 ms), which is influenced by both sensory-automatic and cognitive timing mechanisms. Among the brain regions related to motor timing, frontal and posterior parietal areas of the cortex trigger in the supra-second range, where more top-down attention is involved, whereas cerebellum and basal ganglia trigger in the peri-second range, where less top-down attention is involved (Lewis & Miall, 2003). It is anticipated that there will be strong perception-action coupling and that temporal processing will be influenced by a frequency characteristic of auditory stimuli (responsible for sensation) in the peri-second range. This is because both sensory-automatic and cognitive timing mechanisms are temporally triggered by the peri-second range (Murai & Yotsumoto, 2016).

An individual's tapping performance and listening effort depend on efficiently processing the temporal information in an isochronous sound sequence (Sanjram & Shivhare, 2021). The frequency displayed in the isochronous sound sequence is modulated together with the ISI, which affects how temporal information is processed. Further, the perceived duration of a tone is influenced by frequency because high-pitched tones elicit more tension arousal (Ilie & Thompson, 2006). The expansion and reduction of perceived duration are related to changes in physiological arousal (Droit-Volet & Gil, 2009; Lake et al., 2016). Pure and complex tones, with high-frequency (HF) tones, it is perceived as longer than low-frequency (LF) tones (Burghardt, 1972). Given this knowledge, there is a greater likelihood that tone frequency will influence how the ISI (interval between tones) is perceived, altering the motor timing so that the ISI will either be under or over-translated, manifesting as ITI in the peri-second range. In addition, because HF tones are seen as lasting longer than LF tones, it is anticipated that ISI will be under-translated and over-translated with LF and HF tones, respectively (Jeon & Fricke, 1997).

The current study examined perception-action coupling over two different time scales, including evaluations of the subjective translation index (STI =ITI/ISI). During the perception-action coupling, the perceptual characteristics of external stimuli would impact the action. Mean asynchrony helps to understand how the motor response to a tone is aligned temporally. However, it does not provide information on how the motor timing changes during the perception-action coupling process. The tone frequency affects the perceived translation of ISI into ITI during perception-action coupling. The STI was used to examine the subjective translation of ISI into ITI. As a result, the STI index provides information on how the subject interprets the perceived ISI as ITI (over-translation or under-translation). Because STI offers information about the direction of transformation (of ISI into ITI), it is preferred over “mean asynchrony difference” (i.e., the temporal difference between stimulus and tap onset times). When the STI is greater than 1, the participant over-translated the ISI, whereas STI less than 1 indicates that they under-translated the ISI.

6.2. Results

An alpha value of 0.05 was used to test the statistical significance in all the analyses throughout this paper.

6.2.1. Tapping Analysis

STI was computed by dividing each ITI by ISI. Further, the mean STI was calculated for each participant.

- STI>1 indicates subjects over-translated the ISI.
- STI<1 indicates subjects under-translated the ISI.

A repeated measures ANOVA was carried out for 2 (Frequency: LF vs HF) × 2 (ISI: Short ISI vs Long ISI) mixed factorial design with the second factor as a between-groups variable on STI. The analysis revealed a significant main effect of frequency on STIs. Investigators found that there was a significant difference in the mean STI for LF (*M* = 1.00, *SD* = 0.0017) and HF (*M* = 0.99, *SD* = 0.003) conditions (see Figure 3), F (1, 28) = 4.569, p = 0.04, η_p^2 = 0.14. There was no main effect of ISI (Short ISI, *M* = 1, Long ISI, *M* = 1), F (1, 28) = 2.157, p = 0.153, η_p^2 = 0.072). Levene's test indicated equal variances across groups under both LF (F = 1.277, p = 0.26) and HF (F = 0.013, p = 0.90) conditions.

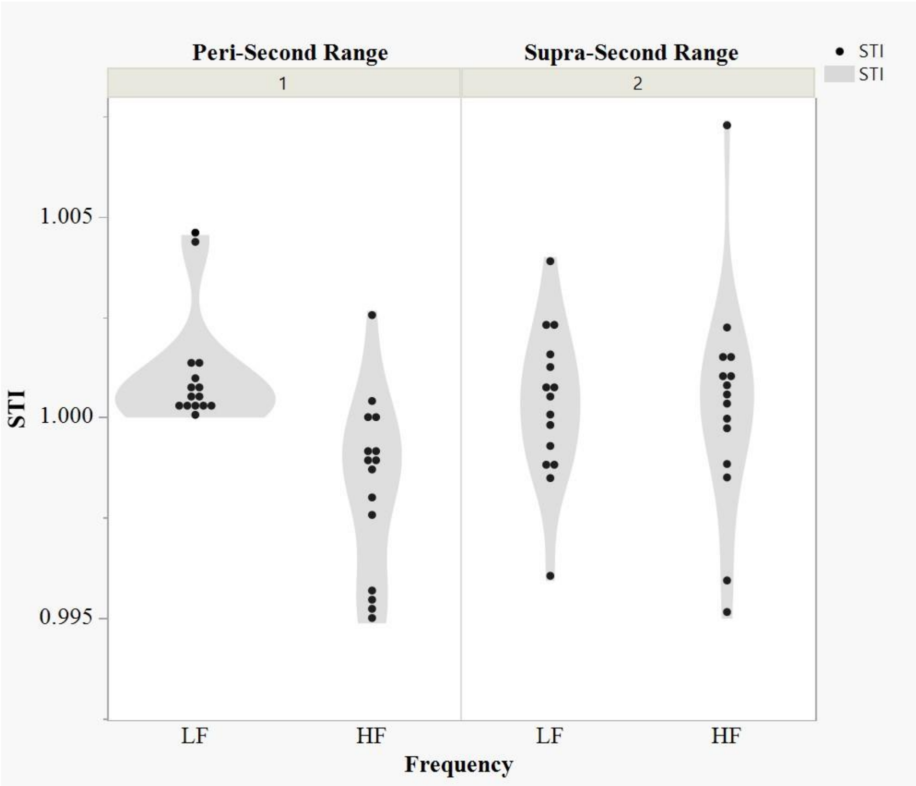


Figure 10. STI under two frequency conditions.

As shown in Figure 4, there was a significant interaction between frequency and ISI, F (1, 28) = 4.277, p = 0.048, η_p^2 = 0.132), such that under short ISI (peri-second range), participants showed higher STI with LF tone as compared to HF tone. With long ISI (supra-second range), this trend was absent. This indicates that under short ISI—with LF tone, participants over-translated ISI (i.e., STI >1), and participant's mean ITI was higher than ISI in contrast to HF tone, where participants under-translated ISI (i.e., STI<1).

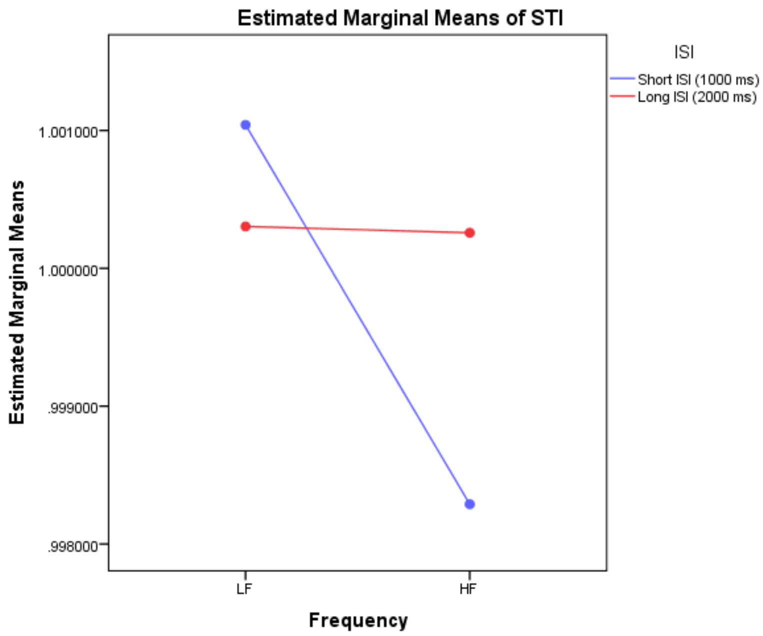


Figure.11 Significant Interaction between frequency and ISI (under Short ISI: with LF, STI >1 in contrast to STI <1 with HF). . Error Bars denotes Standard Error.

Furthermore, a paired sample *t*-test was carried out to analyze the effect of tone frequency on ITI under short and long ISI. Investigators found that there was a significant difference in the mean ITI for LF (M = 1001.04 ms, SD = 1.50 ms) and HF (M = 998.28ms, SD = 2.21 ms) conditions; t(14) = 2.69, bootstrapped p = 0.01 (Figure 5) under short ISI, but this trend was absent under long ISI (Figure 6) since no significant difference was found in the mean ITI for LF (M = 2000.60 ms, SD = 3.81 ms) and HF (M = 2000.51 ms, SD = 5.54 ms) conditions; t(14) = 0.04, p = 0.96. These results suggest that there is a relation between the frequency of the tone and motor timing in the short ISI, and there is a connection between what people perceive (hear) and their actions (motor timing). Therefore, there is a strong coupling between perception and action timing in the short ISI as compared to the long ISI. The mean ITI for LF is significantly more than HF under the short ISI. This implies that participants over-translated the ISI = 1000 ms to 1000.64 ms with LF tone and under-translated to 998.83 ms with HF tone in short ISI (which comes under the peri-second range).

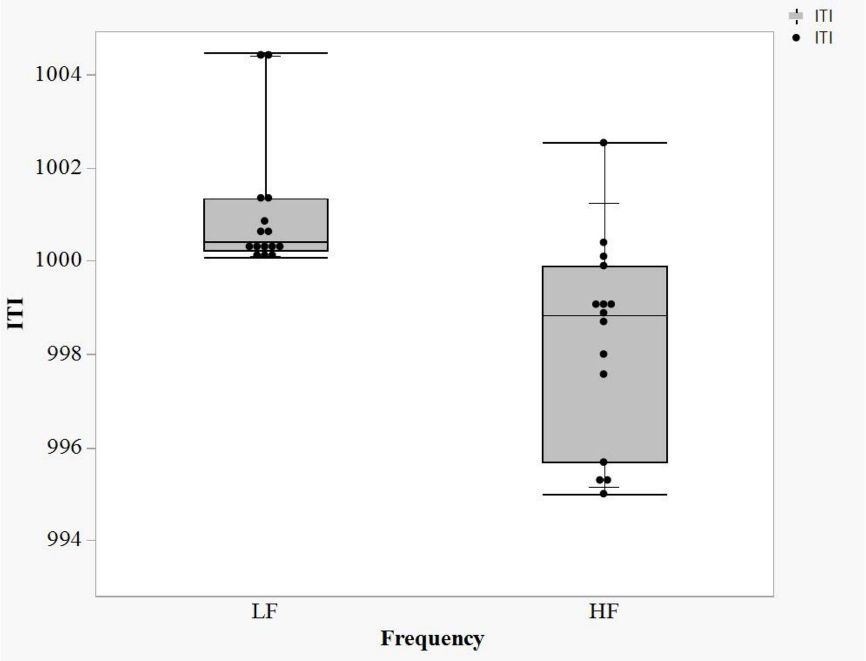


Figure 12. Frequency and ITI under short ISI condition.

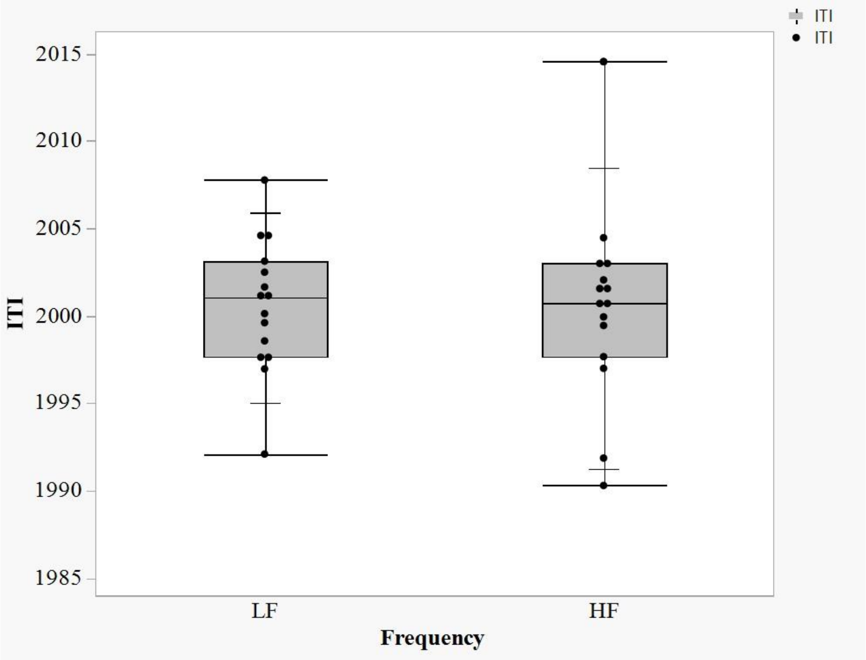


Figure 13. Frequency and ITI under long ISI condition.

6.3. Discussion:

This study used a synchronized tapping task to explore the relationship between perception, cognition, and action in the context of auditory-motor synchronization. It signified how people interpreted information from the ISI into actions that were represented in the ITI and how auditory information—such as the tone's frequency—affected the timing of perception and action over a range of temporal scales. Additionally, it reported the mean ITI (representing motor timing) and STI, which showed how accurately the participant translated the ISI.

The findings also revealed that motor production was perceptually sensitive in the peri-second range. The tone's frequency considerably impacted STI, and there was evidence of a cross-over interaction between the tone's frequency and ISI. Participants over-translated ISI with LF tone and under-translated with HF tone, as evidenced by higher STI with LF tone than HF tone. The cross-over interaction between tone frequency and ISI showed that tone frequency significantly affected ITI, which indicated motor timing under the short ISI, falling under the peri-second range. However, no such effect was observed in the supra-second range of the long ISI.

ISI had no significant impact on motor timing (i.e., ITI), implying that participants showed similar motor timing in both short and long ISI. Furthermore, this indicated that the peri-second range was also involved in cognitive control. As frequency affected the motor timing in the peri-second range (unlike the supra-second range), it is clear that the range is sensory-automatic (Rammsayer & Troche, 2014). In the peri-second range, motor production was perceptually sensitive. This range was automatic and frequency-sensitive as it did not involve much top-down attention, unlike the supra-second range, which was slow (because of higher ISI) and driven by top-down attention. There are two potential causes: subjective rhythmization (SR), a phenomenon in which a participant would have perceived the sounds in an isochronous sound sequence as subjectively different despite their objectively equal amplitude. It was also discovered that the SR experience ceases when the ISI between consecutive sound onsets increases beyond 1600 ms(Fraisse, 1982), which reflected the results of this study as there is no significant effect of tone frequency on motor timing (2000 ms).

The peri-second range is a transit zone between the sub-second and supra-second ranges. The results support the research conducted by Michon (1985), which explains that, in short intervals, the processing of temporal information is sensory-automatic, independent of top-down control, and regulated by the reticular activating system (RAS), whereas, in longer intervals, the processing of temporal information necessitates cognitive control. The research showed that tone frequency was the primary factor affecting motor timing in the peri-second region, and this range that motor timing was impacted by perceptual timing as the frequency of the tone is a perceptual attribute of sound and is effective in affecting motor timing, as opposed to the supra-second range.

The central nervous system is less active in the peri-second range because, unlike feedforward motor control, the signal does not directly travel from the centralized brain to the finger muscles (Seidler et al., 2004). Participants overestimated the goal time of 1000 milliseconds in this range to 1001.042 milliseconds with a low-frequency tone and underestimated it to 998.28 milliseconds with a high-frequency tone. This finding supports the ideomotor principle, which states that perception will modulate action where overlap exists between perceptual and motor representations of action (Ammirante et al., 2011). The reason behind the occurrence of this phenomenon might lie in the movement execution process in a synchronized tapping task. In a synchronized tapping task, movement execution completes in three steps of a single cycle of finger tapping. In the first step, the participants will move down their finger, tap on the surface, and then hold their finger for some time, called *dwell time 1*. Later, the participants will release and extend the finger by recruiting extensor muscle, then again down their finger and tap on the surface. This includes the initial cycle of continuous tapping, and the entire period will be counted as ITI, as well as the duration during which the finger remains in contact with the surface following the initial touch before being released and moving vertically up before moving down for the subsequent tap (excluding the time of the next successive tap). It suggests that participants took their time moving their fingers vertically up and down after estimating the tone length by keeping them on the surface for a while.

Accuracy in motor timing is necessary for better SMS performance. The shift in arousal prevents it from happening because it can only be completed at the optimal degree of arousal. This is because performance only temporarily rises with physiological or mental stimulation. An excessive amount of arousal lowers performance. Since time perception and motor timing are interwoven, physiological arousal changes time perception, which changes motor timing. Previous research revealed that auditory arousal is influenced by the pitch of the signal, which is a psychological correlate of tone frequency. It was discovered that low-frequency square waves elicited considerably lower auditory arousal thresholds (AATs) than other tonal frequencies. It is seen that tension arousal increases by high-pitch level tones (Ilie & Thompson, 2006). High-pitched music has been correlated with high-arousal words such as excitation, exhilaration, animation, elation, movement, strength, alertness, and higher excitation emotions such as high spirits, rage, anger, and fear (Coutinho & Cangelosi, 2009; Eitan & Timmers, 2010). Perceived duration lengthens with an increase in physiological arousal and shortens with a decrease in physiological arousal (Droit-Volet & Gil, 2009; Lake et al., 2016), and this physiological arousal alters the time perception and leads to changes in motor timing as both are intertwined with each other (O'Regan et al., 2017). According to Yerkes-Dodson law, peak performance occurs at optimal arousal level (Cohen, 2018), so it is assumed that at optimal arousal level, the mean inter-tap interval would be closer to ISI. Nonetheless, when arousal level shifts towards either direction from optimum arousal, then and are likely to increase or decrease. Here t_1 refers to dwell time that is sensitive to the attention, which impacts the judgment of frequency and duration

(Winkler et al., 2015), and t_2 is the movement time taken for going up and down as the tap velocity is constant because of the fixed period.

$t_1 + t_2$ = ISI=ITI (at optimal arousal)

t_1'+ t_2'=ITI (when arousal shifts from optimal arousal)

(t_1'-t_1) + ($t_2' - t_2$)= ITI-ISI=Δ (difference between ITI and ISI)

If t_1'>t_1 and t_2'>t_2 then Δ>0

ITI-ISI=Δ

ITI=ISI+Δ

If $t_1' < t_1$ and t_2'<t_2 then Δ<0

ITI=ISI-|Δ|

Participants demonstrated greater ITI than ISI with LF tone and lesser than ISI with HF tone. Therefore, with LF tone, Δ was positive; it happened only when the dwell time for LF tone was higher than HF tone. The dwell time was influenced by sensory and cognitive processing. It was also related to the level of motivation, engagement, and movement time that had come under the transition phase unaffected by attention, so it would be the same regardless of the frequency of auditory stimuli. Therefore, only dwelling time, not moving time, would cause a change in the pace of the taps. Participants felt more engaged in the tapping activity when using LF tone with short ISI than HF tone because dwell duration was relatively longer with LF tone than with HF tone. This indicates that the individual liked the LF tone more. The only visible representation of the listener's subjective experience of tone frequency was this perception and action link, which caused this motion and minor tensions in the listener's finger muscles (Repp, 1993).

Future studies on the tonal frequency and dwell time in the sub-second and peri-second regions should give the researchers a better understanding of different brain parts. It's vital to remember that only one pitch of the sound was used in the current investigation. Additional insights can be gained by including an isochronous sequence of tones with varied pitches. Furthermore, because just one frequency level in the LF and HF ranges was used in this study, it will be intriguing to investigate this aspect over a wider frequency range. Another crucial and pertinent subject in this context is what would happen to motor timing when two subsequent tones of different pitches are provided in the same metronome differently.

Conclusion

The current study shows that perceptual-motor interaction is more common in the peri-second range, where auditory-motor synchronization will function better than in the supra-second region. In summary, the investigators conclude that, on average, participants over-translated the ISI with a low-frequency tone and under-translated with a high-frequency tone in the peri-second range. At the same time, this trend is absent in the supra-second range. The findings imply that tone frequency substantially affects motor timing in the peri-second region more than in the supra-second range. The results of this study suggest that ISI should, therefore, be a crucial consideration when creating sound sequences for auditory-motor synchronization.

Chapter 7

Motor Variability in Auditory-Motor Synchronization

7.1. Introduction

Sensorymotor synchronization (SMS) is the rhythmic coordination between perception and action in which individuals try to align their action such as motor response with a tone. In attempt to do so, they adjust their internal timekeeper with the rate at which stimuli occur in a sound sequence. This involves trial and error process manifested in the form of motor variability (MV). It is because, initially individuals go through learning process. In acquiring learning, MV is prominent and once they acquired learning, variability declines. For learning, MV is needed in the same way, in which genetic variation is necessary for evolution. MV is an essential component and plays an important role in adapting behaviors (Skinner, 1981).

Variability in the motor system normally occurs at the motor periphery (Van beers et al., 2004; Clamann, 1969; Hamilton et al., 2004; Jones et al., 2002). During force production, signal dependent noise occurs in the form of trial to trial variability fluctuation that scale in linear manner with mean force reflect a basic property of the muscle function (Clamann, 1969; Hamilton et al., 2004). In the context of auditory-motor synchronization (AMS), MV occurs in the form of tap to tap variability in response to auditory metronome. MV is depends on the stimulus features like inter-stimulus interval (ISI) and tone frequency because these parameters likely to modulate the temporal information processing. In a synchronization paradigm study performed by Peters (1989), subjects were asked to track the intervals ranges from 180 ms to 1000 ms. The results suggested that there was a sudden increase of variability of inter-tap interval (ITI) near 300 ms because of the transition from automatic to controlled movement. Miyake et al. (2002) found that temporal information processing in the ISI range between 1800 to 3600 ms is influenced by top-down attention but it is unaffected by top-down attention in the 450 to 1500 ms range. This implies that supra-second range involves central planning circuits, and it is said that variability is said to be better suited for learning-related motor exploration in central planning circuits. These circuits are more accessible to reinforcement signals (Björklund & Dunnett, 2007; Skinner, 1981) and show more experience-dependent plasticity (Doyon & Benali, 2005; Nudo et al., 1996; Schultz, 1998). Reinforcement learning states that when the stakes are high, exploitation will happen otherwise exploration will occur (Van beers et al., 2004). It is found that the central nervous system (CNS) regulates MV in context dependent manner by responding to rewarding situations. The pleasant and motivational stimuli act as reward for an individual, which cause adaptive responses and affect different cognitive functions. Growing evidence suggests that stimuli with unique motivational values can strongly modulate perception and attention.

Sound can be reinforcing stimulus that is used to direct flexible behavior—as such, its meaning is often dependent on contextual cues. Basically, sound has 3 important features: one is frequency, amplitude and other is duration. Sound pleasantness depends on the frequency (which is a perceptual feature of sound) because high frequencies are associated with unpleasantness by the individuals as compared to low-frequency spectral components (Halpern et al., 1986). In this way, motor variability seems to be connected with the frequency of auditory stimuli and ISI. In a synchronization paradigm study performed by Peters (1989) subjects were required to track the intervals ranging from 180 ms to 1000 ms. The results suggested that there was a sudden increase of variability of ITI near 300 ms because of the transition from automatic to controlled movement. Since in short ISI, less amount of top-down attention involved than long ISI so it is possible that there will be lesser MV in short ISI as compared to long ISI.

If an individual shows lesser variability in ITI then it means that they perform action in a consistent manner and they does not change learning strategy and manipulates and selectively change their action in reinforcement manner, instead they uses the past information or memory and translates into the action in a predictive manner or automated manner where no error feedback mechanism is involved. Lesser variability in ITI in learning phase has a disadvantage in such a manner that it leave less scope for an individual to adjust his/her internal timekeeper in order to achieve synchronization. In this way, in case of lesser variability of ITI, individual will respond and tap to the tone (present in isochronous sound sequence) in predictive mode under short ISI irrespective of frequency of tone under peri-second and supra-second range (which involves different level of cognitive resources). Furthermore, under long ISI, which comes under the supra-second range where more central planning is involved, MV is better suited for learning-related motor exploration and helps in copying the metric pattern of metronome into the muscle memory during sensory-motor learning. A central planning circuit (which is part of central nervous system) has better accessibility to reinforcement signals and demonstrates experience-dependent plasticity (Björklund & Dunnett, 2007; Nudo et al., 1996; Sanes & Donoghue, 2000). Under long ISI, MV likely to be modulated by the tone frequency unlike short ISI which is very fast and automatic in nature. Under short ISI, there is lesser degree of motor controllability, because here variability is derived from periphery and often referred to as 'execution noise' (Van beers et al., 2004) which is likely to be unmodulated by tone frequency.

Since low-frequency spectral components that lead to the 'chilling' effect of fingernails scratching on a blackboard, not the high frequencies that most people associate with the unpleasant nature of that sound (Halpern et al., 1986) so low-frequency tones are more rewarding than high-frequency tones. This implies that subjects will orient their top-down attention more to low-frequency tone than high-frequency tone. Therefore, it is expected that there will high MV with low-frequency tones as compared to high-frequency tones.

Investigators of the present study have tried to see the effect of tone frequency on predictive tapping and MV under two different time scale by creating an experimental protocol in which the participant has to listen to a series of tones separated by a certain ISI in the isochronous sound sequence and tap with the tone accurately. In this current study, there are two isochronous sequences of different ISI, one is a short ISI of 1000 ms in peri-second range, which does involve less subject's top-down attention and another one is a long ISI of 2000 ms in supra-second range that included the subject's top-down attention. Each sequence is divided into two blocks and each block has 35 identical tones. Tone frequency is different in both blocks. Based on previous explanations, the investigator selected HF tone of 1000 Hz (>800 Hz) in one block and the other has LF tone of 100 Hz (<300 Hz).

On the whole, this study emphasizes the role of frequency of tone on predictive tapping and MV under short and long ISI, respectively. The following hypotheses are tested in this study:

- There is higher MV in long ISI as compared to short ISI.
- There is predictive tapping under short ISI irrespective of the frequency of tone.
- Under long ISI, MV is higher with LF tone than the HF tone.

7.2. Results

An alpha value of 0.05 was used to test the statistical significance in all the analyses throughout this paper.

7.2.1. Predictive Tapping

An ANOVA was carried out for 2 (ISI: short ISI vs. long ISI) $\times$2 (Frequency: LF vs. HF) mixed factorial design with the first factor as a between-groups variable. The analysis revealed that there is a main effect of ISI on predictive tapping (short ISI, M = 35.7, and long

ISI, M = 20.4, see Figure. 14), F (1, 28) = 29.517, p = 0.0001, η_p^2= .513. There was no main effect of frequency (LF, M = 28.06, HF, M =27.03), F (1, 28) = 0.054, p = 0.818, η_p^2 = 0.002). Individuals made significant number of predictive tapping responses under short ISI as compared to long ISI irrespective of tone frequency which has no significant effect on predictive tapping

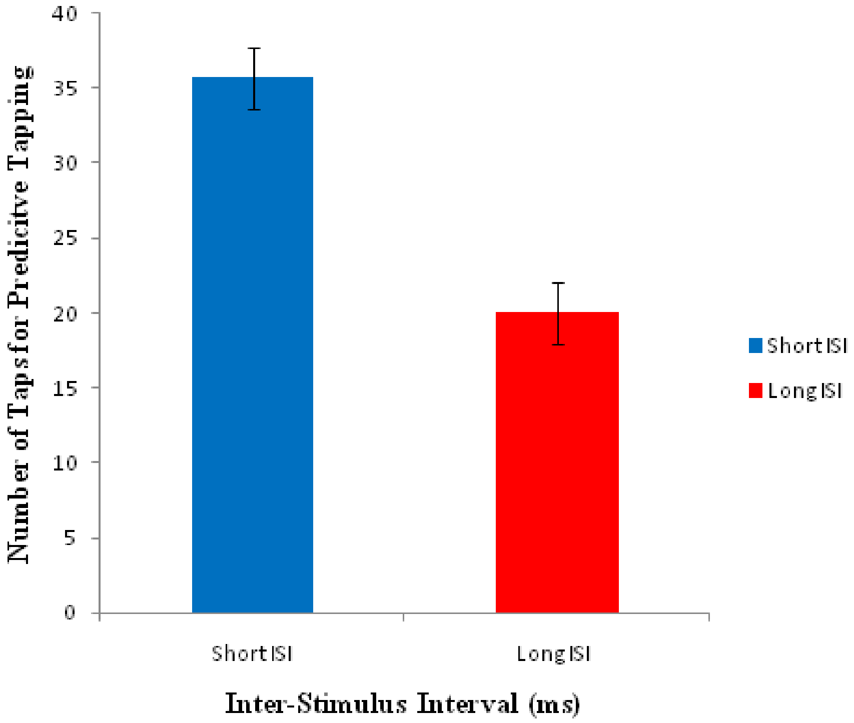


Figure.14 Predictive Tapping under short and long ISI. Error bars denote Standard Error.

7.2.2. Motor Variability (MV)

In order to examine MV, an ANOVA was carried out for 2 (ISI: Short ISI vs. Long ISI) ×2 (Frequency: LF vs. HF) mixed factorial design with the first factor as a between-groups variable. With regard to MV, significant frequency by ISI cross -over interaction was found in ANOVA for the LF and HF based conditions [F (1, 28) = 4.689, p = 0.039, η_p^2= 0.143) (see Figure.15). It appears that participants showed higher MV with LF tone as compared to HF tone under long ISI unlike short ISI where reverse effect was observed.

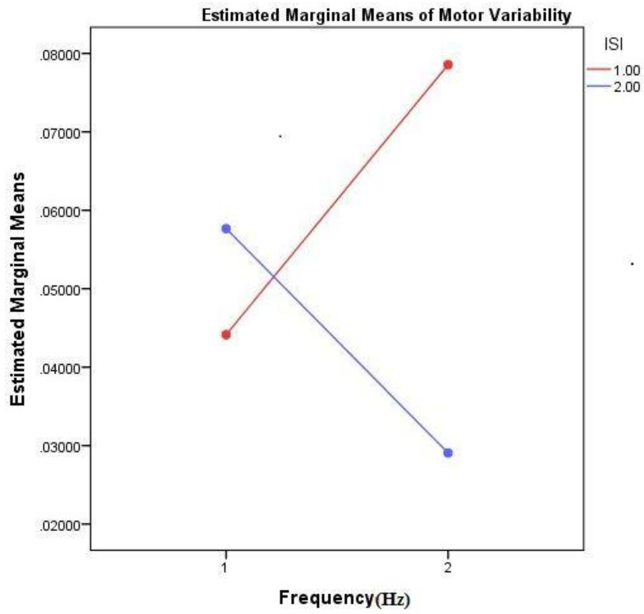


Figure.15 Significant Interaction between frequency and ISI.

Furthermore, a paired sample *t*-test was carried out to check the significance level of difference between the means of MV for LF and HF conditions under both ISI. Investigators found that there was a significant difference in the mean MV for LF (*M* = 0.069, *SD* = .056) and HF (*M* = 0.029, *SD* = 0.025) conditions; *t*(14) = 3.427, *p* = 0.004 (Figure.16) under long ISI (2000 ms) but this trend absent under short ISI (1000 ms) since no significant difference was found in the mean MV for LF (*M* = 0.044, *SD* = .0144) and HF (*M* = 0.078, *SD* = .126) conditions; *t*(14) = -1.061, *p* = 0.307. These results suggest that the mean MV for LF significantly more than HF under the long ISI (Figure.16). This indicates that under long ISI—with LF tone, participant's motor learning is higher as compared to HF tone. This result strengthens our argument that MV does not get influence by the frequency of auditory stimuli under short ISI as compared to long ISI.

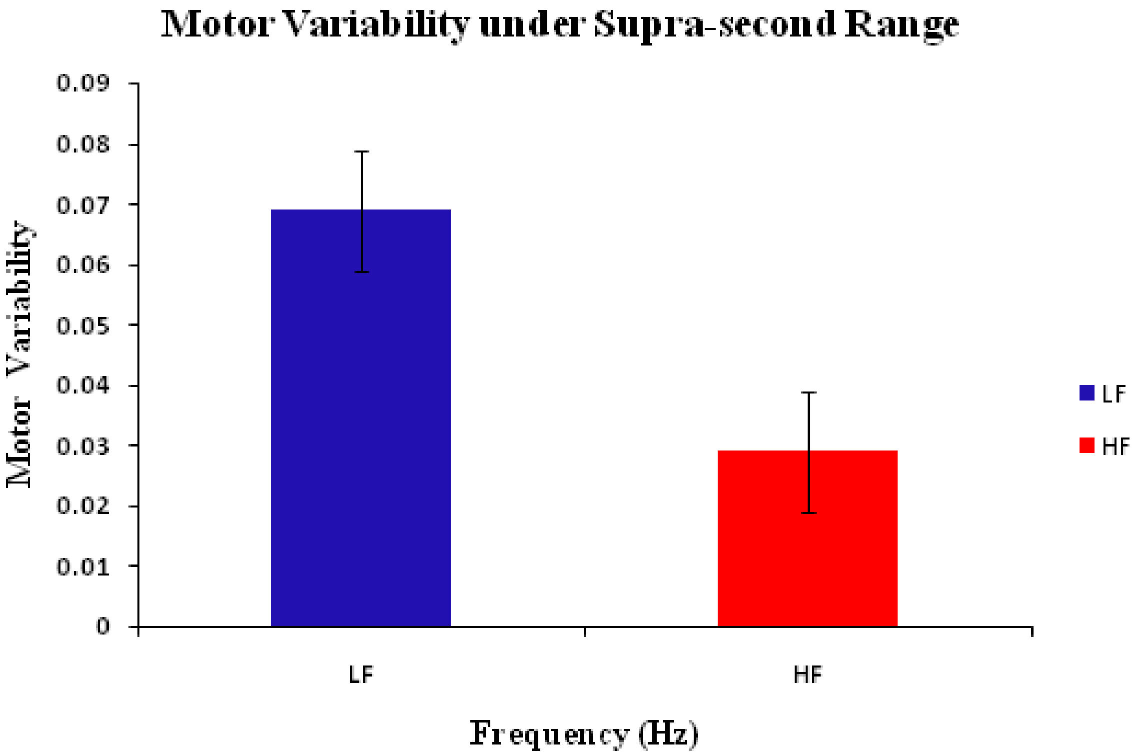


Figure.16 Motor Variability with LF and HF tone under short ISI (which comes under supra-second range). Error bars denote Standard Error.

7.3. Discussion

In this study, investigators examined the role of tone frequency and ISI in predictive tapping and MV under two different time scale (peri-second vs. supra-second range). Predictive tapping was analyzed in terms of negative asynchrony where tap (response) occurred before the tone under which participants tap in a predictive-based manner without any feedback. MV was analyzed in terms of coefficient of variation of the ITI. The results showed that there were a significant number of cases for predictive tapping under short ISI as compared to long ISI irrespective of the tone frequency which has no significant effect on predictive tapping. Also, tone frequency showed a significant effect on MV under supra-second range unlike peri-second range. This implies that the tone frequency affects an individuals' external focus of attention, and regulates motor control (Lohse et al., 2014). Therefore, individuals' motor responses were reactive to the tone rather than predictive under long ISI. Under short ISI, where there were higher number of cases for predictive tapping, MV was unaffected by tone frequency. Motor control is not regulated by tone frequency but rather by a prediction-based feed-forward mechanism. As a result, individuals made a predictive tapping response before the tone. Higher MV under low-frequency tone as compared to high-frequency tone indicates that — (a) low-frequency tone captures more external focus of attention than high-frequency tone and (b) low frequency tones are better for auditory-motor learning under supra-second range. There is prediction based learning instead of stimulus-driven learning under short ISI because lateral side of cerebellum plays an significant role in initiating and coordinating predictive movements (Bares et al.,2007, Xu, 2006) and is well connected with motor cortices via thalamus which involved in predicting movements (Dum & Strick, 2005; Strick et al., 2009).

In terms of temporal processing of events, involvement of cerebellum is mostly limited to sub and peri-second range (Ivry & Spencer, 2004; Lewis & Miall, 2003) but not in supra-second range. Therefore, predictive tapping was dominant in contrast to reactive tapping under short ISI irrespective of tone frequency. The participants made tapping decision in a predictive-based manner not reinforced by the perceptual property of the auditory stimuli. The results revealed that tone frequency showed a significant effect on MV under supra-second range. However, no such effect was found in peri-second range. Supra-second range involves more top-down attention and central planning circuits and variability in these circuits is better suited for learning-related motor exploration. It is worth noting that these circuits are aided by easier access to reward signals (Björklund & Dunnett, 2007; Schultz, 1998), and exhibit more experience-dependent plasticity (Doyon & Benali, 2005; Nudo et al., 1996; Sanes & Donoghue, 2000). The motivation characteristics of stimuli impact stimulus-driven reinforcement learning. Adaptive reactions and cognitive functioning are influenced by motivational stimuli such as rewards. In example, there is mounting evidence that stimuli with specific motivational values can significantly influence perception and attention (Bourgeois et al., 2016). Both selective top-down and stimulus-driven attentional orienting are similar to these effects.

Under long ISI (within supra-second range), participants showed more MV with low-frequency tones as compared to high-frequency tones. Since high-frequency tones increase stress arousal (Ilie & Thompson, 2006), participants avoided them, while they responded and reacted to low-frequency tones. Hence, low-frequency tones are more rewarding than high-frequency tones. Further, LF tone has a pleasing effect, so an individual's attention will orient to it selectively, so variability with LF will be comparatively more than the HF tone. This implies that LF tones are best suited for motor learning under the supra-second range.

In the supra-second range, participants showed higher MV with low-frequency tones as compared to high-frequency tones because MV is context-sensitive across basal ganglia network. Basal ganglia play an important role in supra-second timing, unlike sub-second and peri-second range. As a result, tone frequency (which is perceptual property of the sound) influenced the MV in supra-second range unlike sub-second and peri-second range. In supra-second range, variability facilitates learning because the subject's attention is more involved. As a result of this, motor control is regulated by a feedback mechanism, which is a closed-loop system. In the feedback system, error fed back to the input to reduce the error between target and input and causes the MV. MV is required during the transient phase to aid learning and motor exploration, which is a trial-and-error process aimed at improving motor outcomes. It is well recognized that it plays an important role in motor learning. This is especially true during reinforcement learning, when actions that result in a successful outcome are encouraged while acts that result in a unsuccessful actions are avoided.

Since learning begins with the conscious mind and requires top-down attention, and then transfers into the unconscious mind after a period of time (McLaughlin, 1990). As a result, it is recommended that the subject be asked to synchronize with slow tempo (long ISI) first with LF tone, and then gradually progress to high tempo (short ISI).

7.4. Conclusion

This paper demonstrates that participants showed predictive tapping under short ISI (peri-second range) irrespective of tone frequency. Under short ISI (within supra-second range), tone frequency significantly influenced the MV but this trend was absent under long ISI. Therefore, it is evident that motor control is regulated by prediction-based feed-forward mechanism under short ISI. As a result, individuals made a predictive tapping response before the tone. Under long ISI, participants made a reactive response to the tone in a reinforced manner unlike short ISI where they made a predictive tapping response to the tone without being affected by tone frequency. Participants showed higher MV with low-frequency tone than high-frequency tone in the long ISI (in supra-second range) while this trend is absent in short ISI (in peri-second range). Hence, low-frequency tones are more rewarding than high-frequency tones. In terms of implications, results suggest that supra second range time window is best suited for motor learning, which is reward-based and low-frequency tones are best suited for motor learning as compared to high-frequency tones. For better motor learning, supra-second time range will be beneficial in rhythmic auditory stimulation and stimulated by low-frequency tones For future directions, investigation is needed to examine how the low-frequency tone will affect AMS if participants switch over from learning phase (in supra-second range) to synchronization phase (in sub and peri-second range) or it will be worth knowing that whether this training pattern will improve AMS or not.

Chapter 8

General Discussion

The present dissertation addresses the role of tone frequency (Low vs. High) in motor timing, tapping behavior, synchronization accuracy and listening effort under peri-second and supra-second interval timing. The cognitive processes underlying auditory motor synchronization (AMS) are related to temporal information processing. As evident from previous chapters, increasing line of research has contributed to our knowledge of various aspects of AMS and the current research is an effort to pursue this avenue for further scientific progress by taking into consideration different processes of tapping behavior to isochronous sound sequence.

In particular, this research emphasizes on how the tone frequency under peri-second and supra-second interval timing affect individuals' tapping behavior, tapping performance (synchronization error) and listening effort in a controlled listening environment. The current chapter unifies and reflects on the findings of the present research in the light of the available literature and relevant insights.

This study focuses on the issues pertaining to SMS in isochronous sound sequence. Tone frequency and ISI are an important aspects concerning perception-action coupling, tapping behavior (comprise of timing and variability), synchronization error and listening effort. In this study, investigator examines the role of tone frequency affect these aspects under two different time scale (peri-second vs supra-second range)

As per the results, participants showed more temporal alignment with low-frequency tones as compared to high-frequency tones under short ISI and long ISI. It implies that low-frequency tones demand less attentional resources for efficient temporal information processing as compared to high-frequency tones. On the other hand, lower listening effort was observed with low-frequency tones as compared to high-frequency tones under both ISIs. the dominance of anticipatory tapping over reactive tapping under short ISI implies the predictive quality of a synchronized motor action with an anticipated stimulus]unlike long ISI. On the other hand, higher occurrence of reactive tapping under long ISI as compared to short ISI indicates a reflex activity. This occurs in reaction to a stimulus associated with feedback motor control relying on sensory feedback for coordinating movements.

This study emphasizes the issues pertaining to SMS in isochronous sound sequence. ISI is an important aspect concerning response asynchrony.

As per the results, the dominance of anticipatory tapping over reactive tapping under short ISI implies the predictive quality of a synchronized motor action with an anticipated stimulus (Repp, 2005; Ascherzleben & Prinz, 1995; Ascherzleben, 2002) unlike long ISI. On the other hand, higher occurrence of reactive tapping under long ISI as compared to short ISI indicates a reflex activity. This occurs in reaction to a stimulus associated with feedback motor control relying on sensory feedback for coordinating movements (Iosa et al., 2015). The investigator found an interesting finding concerning motor timing is that tone frequency has shown significant effect on motor timing under 1-second range. But no such effect was found in Supra-second range. While ISI has no significant impact on motor timing which implies that participants showed similar motor timing in both short ISI and long ISI. It implies that peri-second range also involves some top-down attention. As frequency affected the motor timing in this range unlike supra second range so it is clear that this range involve perceptual process. On the whole, peri-second range (1-second) involves both cognitive and perceptual process whereas only cognitive process are involved in supra-second range.

8.1. Tapping behavior

As per the results, the dominance of anticipatory tapping over reactive tapping under short ISI implies the predictive quality of a synchronized motor action with an anticipated stimulus unlike long

ISI. On the other hand, higher occurrence of reactive tapping under long ISI as compared to short ISI indicates a reflex activity. This occurs in reaction to a stimulus associated with feedback motor control relying on sensory feedback for coordinating movements.

In case of short ISI, individuals tend to tap before the onset of the tone by anticipating it, instead of tapping in a symmetric uniform distribution around original (irregular) stimuli. Therefore, investigators conclude that anticipation plays an important role in the short ISI to predict the next event while synchronizing with the external event. Anticipation happens in the absence of required attentional resources. It conforms to the understanding that lesser attentional demand is required in the short ISI. In this context, a relevant issue is in terms of how anticipation takes place when there is a lesser involvement of top-down attention. The point is anticipation takes place when activities rely not only on previous and present stimuli but also on future projections, expectations, or beliefs. Therefore, the current study indicates that the role of sequential decision making becomes very important in synchronizing with the tones present in isochronous sound sequence and anticipation is a critical part of this process.

In sequential decision making, individuals make decisions based on anticipation involving memory (i.e., retrieval of information by attractor dynamics if a given input is sufficiently similar to a stored attractor state). During decision making between sequential auditory stimuli, the first stimulus is held in memory and then compared with the subsequent similar stimuli. Attractor neural activity of second stimuli depends not only on the present stimuli but also on the sensory memory of the first sound event. Anticipation cannot happen without memory because it plays an important role in prediction and it gives the previous experience to the individual for anticipation of the next stimulus from its current state. Auditory sensory memory is the shortest element of memory and the critical first stage in auditory perception . It gives the cognitive system- to retain representations of sensory information after the previous auditory stimuli have ended. It acts as a buffer for the stimuli received through auditory senses, which is very briefly (but accurately) retained.

In case of short ISI, next stimulus occurred before the decaying of sensory memory, so this past sensory memory helped to predict the next event via anticipation mechanism and the ‘prediction’ is a preplanned behavior that activates the feedforward motor control mechanism and it is unaffected by peripheral feedback. Feed-Forward motor control is an open-loop system for motor control that controls fast, ballistic movements before any sensory information is processed. It requires prior knowledge of the upcoming stimulus. In this automatic mechanism, some reflex loops require lower brain areas and are independent of perceptual processing active control but can be affected by prior commands; while the other loops are directed by the spinal cord with no active control of the brain.

In case of long ISI, participants showed reactive tapping behavior. Reactive response is a reflexive response which is not based on anticipation. It is because of memory traces decline (i.e., decaying) over a short span of time until a certain limit is reached, and that memory becomes unreliable. Under such circumstances, it becomes difficult for an individual to retain representation of auditory information for a longer duration because memory decays very quickly, in the region between 200-500 milliseconds (less than a second). Moreover, in the absence of anticipation mechanism, top-down attention plays a critical role which switches an individual’s motor action to feedback mode (in contrast feed-forward mode) for error compensation. In other words, feedback motor control relies on sensory feedback for coordinating movements. Reactive tapping is a reflex activity which occurs in reaction to a stimulus due to feedback motor control. As reactive tapping in long ISI (2000 ms) occurs under supra-second range and anticipatory tapping in short ISI (1000 ms) occurs in 1-second range, it indicates that as the ISI increases from 1-second to supra-second range, there is a transition in tapping from anticipatory to reactive.

In future studies, it will be important to examine issues concerning sensory-automatic processing in 1-second range which is a transit zone between sub-second range and the supra-second range.

8.2. Tapping Performance

During the experiment, it was observed that participants were unable to catch the first beat but they were able to anticipate the next beat under both conditions of ISI. While debriefing, no variation in tapping was observed because of the fixed ISI in the isochronous sound sequence. Investigators did not observe a significant difference between the synchronization error means under two different ISI. This implies that 1-second range is also influenced by attention which corroborates with Rammsayer & Troche (2014).

The investigators observed two different types of tapping under two different ISIs. In short ISI=1000 ms, anticipatory tapping is more prevalent. Under long ISI, there were more number of cases for reactive tapping in long ISI as compared to short ISI. The investigators observed a significant effect of frequency on synchronization error and found that it was significantly higher in the sound frequency of 1000 Hz than the low sound frequency of 100 Hz which comes under bass range. It implies that subjects' tapping were comparably more aligned with the low-frequency sound than with the high-frequency tone. This finding corroborates with the previous research which asserts that when people are presented with music, they have a strong tendency to align their body movement with the low-frequency sounds (Burger et al., 2013) and increased body movements are associated with an increase in loudness of low-frequency sounds (Dyck et al., 2013), and there is one transcranial magnetic stimulation (TMS) study also substantiates the findings which show that songs which have a more spectral flux in the low-frequency region activate the motor system in higher magnitude (Stupacher et al., 2013).

Following frequency rules, frequency of the auditory nerve's impulses corresponds to the tone frequency, which allows us to detect its pitch. The point is that these low-frequency sounds in the range of 20 to

1500 to cycles for each sound can cause volleys (group of auditory neurons) of impulses resonating at same frequencies, and the cochlear nerve transmits these volleys into the brain's cochlear nuclei. It is assumed that the brain's cochlear nuclei can recognize the different frequencies of the volleys (Bamne et al., 2015).

Such findings expand previous ones in relation to the synchronization activity relationship between low-frequency spectral flux and general music-induced movement (Burger et al., 2013). In the debriefing session, the investigators observed that majority of the participants showed significant discomfort in listening to the high-frequency tones than to the low-frequency tones, This implies that the listening task was more demanding with the high-frequency tones than the low-frequency tones.

The finding of the listening effort concluded that there was a significant effect of frequency on the participant's listening effort. Listening effort was significantly higher in the case of HF as compared to LF which indicates that participants felt less listening discomfort with the low-frequency tones than with high-frequency tones which showed an agreement with participant's subjective ratings of listening effort. This is because low-frequency tones are perceived to be less loud than high-frequency tones at the same intensity and it requires less attentional demand for efficient temporal information processing leaving more attentional resources for tapping task as compared to high-frequency tones.

With low-frequency tones, higher performance can be achieved with the minimum mental effort required for listening. In other words, higher movement efficiency can be achieved with low-frequency tones. Movement efficiency has been defined as the ratio of mechanical work to the metabolic energy expended (Cavanagh & Kram, 1985).

8.3. Motor Timing

This study unravels the perception-cognition-action link in context of AMS via synchronized tapping paradigm. It examines how the subject subjectively perceive and translate the ISI into action reflected by inter-tap interval (ITI) and what is the role of auditory information such tone frequency in affecting this perception-action process under peri-second and supra-second interval timing which recruits different number of cognitive resources. This study reports the mean inter-tap interval (representing motor timing) and subjective translation index (STI) indicating whether subject has over and under translated the ISI.

The results revealed that tone frequency had shown a significant effect on STI and cross-over interaction was found between tone frequency and interval timing. Participants showed higher STI (>1) with LF tone than HF tone (<1) indicating that participants over-translated ISI with LF tone and under-translated with HF tone. The cross over interaction between tone frequency and ISI signifies that tone frequency had shown a significant effect on ITI (represents motor timing) under the short ISI which comes under peri-second range. But no such effect was found in the long ISI that comes within supra-second range. ISI has no significant impact on motor timing (i.e., ITI). This implies that participants showed similar motor timing in both short ISI and long ISI. Furthermore, this indicates that the peri-second range also involves cognitive control. As frequency affects the motor timing in peri-second range (unlike supra-second range), so it is clear that this range is sensory-automatic (Rammasyer, 2014).

This implies that motor production is perceptually sensitive in the peri-second range. The peri-second range is susceptible to frequency influences and automatic as it does not involve much top-down attention unlike supra second range which is slow (because of higher ISI) and driven by top-down attention (which involves central planning).There are two possible reasons: one is subjective rhythmization (SR), it is the phenomenon in which subject perceive the sounds in an isochronous sound sequence subjectively different despite their objectively equal amplitude and it is found that SR experience ceases when the inter-stimulus interval (ISI) between consecutive sound onsets increases beyond 1600 ms (Bolton,1894; Fraisse, 1982), which is reflected in our findings as there is no significant effect shown by frequency on motor timing in long ISI (2000 ms).

The findings support the theoretical framework laid by Daniel Kahneman (Arvai, 2013) which states that there are two different systems for information processing: One is System 1 which is automatic, and influenced by external factors; another one is System 2 which is slow in processing, reflexive, and driven by top-down attention and intentions. It means that system 1 is largely associated with sub-second and peri-second ranges, and system-2 is associated with supra second-time system.

If we talk about neural correlates about sub-second and supra second-time system, there are two different neural networks for these time system (Lewis & Miall, 2003b; Pouthas et al., 2005; Jahanshahi et al., 2006; Wiener et al., 2010). One of the regions triggered in supra-second timing is frontal and posterior parietal areas of the cortex, which are connected to working memory or attention. The other region activated in the sub-second and peri-second timing range is the motor and somatosensory systems, including the pre-central and post-central regions, and supplementary motor area (SMA). The peri-second range is a transit zone that comes in between the sub-second and supra-second range. These findings are in corroboration with Michon (1985) that processing of temporal information in short intervals is sensory-automatic and beyond the top-down control and regulated by reticular activation system whereas, in longer intervals, temporal information processing requires cognitive control. The analysis revealed that tone frequency played the main role in influencing the motor timing in the peri-second range. As the tone frequency is a perceptual property of sound, and it is effective in influencing motor timing in peri-second range unlike supra-second range, so it implies that in peri-second range, motor timing is influenced by perceptual timing.

In the peri-second range, there is less involvement of the central nervous system so signal does not reach to the finger muscles from a centralized brain directly like a feedforward motor control (Seidler, 2004). In this range, participants over-translated the objective time of 1000 ms to 1000.642 ms with low-frequency tone and under-translated to 998.83 ms with a high-frequency tone. This finding supports the ideomotor principle, which states that perception will modulate action where overlap exists between perceptual and motor representations of action (Ammirante et al., 2011). The reason behind the occurring of this phenomenon might be lies with the movement execution process in synchronized tapping task. In a synchronized tapping task, movement execution completes in three steps of a single cycle of finger tapping: in the first step, the subject will move down his/her finger and tap on the surface and then hold his/her finger for some time called as dwell timet_1 and then he/she will release and extend the finger by recruiting of extensor muscle then again will down his finger and tap on surface. This includes the first cycle of continuous tapping, and the whole interval will be count as the inter-tap interval. This includes the time when the finger stays in contact with the surface after the first tap and then is released, and it moves vertically up, and again moves down for the next tap (excluding the time of next successive tap). It implies that subject has kept his finger on the surface for some time to estimate the duration of tone denoted as t_1, and then he/she has taken time t_2 in moving finger vertically up and down.

Accuracy in motor timing is necessary for a better SMS performance and can only be accomplished at an optimum level of arousal, but it does not happen because of the shift in arousal. It is because performance increases with physiological or mental arousal, but only up to a point. When levels of arousal become too high, performance decreases (Cohen, 2018). Physiological arousal alters the time perception, and this leads to change in motor timing as both are intertwined with each other (O’Regan et al., 2017).It is seen that tension arousal increases by high pitch level tones (Ilie & Thompson, 2006). High-pitched music has been correlated with high-arousal words such as excitation, exhilaration, animation, elation, movement, strength, alertness, and higher excitation emotions such as high spirits, or rage, anger and fear (Coutinho & Cangelosi, 2009; Eitan &

Timmers, 2010). Perceived duration lengthens with an increase in physiological arousal and shortens with a decrease in physiological arousal (Droit, 2007; lake, 2016). There is hypothetical pacemaker which has a natural rate which increases or decreases by the change in arousal level which leads to a longer and shorter perception of the duration (Schwarz, et al.,2013) and this physiological arousal which alter the time perception and this leads to change in motor timing as both are intertwined to each other (O'Regan et al.,2017). It is assumed that at optimal arousal level, individual mean inter tap interval will be close to ISI, but when arousal level shifts towards either direction from optimum arousal, then t_1 and t_2 are likely to increase or decrease. Here t_1 refers to dwell time is the perceptual time judgment of the tone, and t_2 is the movement time taken for going up and down. As the tap velocity is constant because of the fixed period.

$t_1 + t_2$= ISI=ITI (at optimal arousal)

$t_1' + t_2'$=ITI (when arousal shift from optimal arousal)

$(t_1' - t_1) + (t_2' - t_2)$= ITI-ISI=C

If $t_1' > t_1$ and $t_2' > t_2$ then C>0

ITI-ISI=C

ITI=ISI+C

If $t_1' < t_1$ and $t_2' < t_2$ then C<0

ITI=ISI-C

As participants showed ITI greater than ISI with LF tone and less than ISI with HF tone. So with LF tone, C will be positive; it will happen only in the case when dwell time for LF tone is higher than HF tone.

Here, dwelling time is influenced by sensory and cognitive processing and also relates to the level of motivation and engagement and movement time that comes under the transition phase (which is purely motor function) unaffected by attention (Austin, 2013) so it would be the same regardless of the level of frequency of auditory stimuli. Thus adjustment in the pacing of the taps would only be due to dwelling time, not due to movement time. Due to the fact that dwell time is comparatively more with LF tone than HF tone, participants felt more engagement in tapping activity with LF tone compared to HF tone with short ISI. This means that the participant enjoyed more with the LF tone. This perception and action link, which leads to this motion, is largely unconscious and normally largely suppressed in the listener, "its only outward manifestation [being] subtle tensions of the muscles" (Truslit, 1938; Repp, 1993). Further study is needed to investigate the effect of tone frequency on dwell time under sub-second and peri-second range.

8.4. Motor Variability

In this study, investigator examines the role of tone frequency and ISI on motor variability under two different time scale (1 -second vs supra-second range). Investigators have taken variability of inter-tap interval as an index of variability. The results revealed that tone frequency has shown significant effect on motor variability under supra-second range. But no such effect was found in peri-second range. In supra-second range, participants showed more motor variability with low-frequency tones as compared to high-frequency tones. While ISI has no significant impact on motor variability which rejects our hypothesis that was based on the fact that short ISI (which comes under 1-second range) and long ISI (which comes under supra-second range) has different involvement of attention. This implies that a 1-second range which is a transit zone between sub-second and supra second range also involves cognitive processes like the supra-second range and this corroborates with Rammasyer & Troche (2014) which states that 1-second range involves both cognitive and perceptual process.

The results revealed that tone frequency has shown significant effect on motor variability under supra-second range. But no such effect was found in peri-second range. It is because supra-second range involves more top-down attention and central planning circuits. Variability in central planning circuits may be better suited for learning-related motor exploration. These circuits have more ready access to reinforcement signals (Björklund & Dunnett, 2007; Schultz, 1998), and show more experience-dependent plasticity (Doyon & Benali, 2005; Nudo et al., 1996; Sanes & Donoghue, 2000).

In supra-second range, participants showed more motor variability with low-frequency tones as compared to high-frequency tones because MV is context sensitive across basal ganglia network and basal ganglia plays an important role in supra-second timing unlike sub-second and 1-second range (peri second range that's why frequency of tone (which is perceptual property of the sound) influenced MV in supra-second range unlike Sub-Second and 1-second range. Because in sub and 1-second range only cerebellum is involved.

MV facilitates learning in supra second range where subjects' attention is more involved .Due to this, motor control is regulated by feedback mechanism which is a closed loop system. In the feedback system, error gets fed back to the input to reduce the error between target and input and causes the motor variability. In the transient phase, motor variability is needed to facilitate learning.

Variability in central planning circuits may be more suited for learning-related motor exploration since supra-second range incorporates central planning circuits. More reinforcement signals are available to these circuits. We found that variability with LF tone is comparatively more than that of HF tone. This implies that LF tone captures the individual's selective attention because of its pleasant effect. MV is goal oriented (to minimize the synchronization error) under both LF and HF cases so variability will be there. But since LF one has a pleasing effect so individual's attention will orient to it selectively so variability with LF will be comparatively more the HF tone. This implies that LF tones are best suited for motor learning under supra-second range.

Chapter 9

Summary and Conclusions

This chapter summarizes the major findings of the current research. It also discusses the implications of the findings, limitations of the present research, and provides insights for future research.

9.1. An overview of the findings

This dissertation provides an understanding of the role of frequency of tone in affecting various parameters of AMS such as tapping behavior, motor timing, motor variability (MV), listening effort, and tapping performance in response to isochronous sound sequence. For scientifically investigating the role of frequency of tone under peri and supra-second range involving different level of cognitive resources, this study employs synchronized tapping paradigm.

- The results of the present study demonstrates significant occurrence of anticipatory tapping and reactive tapping under short ISI (1000 ms) and long ISI (2000 ms) respectively. In anticipatory tapping, the feed-forward motor control mechanism plays a critical role (unlike the case of reactive tapping in which the motor control is dependent on the feed-back motor control mechanism). In terms of implications, the findings of this study indicate that ISI in isochronous sound sequence should be given due consideration in designing movement studies and incorporate use of the same for therapeutic purposes such as rhythmic auditory stimulation (RAS).
- It is observed that motor response is more easily synchronized with low-frequency (LF) tones as compared

to high-frequency (HF) tones without expending much listening effort while responding to isochronous sound sequence. In terms of implications, this study suggest that tone frequency in isochronous sound sequence should be given due consideration in designing movement therapies such as RAS.

- Investigator found the strong perception-action coupling in the peri-second range unlike supra-second range. Participants over translated the ISI with LF tone and under translated ISI with HF tone in the peri-second range while this trend is absent in supra-second range.
- Further understanding with respect to motor timing is that individual’s shows higher dwell time for LF tone as compared to HF tone. This implies that individuals show more engagement with low-frequency tone.
- The results suggest that pitch has a greater influence on motor timing within the peri-second range as compared to the supra-second range. In terms of implications, this study's findings indicate that ISI should, therefore, be a significant factor in the design of sound sequence to be used for RAS.
- The results revealed that tone frequency has shown significant effect on MV under supra-second range. But no such effect was found in peri-second range. In supra-second range, participants showed more MV with LF tones as compared to high-frequency tones.
- Further understanding with respect to MV is that MV is error driven not stimulus driven under peri-second range whereas MV is both error and stimulus driven under supra-second range.
- While ISI has no significant impact on MV which implies that participants showed similar motor timing in both short ISI and long ISI. It implies that 1-s range also involves cognitive control.

- In terms of the synchronization error and listening effort, no significant effect of ISI was found, which rejects our hypothesis that was based on the fact that Short ISI (which comes under 1-second range) and long ISI (which comes under supra-second range) has different involvement of attention. This implies that a 1-second range which is a transit zone between sub-second and supra second range also involves cognitive processes like the supra-second range and this corroborates with Rammasyer's finding which states that 1-second range involves both cognitive and perceptual process (Rammasyer & Troche, 2014).

9.2. Limitations and implications of the Study

The thesis besides being an addition to the existing literature on the phenomenon of auditory-motor-synchronization, however, has some limitations that could be overcome in future research.

The participants who voluntarily took part in this research were healthy individuals belonging to the investigator's educational institute. Further investigation involving PD patients would have provided better insights in understanding AMS. This would have led to the possibility of comparing healthy individuals vs. PD patients. In terms of collecting pupillary dilation data, the instrument used in this research collected data at the rate of 30 samples per second, which is considered a poor sampling rate. There are high chances that some significant amount of crucial pupil data might not have been captured by the eye tracking instrument. Eye tracking system with higher sampling rate will benefit future studies.

With respect to the implication of this research, synchronized tapping paradigm could be used by researchers for investigating issues related to AMS. Furthermore, the results of this research signify the importance of LF tones related to AMS under peri-second and supra-second range. In addition to this, the findings of the current research provide further behavioral evidence in support of the dissociation between peri-second and supra-second range

The framework reported in this study could be useful for researchers who are interested in investigating issues related to AMS. Furthermore, the results of this research signify the importance of LF tones in regulation of motor control and MV under supra-second range. Since, supra-second range is suitable for motor learning it will be interesting to investigate how participants' tapping ability is altered under short ISI after receiving motor training in the long ISI. This study was mainly confined to movement timing and its relation to tone frequency and ISI in isochronous sound sequence.

Future study needs to explore non-isochronous sound sequence. It is easy to synchronize with the isochronous sound sequence because of its fixed ISI. But it becomes comparatively difficult to synchronize and adapt with the non-isochronous sound sequence because of its varied tempo. While synchronizing with isochronous sound sequence, phase correction response is required unlike non-isochronous sound sequence. Another aspect of future investigation is movement kinematics i.e., emphasizing on the geometric and time-dependent aspects of motion which will throw more light on the relationship between the pitch contour and the kinematics of movement.

9.3. Scope for Future Research

In future studies, it will be important to examine issues concerning sensory-automatic processing in peri-second range which is a transit zone between sub-second range and the supra-second range. In this current study, participants were given auditory feedback but what will happen when they will be given auditory and tactile feedback simultaneously. Since low-frequency tones has vibrational properties as compared to HF tones so it is most likely that LF tones will be suitable for auditory-tactile feedback but investigation is needed in further studies. Also, it is widely known that during grip strength training that individuals feels mental and physical fatigue which can be reduce via low-frequency driven auditory -tactile feedback because individuals shows less listening effort with low-frequency tones.

- This study has some limitations, such as the inability to access the neural correlates behind this phenomenon due to limited resources. Second, it was being done with a monotone sequence, not with an isochronous series of tones having different pitches. Another is the sample size, as only 15 participants have gone through a single ISI. So this finding cannot generalize as the sample size is not large enough. As only a single level of frequency had been taken in both LF and HF range, so it will be interesting to see that, up to what range of tone frequency, this behavior is shown needs to be investigated in the future. Another research question that came out from our study is that what will happen to motor timing when two successive tones are given in the same metronome in an alternative manner?
- This study was mainly confined to movement timing and its relation to tone frequency and ISI in isochronous sound sequence. But how these parameters affects the movement timing in non-isochronous sound sequence is yet to be explored. It is easy to synchronize with the isochronous sound sequence because of its fixed ISI. But it becomes comparatively difficult to synchronize and adapt with the non-isochronous sound sequence because of its varied tempo. While synchronizing with isochronous sound sequence, only phase correction response is required unlike non-isochronous sound sequence. Because it requires additional error correction which is called as period correction response this requires conscious control. Some studies has been done on AMS parameter like period correction and phase correction response and its relation to the tempo changes in auditory sequence.
- Another area of future investigation is movement kinematics which is not been addressed in this study. Movement

Kinematics is the study of the geometric and time-dependent aspects of motion without analyzing the forces causing the motion. There have been many studies that suggest a relationship between the pitch contour and the kinematics of the movement.

- Results suggests that supra second range is the best temporal window for auditory driven motor learning via low-frequency tones but it require further studies for reaching to an an appropriate conclusion.
- Since our study involves only peri-second and supra second temporal window not sub second window so future studies should take all the possible temporal window to analyses the motor variability (an indicator of learning). This study suggests that supra -second and peri-second window is appropriate window for cognitive and associative phases whereas sub-second temporal window is needed for autonomous phase for motor learning.
- Further study is needed to examine the effect of low-frequency tone on dwell and movement time.

Chapter 10

Appendices

Informed consent form

You are being invited to participate in a research study, which the Institutional Human Ethics Committee (IHEC), Indian Institute of Technology Indore has reviewed and approved for conduct by the investigators named here. The Investigator will describe this study to you and answer any of your questions. If you have any questions or complaints about the informed consent process of this research study or your rights as a subject, please contact the IHEC at the Indian Institute of Technology, Indore

Thank you for agreeing to participate in this research project. This study involves scientific enquiry concerning effect of tone frequency and inter-stimulus interval on AMS. There will be no risks with participation in this research study. Your details will not be revealed when the data is presented or reported. All data will be the property of Human Factors & Applied Cognition Lab IIT Indore so you cannot claim any share or right in terms of acknowledgment or authorship of the research paper(s) published on the basis of the data collected in this study.

Your participation in this study is completely voluntary. Should you decide to discontinue participation or decline to answer any specific part of the study. You will still receive a T-Shirt as compensation that in exchange for your participation. You will also receive a thank you card with experimenter's contact details. Should you have any further questions, please feel free to contact the experimenter.

Consent statement

I, ___________________________________, hereby give my consent to participate in the research study entitled "effect of tone frequency present in isochronous sound sequences and Inter-stimulus interval on synchronization error". I have read the above information provided by the experimenter for the participation in their study and I agree to participate.

Participant's signature Date

I hereby certify that I have given an explanation to the above individual of the contemplated study and its risks and potential.

Experimenter's signature Date

Instructions

You are most welcome to our HFAC Lab for Auditory cognition study. The Isochronous sound sequence is designed in HFAC lab and you are not supposed to disclose the details of the isochronous sound sequence to other participants to avoid the presumption. For this study, you are required to listen to the isochronous sound sequences and tap the tapping board by your dominant index finger in response to tone onset present in rhythm track in a comfortable manner. During the listening, we will also record your pupil diameter from Tobii TX30 eye tracker in the continuous manner which needs calibration which may take 2-5 minutes.

There will be a rhythm track consisting of isochronous sound sequences of different ISI. You have to listen to the rhythm track and follow the instructions which are mentioned below.

- Sit on the chair comfortably. Keep your spine in a vertical position.
- Look into the center point of the screen.
- Wear the headset.
- While calibration, you need to gaze your eyes to the red point which will move towards all corners of the screen.
- Keep your finger on the piezoelectric sensor. Tap your finger on the middle white circular part of the piezoelectric sensor.
- A video with black screen augmented with an auditory metronome will be presented in the form of visual stimuli in front of you on the Tobii screen. First, 33 seconds of

video will contain no auditory information and visual information. This offset time will be allotted for dark adaptation of the pupil.

- You will be listening sound sequence after 33 seconds and you will have to respond to the sound by tapping your finger on the piezoelectric sensor.
- In between the recording, don't shake your head and fix your chin with the support of your hand.
- You have to place your cell phone, other magnetic/electronic devices; metallic objects out of the recording room, in a locker and you have the key with yourself.
- Don't do eye makeup like mascara etc.
- If you need help or feel discomfort at any point in between the experiment, you can raise your hand.
- If you want to escape from the experiment because of any discomfort with listening, you can withdraw from the study at any time.

Thank you for your acceptance to participate in our experiment

References

Alain, C., Woods, D. L., & Knight, R. T. (1998). A distributed cortical network for auditory sensory memory in humans. *Brain Research*, *812*(1-2), 23–37. https://doi.org/10.1016/s0006-8993 (98)00851-8

Albantakis, L., & Deco, G. (2011). Changes of Mind in an Attractor Network of Decision-Making. *PLoS Computational Bolgy*, *7*(6). https://doi.org/10.1371/journal.pcbi.1002086

Ammirante, P., Thompson, W. F., & Russo, F. A. (2011). Ideomotor effects of pitch on continuation tapping. *Quarterly Journal of Experimental Psychology*, *64*(2), 381-393. https://doi.org/10.1080/17470218.2010.495408

Arvai, J. (2013). Thinking, fast and slow, Daniel Kahneman, Farrar, Straus & Giroux. *Journal of Risk Research*, *16*(10), 1322-1324. https://doi.org/10.1080/13669877.2013.766389

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition*, *48*(1), 66-79. https://doi.org/10.1006/brcg.2001.1304

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception & Psychophysics*, *57*(3), 305-317. https://doi.org/10.3758/bf03213056

Austin, D. (2013). *On the role of attention in finger tapping and towards continuous, objective, and unobtrusive measurement of motor speed* (Doctoral dissertation, PhD Thesis, Department of

Biomedical Engineering, Oregon Health & Science University, Portland, OR).

Balasubramaniam, R., Wing, A., & Daffertshofer, A. (2004). Keeping with the beat: Movement trajectories contribute to movement timing. *Experimental Brain Research*, *159*(1), 129-134. https://doi.org/10.1007/s00221-004-2066-z

Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *Journal of Neuroscience*, *34*(11), 3910-3923. https://doi.org/10.1523/jneurosci.2679-13.2014

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276-292. https://doi.org/10.1037/0033-2909.91.2.276

Beckers, T., De Houwer, J., & Eelen, P. (2002). Automatic integration of non-perceptual action effect features: The case of the associative affective Simon effect. *Psychological Research*, *66*(3), 166-173. https://doi.org/10.1007/s00426-002-0090-9

Blais, M., Albaret, J.-M., & Tallet, J. (2015). Is there a link between sensorimotor coordination and inter-manual coordination? Differential effects of auditory and/or visual rhythmic stimulations. *Experimental Brain Research*, *233*(11), 3261–3269.

Boasson, A. D., & Granot, R. (2012, July). Melodic direction's effect on tapping. In *12th International Conference on Music Perception and Cognition, Thessaloniki, Greece*.

Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology*, *6*(2), 145. https://doi.org/10.2307/1410948

Bradley, M. M., Sapigao, R. G., & Lang, P. J. (2017). Sympathetic ANS modulation of pupil diameter in emotional scene perception: Effects of hedonic content, brightness, and contrast. *Psychophysiology*, *54*(10), 1419–1435.

Brass, M., Zysset, S., & Von Cramon, D. (2001). The inhibition of imitative response tendencies. *NeuroImage*, *14*(6), 1416-1423. https://doi.org/10.1006/nimg.2001.0944

Brower, C. (1993). Memory and the perception of rhythm. *Music Theory Spectrum*, *15*(1), 19-35. https://doi.org/10.1525/mts.1993.15.1.02a00020

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, *10*(1), 12–21. https://doi.org/10.1080/17470215808416249

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience*, *20*(2), 204-214. https://doi.org/10.1162/jocn.2008.20017

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755-765. https://doi.org/10.1038/nrn1764

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00183

Burghardt, H. (1972). [Simple function scheme for the description of the subjective duration of sound impulses and pauses]. *Kybernetik*, *12*(1), 11–29. https://doi.org/10.1007/bf00289233

Butz, M. V., Sigaud, O., & Gérard, P. (2003). Anticipatory behavior: exploiting Knowledge about the Future to Improve Current Behavior. *Anticipatory Behavior in Adaptive Learning Systems Lecture Notes in Computer Science*, 1–10. doi:10.1007/978-3-540-45002-3_1

Carlini, A., & French, R. (2014). Visual tracking combined with hand-tracking improves time perception of moving stimuli. *Scientific Reports*, 4(1). https://doi.org/10.1038/srep05363

Cavanagh, P. R., & Kram, R. (1985). The efficiency of human movement-a statement of the problem. *Medicine & Science in Sports & Exercise*, *17*(3), 304-308. https://doi.org/10.1249/00005768-198506000-00002

Chen, Y., Repp, B. H., & Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, *21*(4), 515–532.

Cohen, R. A. (2018). Yerkes-Dodson law. *Encyclopedia of Clinical Neuropsychology*, 3761-3762. https://doi.org/10.1007/978-3-319-57111-9_1340

Coull J. T., Cheng R. K., Meck W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology*, 36, 3–25. 10.1038/npp.2010.113

Coull, J. T., Charras, P., Donadieu, M., Droit-Volet, S., & Vidal, F. (2015). SMA selectively codes the active accumulation of temporal, not spatial, magnitude. *Journal of Cognitive Neuroscience*, *27*(11), 2281-2298. https://doi.org/10.1162/jocn_a_00854

Coutinho, E., & Cangelosi, A. (2009). The use of spatio-temporal Connectionist models in psychological studies of musical emotions. *Music Perception*, *27*(1), 1-15. https://doi.org/10.1525/mp.2009.27.1.1

Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1673-1692. https://doi.org/10.1037/0096-1523.25.6.1673

Crowe D. A., Zarco W., Bartolo R., Merchant H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J. Neurosci, 34,* 11972–11983.

Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial Premotor cortex. *Journal of Neuroscience*, *34*(36), 11972-11983. https://doi.org/10.1523/jneurosci.2177-14.2014

Darwin, C., & Darwin, F. (1896). The life and letters of Charles Darwin: Including an autobiographical chapter /. https://doi.org/10.5962/bhl.title.49491

Deco, G., Rolls, E. T., & Romo, R. (2010). Synaptic dynamics and decision making. *Proceedings of the National Academy of Sciences*, *107*(16), 7545–7549. doi:10.1073/pnas.1002333107

Droit-Volet, S., & Meck, W. H. (2007). How emotions color our perception of time. *Trends in Cognitive Sciences*, *11*(12), 504-513. https://doi.org/10.1016/j.tics.2007.09.008

Dunlap, K. (1910). Reaction to rhythmic stimuli with attempt to synchronize. *Psychological Review*, *17*(6), 399–416. https://doi.org/10.1037/h0074736

Eckert, M. A., Teubner-Rhodes, S., & Vaden, K. I., Jr. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear and Hearing*, *37*, 101S–110S.

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, *114*(3), 405-422. https://doi.org/10.1016/j.cognition.2009.10.013

Fairclough, S. H., & Houston, K. (2004). A metabolic measure of mental effort. *Biological Psychology*, *66*(2), 177-190. https://doi.org/10.1016/j.biopsycho.2003.10.001

Fraisse, P. (1982). Rhythm and tempo. *Psychology of Music*, 149-180. https://doi.org/10.1016/b978-0-12-213562-0.50010-3

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211. https://doi.org/10.1098/rstb.2008.0300

Gernigon, C., D'Arripe-Longueville, F., Delignières, D., & Ninot, G. (2004). A dynamical systems perspective on goal involvement states in sport. *Journal of Sport and Exercise Psychology*, *26*(4), 572-596. https://doi.org/10.1123/jsep.26.4.572

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). undefined. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252-269. https://doi.org/10.3758/cabn.10.2.252

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, *21*(2), 90-95. https://doi.org/10.1177/0963721412436811

Gooch, C. M., Wiener, M., Wencil, E. B., & Coslett, H. B. (2010). Interval timing disruptions in subjects with cerebellar lesions. *Neuropsychologia*, *48*(4), 1022–1031. https://doi.org/10.1016/j.neuropsychologia.2009.11.028

Grondin, S. (2014). About the (Non)scalar property for time perception. *Advances in Experimental Medicine and Biology*, 17-32. https://doi.org/10.1007/978-1-4939-1782-2_2

Harrington, D. L., Haaland, K. Y., & Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *The Journal of Neuroscience*, *18*(3), 1085-1095. https://doi.org/10.1523/jneurosci.18-03-01085.1998

Hawkins, J., George, D., & Niemasik, J. (2009). Sequence memory for prediction, inference and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1209. https://doi.org/10.1098/rstb.2008.0322

Hayashi M. J., Kantele M., Walsh V., Carlson S., Kanai R. (2014). Dissociable neuroanatomical correlates of subsecond and suprasecond time perception. *J. Cogn. Neurosci., 26*, 1685–1693. 10.1162/jocn_a_00580

Hazeltine, E., Helmuth, L. L., & Ivry, R. B. (1997). Neural mechanisms of timing.*Trends in Cognitive Sciences*, *1*(5), 163-169. https://doi.org/10.1016/s1364-6613(97)01058-9

Hellbrück, J., & Moser, L. M. (1986). Psychophysical measurements of inner ear processes with category scaling of loudness. *ORL*, *48*(2), 135-136. https://doi.org/10.1159/000275858

Hellström, Å., & Rammsayer, T. (2002). Mechanisms behind discrimination of short and long auditory durations. In *Annual meeting of the International Society for Psychophysics* (pp. 110-115). The International Society for Psychophysics Rio de Janeiro, Brazil.

Herwig, A., & Waszak, F. (2012). Action-effect bindings and ideomotor learning in intention- and stimulus-based actions. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00444

Hess, T. M., & Ennis, G. E. (2011). Age differences in the effort and costs associated with cognitive activity. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67*(4), 447-455. https://doi.org/10.1093/geronb/gbr129

Hommel, B. (2009). Conscious and unconscious control of spatial action. *Encyclopedia of Consciousness*, 171-181. https://doi.org/10.1016/b978-012373873-8.00022-0

Hommel, B., Alonso, D., & Fuentes, L. (2003). Acquisition and generalization of action effects. *Visual Cognition*, *10*(8), 965-986. https://doi.org/10.1080/13506280344000176

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849-878. https://doi.org/10.1017/s0140525x01000103

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

Hornsby, B. W., Naylor, G., & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing*, 37, 136S–144S.

Hove, M. J., Marie, C., Bruce, I. C., & 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*, *111*(28), 10383-10388. https://doi.org/10.1073/pnas.1402039111

Hyönä, J., Tommola, J., & Alaja, A. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology Section A*, *48*(3), 598-612. https://doi.org/10.1080/14640749508401407

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception, 23*(4), 319- 330. https://doi.org/10.1525/mp.2006.23.4.319

Iosa, M., Gizzi, L., Tamburella, F., & Dominici, N. (2015). Editorial: Neuro-motor control and feed-forward models of locomotion in humans. *Frontiers in Human Neuroscience, 9*, https://doi.org/10.3389/fnhum.2015.00306

Ishikawa, T., Tomatsu, S., Izawa, J., & Kakei, S. (2016). The cerebro-cerebellum: Could it be loci of forward models? *Neuroscience Research*, *104*, 72–79. doi: 10.1016/j.neures.2015.12.003

Ito, M. (2005). Bases and implications of learning in the cerebellum — adaptive control and internal model mechanism.*Progress in Brain Research Creating Coordination in the Cerebellum*, 95–109. https://doi.org/10.1016/s00796123 (04)48009-1

Jahanshahi, M., Jones, C. R., Dirnberger, G., & Frith, C. D. (2006). The Substantia Nigra pars Compacta and temporal processing. *Journal of Neuroscience, 26*(47), 12266-12273. https://doi.org/10.1523/jneurosci.2540-06.2006

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of*

Experimental Psychology: General, *141*(1), 54-75. https://doi.org/10.1037/a0024208

Jeon, J. Y., & Fricke, F. R. (1997). Duration of perceived and performed sounds. *Psychology of Music*, *25*(1), 70-83. https://doi.org/10.1177/0305735697251006

Jin, X., Wang, B., Lv, Y., Lu, Y., Chen, J., & Zhou, C. (2019). Does dance training influence beat sensorimotor synchronization? Differences in finger-tapping sensorimotor synchronization between competitive ballroom dancers and nondancers. *Experimental Brain Research*, *237*(3), 743–753. https://doi.org/10.1007/s00221-018-5410-4

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 56-88. https://doi.org/10.1080/14639220210159735

Kacelnik A., Brunner D. (2002). Timing and foraging: Gibbon’s scalar expectancy theory and optimal patch exploitation. *Learn. Motiv.*, *33*, 177–195. https://doi.org/10.1006/lmot.2001.1110

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall

Kawaguchi, H., Matsumoto, T., Takeuchi, Y., Kudo, H., & Ohnishi, N. (2010). Factors of unpleasantness in stick-slip sounds —Peak frequency and envelope. *IEEJ Transactions on Electronics, Information and Systems*, *130*(12), 2180-2181. https://doi.org/10.1541/ieejeiss.130.2180

Keele, S. W., Nicoletti, R., Ivry, R. I., & Pokorny, R. A. (1989). Mechanisms of perceptual timing: Beat- based or interval-based judgements? *Psychological Research*, *50*(4), 251-256. https://doi.org/10.1007/bf00309261

Keller, P. E., & Rieger, M. (2009). Special issue—Musical movement and synchronization. *Music Perception*, *26*(5), 397-400. https://doi.org/10.1525/mp.2009.26.5.397

Klein, R. M. (1976). Attention and Movement. *Motor Control*, 143–173. https://doi.org/10.1016/b978-0-12-665950-4.50011-0

Knuf, L., Aschersleben, G., & Prinz, W. (2001). An analysis of ideomotor action. *Journal of Experimental Psychology: General*, *130*(4), 779-798. https://doi.org/10.1037/0096-3445.130.4.779

Koch, G., Oliveri, M., Torriero, S., Salerno, S., Gerfo, E. L., & Caltagirone, C. (2006). Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research*, *179*(2), 291–299. https://doi.org/10.1007/s00221-006-0791-1

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hearing Research*, *312*, 114-120. https://doi.org/10.1016/j.heares.2014.03.010

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, *97*(2), 253-270. https://doi.org/10.1037/0033-295x.97.2.253

Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Kuik, D. J. (1997). Assessing aspects of auditory handicap by means of pupil dilatation. *International Journal of Audiology*, *36*(3), 155-164. https://doi.org/10.3109/00206099709071969

Kurakata, K., Mizunami, T., & Matsushita, K. (2013). Sensory unpleasantness of high-frequency sounds. *Acoustical Science and Technology*, *34*(1), 26-33. https://doi.org/10.1250/ast.34.26

Lake, J. I., LaBar, K. S., & Meck, W. H. (2016). Emotional modulation of interval timing and time perception. *Neuroscience & Biobehavioral Reviews*, *64*, 403-420. https://doi.org/10.1016/j.neubiorev.2016.03.003

Lehiste, I. (1976). Influence of fundamental frequency pattern on the perception of duration. *Journal of Phonetics*, *4*(2), 113-117. https://doi.org/10.1016/s0095-4470(19)31231-8

Leman, M. (2008). *Embodied music cognition and mediation technology*. MIT Press.

Lemon, R. N. (2008). Descending pathways in motor control. *Annual Review of Neuroscience*, *31*(1), 195-218. https://doi.org/10.1146/annurev.neuro.31.060407.125547

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current Opinion in Neurobiology*, *13*(2), 250-255. https://doi.org/10.1016/s0959-4388(03)00036-9

Lewis, P. A., & Miall, R. C. (2006). Remembering the time: A continuous clock. *Trends in Cognitive Sciences*, *10*(9), 401-406. https://doi.org/10.1016/j.tics.2006.07.006

Lewis, P., & Miall, R. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia*, *41*(12), 1583-1592. https://doi.org/10.1016/s0028-3932(03)00118-0

Lezama-Espinosa, C., & Hernandez-Montiel, H. (2020). Neuroscience of the auditory-motor system: How does sound interact with movement? *Behavioural Brain Research*, *384*, 112535. https://doi.org/10.1016/j.bbr.2020.112535

Lu, Z. L., & Sperling, G. (2003). Measuring sensory memory: Magnetoencephalography habituation and psychophysics. *Magnetic source imaging of the human brain*, 319-342. https://doi.org/10.4324/9781410609175-18

Malapani, C., Dubois, B., Rancurel, G., & Gibbon, J. (1998). Cerebellar dysfunctions of temporal processing in the seconds range in humans. *Neuro Report*, *9*(17), 3907–3912. https://doi.org/10.1097/00001756-199812010-00026

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978. https://doi.org/10.1080/01690965.2012.705006

Mazzoni, D., Dannenberg, R. Audacity [software]. The Audacity Team. Pittsburg, PA, USA, 2000. Retrieved, June 20, 2020 from https://www.audacityteam.org/

McCown, W., Keiser, R., Mulhearn, S., & Williamson, D. (1997). The role of personality and gender in preference for exaggerated bass in music. *Personality and Individual Differences*, *23*(4), 543-547. https://doi.org/10.1016/s0191-8869 (97)00085-8

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British society of audiology cognition in hearing special interest group 'white paper'. *International Journal of Audiology*, *53*(7), 433-445. https://doi.org/10.3109/14992027.2014.890296

Meltzoff, A. N., & Prinz, W. (2002). *The imitative mind: Development, evolution and brain bases*. Cambridge University Press.

Merchant H., Georgopoulos A. P. (2006). Neurophysiology of perceptual and motor aspects of interception. *Journal of Neurophysiology, 95,* 1–13. https://doi.org/10.1152/jn.00422.2005

Merchant H., Harrington D. L., Meck W. H. (2013a). Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci., 36*, 313–336. 10.1146/annurev-neuro-062012-170349

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

Merchant, H., Zarco, W., & Prado, L. (2008). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *Journal of Neurophysiology*, *99*(2), 939-949. https://doi.org/10.1152/jn.01225.2007

Merker, B. (1999). Synchronous chorusing and the origins of music. *Musicae Scientiae*, *3*(1_suppl), 59-73. https://doi.org/10.1177/10298649000030s105

Michon, J. A. (1985). The complete time Experiencer. *Time, Mind, and Behavior*, 20-52. https://doi.org/10.1007/978-3-642-70491-8_2

Miyake, I. (1902). Researches on rhythmic action. Studies from the Yale Psychology Laboratory, 10, 1–48.

Miyake, Y., Onishi, Y., & Pöppel, E. (2002). Two Types of Anticipatory-Timing Mechanisms in Synchronization Tapping.*Object Recognition, Attention, and Action*, 231–244. doi: 10.1007/978-4-431-73019-4_16

Montero, B. G. (2011). Effortless bodily movement. *Philosophical Topics*, *39*(1), 67-79. https://doi.org/10.5840/philtopics201139114

Montero, B. G. (2016). Aesthetic effortlessness. *Body Aesthetics*, 180-191. https://doi.org/10.1093/acprof:oso/9780198716778.003.0011

Moore, B. C. (2012). *An introduction to the psychology of hearing*. BRILL.

Morillon, B., Kell, C. A., & Giraud, A. (2009). Three stages and four neural systems in time estimation. *Journal of Neuroscience*, *29*(47), 14803-14811. https://doi.org/10.1523/jneurosci.3222-09.2009

Murai, Y., & Yotsumoto, Y. (2016). Context-Dependent Neural Modulations in the Perception of Duration. *Frontiers in Integrative Neuroscience*, *10*. https://www.frontiersin.org/articles/10.3389/fnint.2016.00012

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 861-872. https://doi.org/10.1037/0096-1523.23.3.861

O’Regan, L., Spapé, M. M., & Serrien, D. J. (2017). Motor timing and Covariation with time perception: Investigating the role of handedness. *Frontiers in Behavioral Neuroscience*, *11*. https://doi.org/10.3389/fnbeh.2017.00147

Oberauer, K., & Lewandowsky, S. (2014). Further evidence against decay in working memory. *Journal of Memory and Language*, *73*, 15–30. https://doi.org/10.1016/j.jml.2014.02.003

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, *134*(3), 368-387. https://doi.org/10.1037/0096-3445.134.3.368

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*(2), 226-238. https://doi.org/10.1007/s00221-004-2159-8

Pearson, K., Gordon, J. Spinal reflexes: Principles of Neuroscience, New York, NY: McGraw-Hill, 2000

Phillips, N. A. (2016). The implications of cognitive aging for listening and the framework for understanding effortful listening (FUEL). *Ear & Hearing*, *37*(1), 44S-51S. https://doi.org/10.1097/aud.0000000000000309

Pichora-Fuller, M. K. (2016). How social psychological factors may modulate auditory and cognitive functioning during listening. *Ear and Hearing*, 37, 92S–100S.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G. (2016). Hearing impairment and cognitive energy. *Ear and Hearing*, *37*, 5S-27S. https://doi.org/10.1097/aud.0000000000000312

Pisotta, I., & Molinari, M. (2014). Cerebellar contribution to feedforward control of locomotion. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00475

Plack, C. J., & Oxenham, A. J. (2005). The psychophysics of pitch. *Pitch*, 7-55. https://doi.org/10.1007/0-387-28958-5_2

Pouthas, V., George, N., Poline, J., Pfeuty, M., VandeMoorteele, P., Hugueville, L., Ferrandez, A., Lehéricy, S., LeBihan, D., & Renault, B. (2005). Neural network involved in time perception: An fMRI study comparing long and short interval estimation. *Human Brain Mapping*, *25*(4), 433-441. https://doi.org/10.1002/hbm.20126

Prassas, S., Thaut, M., McIntosh, G., & Rice, R. (1997). Effect of auditory rhythmic cuing on gait kinematic parameters of stroke patients. *Gait & Posture*, *6*(3), 218-223. https://doi.org/10.1016/s0966-6362(97)00010-6

Pressing, J. (1999). The referential dynamics of cognition and action. *Psychological Review*, *106*(4), 714-747. https://doi.org/10.1037/0033-295x.106.4.714

Prinz, J. J. (2014). *Beyond human nature: How culture and experience shape the human mind*. W. W. Norton & Company.

Prinz, W. (1984). Modes of linkage between perception and action. *Cognition and Motor Processes*, 185-193. https://doi.org/10.1007/978-3-642-69382-3_13

Prinz, W. (1990). A common coding approach to perception and action. *Relationships Between Perception and Action*, 167-201. https://doi.org/10.1007/978-3-642-75348-0_7

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129-154. https://doi.org/10.1080/713752551

Pruszynski, J. A., Kurtzer, I., Nashed, J. Y., Omrani, M., Brouwer, B., & Scott, S. H. (2011). Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature*, *478*(7369), 387-390. https://doi.org/10.1038/nature10436

Rajendran, V. G., & Teki, S. (2016). Periodicity versus Prediction in Sensory Perception. *Journal of Neuroscience*, *36*(28), 7343–7345. https://doi.org/10.1523/jneurosci.1335-16.2016

Rajendran, V. G., Teki, S., & Schnupp, J. W. (2018). Temporal processing in audition: Insights from music. *Neuroscience*, *389*, 4–18.

Ramachandran, V. (1996). Short note: On the unpleasantness of certain harsh sounds. *Medical Hypotheses*, *46*(5), 487. https://doi.org/10.1016/s0306-9877(96)90031-6

Rammsayer T., Lima S. D. (1991). Duration discrimination of filled and empty auditory intervals: cognitive and perceptual factors. *Percept.Psychophys, 50*, 564–57. https://doi.org/10.3758/BF03207541

Rammsayer, T. H., & Troche, S. J. (2014). In search of the internal structure of the processes underlying interval timing in the sub-second and the second range: A confirmatory factor analysis approach. *Acta Psychologica*, *147*, 68–74. https://doi.org/10.1016/j.actpsy.2013.05.004

Rammsayer, T., & Ulrich, R. (2011). Elaborative rehearsal of nontemporal information interferes with temporal processing of durations in the range of seconds but not milliseconds. *Acta Psychologica*, *137*(1), 127–133. https://doi.org/10.1016/j.actpsy.2011.03.010

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics*, *50*(6), 565-574. https://doi.org/10.3758/bf03207541

Ravignani, A., Honing, H., & Kotz, S. A. (2017). The evolution of rhythm cognition: Timing in music and speech. *Frontiers in Human Neuroscience*, *11*, 303.

Refsum Jensenius, A. (2017). Exploring music-related micromotion. *Body, Sound and Space in Music and Beyond: Multimodal Explorations*, 29-48. https://doi.org/10.4324/9781315569628-3

Reichenbach, T., & Hudspeth, A. J. (2012). Discrimination of low-frequency tones employs temporal fine structure. *PLoS ONE*, *7*(9), e45579. https://doi.org/10.1371/journal.pone.0045579

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452. https://doi.org/10.3758/s13423-012-0371-2

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969-992. https://doi.org/10.3758/bf03206433

Repp, B. H. (1993). Music as motion: A synopsis of Alexander Truslit's (1938) Gestaltung und Bewegung in Der Musik. *Psychology of Music*, *21*(1), 48-72. https://doi.org/10.1177/030573569302100104

Repp, B., & Penel, A. (2003). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research Psychologische Forschung*, *68*(4). https://doi.org/10.1007/s00426-003-0143-8

Richter, M. (2016). The moderating effect of success importance on the relationship between listening demand and listening effort. *Ear and Hearing*, *37*, 111S–117S

Robert J. Hockey, G. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, *45*(1-3), 73-93. https://doi.org/10.1016/s0301-0511(96)05223-4

Safranek, M. G., Koshland, G. F., & Raymond, G. (1982). Effect of auditory rhythm on muscle activity. *Physical Therapy*, *62*(2), 161-168. https://doi.org/10.1093/ptj/62.2.161

Sanjram, P. K., & Shivhare, Y. K. (2021). Emergence of Frequency-Dependent Motor Variability Within Supra-Second Auditory Cueing. *Perceptual and Motor Skills*, *128*(6), 2805–2819.

Sarter, M., Gehring, W. J., & Kozak, R. (2006). More attention must be paid: The neurobiology of attentional effort. *Brain Research Reviews*, *51*(2), 145-160. https://doi.org/10.1016/j.brainresrev.2005.11.002

Schack, T., & Mechsner, F. (2006). Representation of motor skills in human long-term memory. *Neuroscience Letters*, *391*(3), 77-81. https://doi.org/10.1016/j.neulet.2005.10.009

Schirmer A. (2004). Timing speech: a review of lesion and neuroimaging findings.*Brain Res. Cogn. Brain Res.,21*, 269–287. https://doi.org/10.1016/j.cogbrainres.2004.04.003

Schmidt, R. A. (1968). Anticipation and timing in human motor performance. *Psychological Bulletin*, *70*(6, Pt.1), 631–646. https://doi.org/10.1037/h0026740

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218. https://doi.org/10.1016/j.tics.2007.02.006

Schwarz, M. A., Winkler, I., & Sedlmeier, P. (2012). The heart beat does not make us tick: The impacts of heart rate and arousal on time perception. *Attention, Perception, & Psychophysics*, *75*(1), 182-193. https://doi.org/10.3758/s13414-012-0387-8

Seidler, R., Noll, D., & Thiers, G. (2004). Feedforward and feedback processes in motor control. *NeuroImage*, *22*(4), 1775-1783. https://doi.org/10.1016/j.neuroimage.2004.05.003

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, *33*(1), 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135

Skinner, B. (1981). Selection by consequences. *Science*, *213*(4507), 501-504. https://doi.org/10.1126/science.7244649

Sohn M. H., Carlson R. A. (2003). Implicit temporal tuning of working memory strategy during cognitive skill acquisition. *Am. J. Psychol*, *116*, 239–256. https://doi.org/10.2307/1423579

Sparrow, A. W., and K. M. Newell. 1998. Metabolic energy expenditure and the regulation of movement economy. Psychon. Bull. Rev. 5:173–196

Sparrow, A. W., K. M. Hughes, A. P. Russell, and P. F. Le Rossignol. 1999. Effects of practice and preferred rate on perceived exertion, metabolic variables and movement control. Hum. Mov. Sci.18:137–153.

Spencer R. M., Zelaznik H. N., Diedrichsen J., Ivry R. B. (2003). Disrupted timing of discontinuous but not continuous movements by cerebellar lesions Science 300, 1437–1439. https://doi.org/10.1126/science.108366

Spencer, H. (1890). I.—The origin of music. *Mind*, *os-XV*(60), 449-468. https://doi.org/10.1093/mind/os-xv.60.449

Spencer, R. M., Verstynen, T., Brett, M., & Ivry, R. (2007). Cerebellar activation during discrete and not continuous timed movements: An fMRI study. *NeuroImage*, *36*(2), 378-387. https://doi.org/10.1016/j.neuroimage.2007.03.009

Spapé, M. M., Harjunen, V. J., & Ravaja, N. (2022). Time to imagine moving: Simulated motor activity affects time perception. *Psychonomic Bulletin & Review*, *29*(3), 819–827. https://doi.org/10.3758/s13423-021-02028-2

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology*, 52(1), 77–86. https://doi.org/10.1016/j.ijpsycho.2003.12.005

Sternad, D. (2000). Debates in dynamics: A dynamical systems perspective on action and perception. *Human Movement Science*, *19*(4), 407–423.

Stevens, L. T. (1886). On the time-sense. *Mind*, *os-XI*(43), 393-404. https://doi.org/10.1093/mind/os-xi.43.393

Strauss, D. J., & Francis, A. L. (2017). Toward a taxonomic model of attention in effortful listening. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(4), 809-825. https://doi.org/10.3758/s13415-017-0513-0

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, *82*(2), 127-136. https://doi.org/10.1016/j.bandc.2013.03.003

Symes, E., Tucker, M., Ellis, R., Vainio, L., & Ottoboni, G. (2008). Grasp preparation improves change detection for congruent objects. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 854-871. https://doi.org/10.1037/0096-1523.34.4.854

Taub, E., Ellman, S. J., & Berman, A. J. (1966). Deafferentation in Monkeys: Effect on Conditioned Grasp Response. *Science*, *151*(3710), 593–594. https://doi.org/10.1126/science.151.3710.593

Teki, S., Grube, M., & Griffiths, T. D. (2012). A Unified Model of Time Perception Accounts for Duration-Based and Beat-Based Timing

Mechanisms. *Frontiers in Integrative Neuroscience*, *5*. https://doi.org/10.3389/fnint.2011.00090

Thaut, M. H., & Abiru, M. (2010). Rhythmic Auditory Stimulation in Rehabilitation of Movement Disorders: A Review of Current Research. *Music Perception*, *27*(4), 263–269. http://doi.org /10.1525/mp.2010.27.4.263

Thaut, M. H. (2005). The future of music in therapy and medicine. *Annals of the New York Academy of Sciences*, *1060*(1), 303-308. https://doi.org/10.1196/annals.1360.023

Thaut, M. H., McIntosh, G. C., Prassas, S. G., & Rice, R. R. (1992). undefined. *Neurorehabilitation and Neural Repair*, *6*(4), 185-190. https://doi.org/10.1177/136140969200600403

Thaut, M., Schleiffers, S., & Davis, W. (1991). Analysis of EMG activity in biceps and triceps muscle in an upper extremity gross motor task under the influence of auditory rhythm. Journal of Music Therapy, 28(2), 64-88. https://doi.org/10.1093/jmt/28.2.64

Treisman, M., Faulkner, A., & Naish, P. L. (1992). On the relation between time perception and timing of motor action: Evidence for a temporal oscillator controlling the timing of movement. *The Quarterly Journal of Experimental Psychology Section A*, *45*(2), 235-263. https://doi.org/10.1080/14640749208401326

Treisman, M., Faulkner, A., Naish, P. L., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception*, *19*(6), 705-742. https://doi.org/10.1068/p190705

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception: An Interdisciplinary Journal*, *30*(4), 349-359. https://doi.org/10.1525/mp.2013.30.4.349

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2012). The impact of the bass drum on human dance movement. *Music Perception*, *30*(4), 349-359. https://doi.org/10.1525/mp.2013.30.4.349

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00577

Wearden, J. H., & McShane, B. (1988). Interval production as an analogue of the peak procedure: Evidence for similarity of human and animal timing processes. *The Quarterly Journal of Experimental Psychology*, *40*(4), 363–375.

Wencil, E. B., Coslett, H. B., Aguirre, G. K., & Chatterjee, A. (2010). Carving the clock at its component joints: Neural bases for interval timing. *Journal of Neurophysiology*, *104*(1), 160-168. https://doi.org/10.1152/jn.00029.2009

Widmann, A., Schröger, E., & Wetzel, N. (2017). Emotional arousal in novelty processing is indicated by the P3a and by the sympathetic componentof the pupil dilation response. Psychophysiology, 54(supplement 1), S73.

Wiener M., Turkeltaub P., Coslett H. B. (2010). The image of time: a voxel-wise meta-analysis. *Neuroimage, 49,* 1728–1740. https://doi.org/10.1016/j.neuroimage.2009.09.064

Wiener, M., Turkeltaub, P., & Coslett, H. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage*, *49*(2), 1728-1740. https://doi.org/10.1016/j.neuroimage.2009.09.064

Wierda, S. M., van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporalresolution. Proceedings of the National Academy of Sciences of the United States of America,109(22), 8456–8460.

Wierda, S. M., Van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences*, *109*(22), 8456-8460. https://doi.org/10.1073/pnas.1201858109

Winkler, I., Glauer, M., Betsch, T., & Sedlmeier, P. (2015). The Impact of Attention on Judgments of Frequency and Duration. *PLOS ONE, 10*(5), e0126974. https://doi.org/10.1371/journal.pone.0126974

Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, *20*, 233121651666972. https://doi.org/10.1177/2331216516669723

Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing*, *36*(4), 153-165. https://doi.org/10.1097/aud.0000000000000145

Witt S. T., Laird A. R., Meyerand M. E. (2008). Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. *Neuroimage*, *42*, 343–356.

Wohlschläger, A. (1999). Synchronization error: An error in time perception. In *Abstracts of the Psychonomic Society* (Vol. 4, p.48)

Wolters, M., Mundt, H., & Riedmiller, J. (2010). Loudness normalization in the age of portable media players. *Audio Engineering Society Convention, 128*.

Woodrow, H. (1932). The effect of rate of sequence upon the accuracy of synchronization. *Journal of Experimental Psychology*, *15*(4), 357-379. https://doi.org/10.1037/h0071256

Yumoto N., Lu X., Henry T. R., Miyachi S., Nambu A., Fukai T., et al. (2011). A neural correlate of the processing of multi-second time intervals in primate prefrontal cortex. *PLoS ONE*6:e19168 https://doi.org/10.1371/journal.pone.0019168

Yumoto, N., Lu, X., Henry, T. R., Miyachi, S., Nambu, A., Fukai, T., & Takada, M. (2011). A neural correlate of the processing of multi-second time intervals in primate prefrontal cortex. *PLoS ONE*, *6*(4), e19168. https://doi.org/10.1371/journal.pone.0019168

Zelaznik, H. N., Spencer, R. M., & Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(3), 575-588. https://doi.org/10.1037/0096-1523.28.3.575

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

Zwickel, J., Grosjean, M., & Prinz, W. (2008). A contrast effect between the concurrent production and perception of movement directions. *Visual Cognition*, *16*(7), 953-978. https://doi.org/10.1080/13506280701653586