EFFECT OF SPEED MANIPULATIONS ON PHONOLOGICAL SHORT-TERM MEMORY

EFFECT OF SPEED MANIPULATIONS ON PHONOLOGICAL SHORT-TERM MEMORY

By JORDAN ZIQI YANG, B.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2023)

Hamilton, Ontario (Department of Linguistics and Languages)

TITLE: Effect of Speed Manipulations on Phonological Short-Term Memory

AUTHOR: Jordan Ziqi Yang, B.A. (McMaster University)

SUPERVISOR: Dr. Elisabet Service; Dr. Daniel Pape

NUMBER OF PAGES: x, 71

Lay Abstract

Our ability to remember spoken language, or phonological content, is closely tied to how we perceive its rhythm and timing. When we hear speech, our brain processes not just the words but also the temporal structure of the context. This temporal organization helps to retain spoken information, and our ability to remember speech in turn predicts how well we can learn new words in a foreign language. This suggests that the rhythmic patterns of words and their sentence contexts might play a crucial role in how we organize and remember linguistic content. The brain's rhythms naturally synchronize with the rhythms of speech, influencing comprehension. However, when this synchronization is disrupted – for instance, when a speaker changes their pace suddenly – comprehension may be impeded. On the other hand, while studies have shown that memory for digits is not affected by their presentations at different rhythmic regularities, the way the brain processes disruption of an internally driven rhythms might differ from how it processes externally driven rhythms. This thesis explores these rhythmic influences on memory by conducting tasks involving repeating sentences in native English and the foreign Urdu language to see how sudden changes in rhythm impact memory. This sheds light on the mechanisms by which the brain handles different time-related aspects of language and how this affects the memory retention, which ultimately shapes language skills and learning abilities.

Abstract

Temporal representation in the brain has been recently acknowledged as a fundamental mechanism underlying short-term memory (STM). Nonetheless, the existing body of research presents conflicting results on the extent of this relationship. Some studies propose that rhythmic disruption adversely affects perception and comprehension, and regular rhythm has been associated with the facilitative impact on STM tasks, while others suggest that its influence on STM tasks might not be as substantial. This thesis delves into rhythmic irregularity's impact on STM, particularly when an internally established rhythm is disrupted. Through two verbal STM tasks – the jabberwocky memory task and the Urdu memory task – conducted with thirty participants (31 females), involving the repetition of sentences in native English or foreign Urdu, this experiment investigates whether manipulating the speed of target sentences in relation to prime phrases affects the accuracy of sentence repetition in the assigned tasks. Our hypothesis posited that STM for sentences presented at altered speeds, either slower or faster, would be compromised compared to sentences at a normal pace. However, the outcomes of our study did not reveal any significant differences in repetition accuracy across the three speed conditions within the two STM tasks. Interestingly, our investigation uncovers two noteworthy findings. Firstly, variations in repetition accuracy among the three speed conditions appear to be influenced by participants' bilingual or multilingual background. Bilingual and multilingual individuals exhibited better performance under slow speed conditions at the syllable level, while monolingual participants displayed enhanced recall accuracy for whole words in the normal speed condition. Secondly, a discrepancy emerges between participants' self-perceived performance across the three speech conditions and their actual performance. These findings emphasize the potential roles of linguistic background and metacognition in shaping both temporal representation and STM performance, thereby prompting further exploration of these intricate interactions.

Acknowledgements

I would first like to thank my supervisor, Dr. Elisabet Service, whose guidance and expertise have been invaluable throughout my graduate study and the writing of this thesis. I extend my thanks to my co-supervisor, Dr. Daniel Pape, for the guidance on the technical aspects of the study. I am truly grateful for the wealth of knowledge I have gained under their mentorship. I would like to express my sincere appreciation to the colleagues of the Memory, Language, and Brain Lab at ARiEAL: Fiza Ahmed and Fareeha Rana who have been providing me with invaluable resources and assistance throughout the project. A special mention goes to the volunteer research assistants at ARiEAL - Dana Alnahhar, Jacquelyn Campbell, Kaelan Crago, Mahnoor Butt, Paloma van Vliet, Shruthi Raghuraman, and Zainab Khan. Their dedication and hard work in assisting with data scoring have been immensely impactful. I am especially grateful to Conan Hu for the unwavering emotional support throughout the challenging yet exciting journey of completing this project. Last but not least, I want to thank my parents and friends for their constant love and support, which has been a driving force behind my academic journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANOVA = analysis of variance

DAT = Dynamic Attending Theory

EEG = electroencephalograph

FWL = foreign word learning

IRR = inter-rater reliability

LM = long-term memory

MEG = magnetoencephalography

MREB = McMaster Research Ethics Board

PD = Parkinson's Disease

STM = short-term memory

SVO = subject-verb-object

WM = working memory

Declaration of Academic Achievement

The conceptualization of this thesis was conducted by myself, in collaboration with Dr. Elisabet Service and Dr. Daniel Pape. I explored and analyzed the data under the supervision of Dr. Elisabet Service. The statistical analysis was completed by myself and Dr. Elisabet Service. The present thesis was written by myself, with comments and feedback from Dr. Elisabet Service, Dr. Daniel Pape, and Dr. Gemma Repiso Puigdelliura.

1 Introduction

Consider the lyrics from a rhythmic rap song, the same lyrics in normal speech, and in rhythmically irregular speech – which version would you be more likely to remember? Our ability to retain phonological content in short-term memory (STM) has been linked to the way we perceive its temporal structure. Research has revealed that the capacity to represent temporal patterns in STM supports our ability to remember spoken content for immediate repetition. Interestingly, individual differences in phonological STM predict how well we learn foreign language word forms. This observation has led to the hypothesis that the prosodic characteristics of words might aid in organizing sounds in our STM (Tierney, White-Schwoch, MacLean, & Kraus, 2017; Service, DeBorba, Lopez-Cormier, Horzum, & Pape, 2022). Neurophysiological studies investigating brain activity offer evidence that our brain rhythms synchronize with speech rhythms, and this synchronization can impact speech comprehension (Teki & Griffiths, 2016; Papagno, Comi, Riva, Bizzi, Vernice, Casarotti & al., 2017). When an established internal rhythm is disrupted, such as when a speaker alters their speech rate, comprehension can be negatively affected (Kösem, Bosker, Takashima, Meyer, Jensen, & Hagoort, 2018). Yet, studies directly comparing regular and irregular intervals between digits in STM tasks have not identified a rhythmic influence on phonological STM (Gorin, 2020; Lui & Wöstmann, 2022). However, there might be a difference between the effects of internally and externally driven rhythms, suggesting that only internal rhythms contribute to the organization of order in STM. The specific conditions of the studies by Gorin (2020) and Lui & Wöstmann (2022) may not have impacted the internal rhythms tied to language comprehension and phonological STM, particularly at the syllable and phrase levels.

This current study consists of two STM repetition tasks - an English pseudo-sentence task and an Urdu sentence task. The experiment aims to explore the consequences of rhythmic disruptions on verbal STM repetition tasks. It delves into the fundamental processes involved in processing absolute and relative time, and their role in shaping the representation of temporal order in STM. Additionally, it investigates whether these processes either enhance or limit the representation of phonological information, which could ultimately influence the efficiency of language functions and developmental processes.

1.1 Phonological Short-Term Memory as a Subsystem of Working Memory

The concept of working memory (WM) was initially elaborated by Baddeley and Hitch (Baddeley & Hitch, 1974). WM is defined as "a temporary storage of information in connection with the execution of other cognitive tasks such as reading, problem-solving, or learning" (Baddeley, 1983). The novel conceptualization of WM aimed to address the limitations of the traditional model of information flow from short-term memory (STM) to long-term memory (LTM) proposed by Atkinson and Shiffrin (Atkinson & Shiffrin, 1968). According to their model, information is first held in the limited-capacity STM store and then directly transferred to LTM for later retrieval and various cognitive functions. However, this two-layered model has its shortcomings. One issue with this model is that it implies that maintaining information in STM automatically leads to its storage in LTM, and conversely, disruption or deficits in STM result in an individual's inability to learn due to unsuccessful encoding into LTM (Baddeley, 2010). Nevertheless, research has shown that maintaining STM for the phonetic aspects of spoken language utterances is less conducive to learning compared to encoding the semantics and emotional valence associated with the spoken language forms (Craik & Lockhart, 1972). Furthermore, case studies have demonstrated that deficits in STM capacity have less impact on LTM encoding than initially suggested by the direct pathway to LTM (Baddeley, 2010).

Due to the limitations of the traditional memory models, a newer and widely recognized framework, known as the WM model, was proposed by Baddeley and Hitch. This model introduces a multi-layered component structure, with the central WM component being a limited-capacity attentional controller, known as the central executive. The central executive is responsible for overseeing various STM subsystems, including the visuo-spatial sketchpad and the phonological loop. The visuo-spatial sketchpad handles STM for visual information, while the phonological loop is responsible for storing STM related to acoustic material (Baddeley, 1983). A third subsystem of STM, known as the episodic buffer, was later introduced. The episodic buffer is domain-general and capable of storing approximately four chunks or episodes of events that can be accessed through conscious awareness (Baddeley, 2007). Unlike the previous single STM system, this model incorporates at least three distinct, yet interconnected, subsystems, allowing for parallel processing across these subsystems. This multi-component structure enables a more intricate and detailed modeling of the individual components within it. The idea of a WM system governing STM subsystems emphasizes that WM's function surpasses mere storage, highlighting its crucial involvement in various cognitive processes (Baddeley, 2010).

While this thesis predominantly focuses on exploring the mechanisms underlying phonological (speech-based) STM, it's essential to contextualize phonological STM within the broader framework of WM, which is the framework assumed by this current thesis. Within this framework, phonological STM is viewed as a subsystem operating in conjunction with other subsystems, such as the visuo-spatial sketchpad and the episodic buffer. These subsystems are regulated by the central executive, which functions as an attention-controlling system. Throughout the remainder of this thesis, we will use the term "short-term memory" (STM) to refer to what is often termed "phonological short-term memory."

1.2 Encoding Temporal and Sequential Information in STM

An unaddressed aspect of STM in Baddeley and Hitch's WM model pertains to the mechanism for maintaining the temporal order within STM. While visual and phonological content unfolds over time, it's essential to encode the temporal and sequential information of this material to maintain an accurate mental representation. The processes involved in encoding such temporal and sequential details in STM continue to be a subject of ongoing research. Several theoretical models have been suggested to explain the time and order tracking mechanism in STM. Burgess and Hitch proposed a network model (Burgess & Hitch, 1992; 1999; 2006). This model assumes two separate mechanisms for encoding order information, and phonological and semantic (item) information. Order information is coded via connections between lexical representations of items and a context-timing signal, whereas item information is coded in terms of associations between lexical and phonological representations (Burgess & Hitch, 2006). Brown, Preece, and Hulme (2000) suggested that the representation of sequential information relies on the alignment between an individual's internal dynamic sequences (referred to as an internal clock) and the sequences of external events. The time signals generated through these dynamic correspondence mechanisms enable the encoding of sequential information. Collectively, various theoretical models all emphasize the necessity of order-tracking in STM (Burgess & Hitch, 2006; Brown & al., 2000; Lewandowsky & Farrell, 2008; Hartley, Hurlstone, & Hitch, 2016), with some suggesting that such order signal relies on a time-keeping mechanism in STM.

Neurocognitive research has shed light on the neural networks responsible for processing temporal information within the brain. Lesion and stimulation studies have provided insights into the crucial roles of parietal and temporal areas in encoding order information. Papagno et al. (2017) conducted digit span assessments on patients undergoing awake surgery for left frontal or temporal glioma removal. Their findings indicated that both item and order storage involve various brain areas, including Broca's area, the inferior parietal lobule, and the supramarginal gyrus. Additionally, stimulation of the anterior arcuate fascicle pathway between inferior frontal areas and temporal areas led to interference in order accuracy (Papagno & al., 2017). An intriguing additional pathway for STM in the phonological loop, particularly for order information, has been proposed by Papagno and her colleagues. This model suggests that auditory information is transferred from the left posterior temporal (Wernicke's) area to the left inferior frontal (Broca's) area. Simultaneously, order information is proposed to be conveyed from the supramarginal gyrus area to Broca's area. Subsequently, the information from both pathways is thought to be integrated to facilitate articulation (Papagno & al., 2017).

Teki and Griffiths (2014) introduced a paradigm to investigate STM for temporal structure. Through a series of experiments, participants were tasked with retaining interval lengths in sequences featuring different levels of temporal regularity. The combined results consistently indicated that the retention of memory for single intervals declines with decreasing temporal regularity in the sequences (Teki & Griffiths, 2014). In subsequent fMRI studies focusing on brain substrates of these STM tasks, the findings revealed a significant negative correlation between the temporal regularity of the stimuli and brain activation in regions such as the cerebellum and striatum. These brain areas exhibited greater activation when processing stimuli with temporal irregularities (Teki & Griffiths, 2016). This observation supports the previous notion of an interconnected relationship between the cerebellum and striatum in relation to time perception. The striatum is suggested to function as an internal clock, while the cerebellum is involved in encoding timing errors corresponding to the striatal clock, and that both types of timing processing may be central to verbal STM (Teki et al., 2012). In this framework, the cerebellum is involved in encoding absolute, duration-based timing, as well as errors compared to the striatal clock, which is involved in encoding relative, rhythm-based timing. The cerebellum is activated when the primary striatal clock deviates from absolute time. This creates a two-layer representation with the striatum encoding the shape of temporal patterns and the cerebellum their exact temporal realization as faster or slower (Teki & Griffiths, 2016). A review conducted by Paton and Buonomono (2018) further suggests that the underlying mechanisms involved in detecting temporal information are rooted in the intrinsic cellular and network dynamics of neural circuits. Through the activation of oscillatory frequency bands, a temporal framework is established, enabling the brain to capture and process temporal information effectively (Paton & Buonomano, 2018). Collectively, the research indicates a plausible mechanism for encoding both time and order information within STM. This involves the activation of "clocks" in the cerebellum and striatal areas by the frontal regions to monitor time information. Subsequently, this information is stored in the inferior parietal and temporal regions before being transferred back to the frontal areas for articulation.

1.3 Rhythmic Ability as a Fundamental Cognitive Process

Several behavioral findings provide support for the interplay between the ability to encode temporal information and both STM and perception capabilities. In a study conducted by Grahn and Brett (2009), the rhythmic discrimination performance of individuals with Parkinson's Disease (PD) was compared to that of healthy controls. Participants were presented with two types of rhythmic stimuli: metric complex and metric simple rhythms. The task was to determine if a target rhythmic stimulus matched a previously heard rhythmic stimulus. Metric simple (beat-based) stimuli were regularly arranged into groups of four units, while metric complex (not-beat-based) stimuli lacked such regular grouping (Grahn & Brett, 2007). The results indicated a significant improvement in performance for rhythms with a regular beat in the control group, as opposed to rhythms without a beat. However, this improvement was less pronounced in the PD group. Notably, the discrimination of beat-based rhythms was impaired in PD patients in comparison to the control group, while the discrimination of non-beat-based rhythms showed no significant differences (Grahn & Brett, 2009). In line with this work, a subsequent study identified a benefit for rhythm discrimination tasks involving both beat-based visual and auditory stimuli (Grahn & Brett, 2012). This observation implies that rhythmic and temporal capabilities could be a cross-modal fundamental cognitive process that underlies elementary cognitive functions, including perception.

In two experiments examining participants' STM for digit span presented with specific temporal structures (Gilbert, Hitch, & Hartley, 2017), it was discovered that increasing domain-specific load on STM led to a decrease in the temporal precision of digit span rehearsal. Temporal precision is defined as the degree of alignment between the time-points at which a digit is expected to be rehearsed and the time-point at which it is actually rehearsed. Moreover, temporal precision emerged as a significant predictor of recall accuracy, even after accounting for factors such as nonverbal IQ and visuo-spatial STM. Investigations into dyslexic individuals (Laasonen, Virsu, Oinonen, Sandbacka, Salakari, & Service, 2012) revealed overlap between verbal memory and memory for non-verbal temporal sequences. The research uncovered a significant correlation between serial memory for non-verbal stimuli (light flashes, tone bursts, finger touches) and phonological STM for verbal stimuli. The correlation was found for both dyslexic and typical readers although the dyslexic participants performed less well in both types of STM tasks. Consistent with these findings, Tierney et al. (2017) found that individuals with enhanced rhythmic sequence memory, defined as the capability to preserve and replicate a provided rhythmic sequence, displayed better performance not only in verbal STM measures but also in other language-related tasks, including reading (Tierney & al., 2017). This observation raises the possibility that the challenges associated with encoding sequence information in STM might be a contributing element to deficits in various linguistic aptitudes within the dyslexic population.

A recent study (Service & al., 2022) further supports the association between rhythmic ability and verbal STM. Participants were tasked with replicating MORSE-like patterns consisting of short and long beeps through tapping. Subsequently, they engaged in verbal STM tasks involving English-like pseudo-sentences (jabberwocky sentences) and real Turkish sentences. Additionally, participants completed a foreign-word learning task where they heard English-word–Turkish-word pairs and were required to recall the associated Turkish word upon hearing the English term. Significant positive correlations emerged between tapping accuracy and STM performance for both English jabberwocky sentences and Turkish sentences, as well as between tapping accuracy and word learning. These findings highlight a significant research area concerning the role of temporal representation and its potential influence on fundamental cognitive processes such as perception, extending to STM and language learning.

1.4 Effects of Rhythmic Manipulations on STM and Perception

Few studies have directly investigated how the rhythmic properties of auditory content affect one's STM for and perception of such content. Kösem et al. (2018) conducted a study to explore whether an individual's perception of target phrases is influenced by manipulations of the preceding speech rate of carrier (priming) phrases, along with the corresponding patterns of neural entrainment in response to these speech rates. In this study, sentences presented in the carrier window were at either a slow rate (3 Hz) or a fast rate (5.5 Hz). The final three words in the target window were presented at the original speech rate. Notably, the final word contained an ambiguous vowel that could be perceived as either a short /ɑ/ or a long /a:/, leading to two distinct Dutch words. Results demonstrated that participants were more inclined to perceive the final word with a long vowel following a fast speech rate, and conversely, they were more likely to perceive the final word with a short vowel following a slow speech rate. Additionally, measurements using magnetoencephalography (MEG) revealed that neural oscillations in the auditory area mirrored the speech rate of sentences presented in the carrier window. Specifically, there was a stronger 3 Hz oscillation power during the slow speech rate condition and a stronger 5.5 Hz oscillation power during the fast speech rate condition. Remarkably, this oscillation pattern persisted for several cycles after the carrier window, suggesting a lingering effect on neural activity. These findings provide support for the notion that the rhythmic properties of speech influence neural oscillations, which in turn have been found to play a role in shaping speech perception and comprehension (Kösem & al., 2018).

A similar result was found in an electroencephalograph (EEG) study that examined how the rhythmical structure of a sentence context affects speech comprehension (Li, Shao, Xia, & Xu, 2019). Participants listened to Mandarin Chinese sentences with regular or irregular rhythms, including critical nouns that were either semantically congruent or incongruent with the preceding context. Findings showed that incongruent nouns in the regular-rhythm conditions triggered a larger N1 and N400 response, while in irregular-rhythm conditions, they led to a larger P600 response. This early onset of neural responses to semantic incongruence in the regular-rhythm condition suggests that rhythmic regularity could accelerate speech comprehension (Li & al., 2019).

Although a number of studies have predicted and found regular rhythms to support verbal STM, this has not always been the case. In a recent investigation by Gorin (2020), the direct influence of rhythmic attributes on STM was explored. The study involved comparing participants' accuracy in recalling sequences of nine digits presented with varying rhythmic structures. The total time of digit presentation remained consistent across all conditions. In the regular condition, digits were presented with a steady rhythm, while in the irregular trials, the time intervals between auditory digit presentation and the duration of the digits themselves differed. Contrary to what was hypothesized, the results indicated no significant difference in recall accuracy between the regular and irregular conditions, and no discernible distinctions in error patterns were observed (Gorin, 2020). Similarly, another study (Lui & Wöstmann, 2022) examined the impact of rhythmically regular versus irregular distractors during the memory encoding phase on participants' accuracy in digit span tasks. The findings revealed no influence on the participants' primary performance measure, which was recall in the digit span task. However, a secondary performance metric, reaction time, did show variation across the conditions. The participants displayed faster responses when faced with regular distractors compared to irregular ones. Additionally, participants' confidence ratings for their responses indicated a trend of higher confidence when regular distractors were presented during the memory encoding phase in comparison to irregular distractors were introduced (Lui & Wöstmann, 2022).

The impact of rhythmic structure on STM remains a topic of ongoing discussion. Some evidence indicates differences in neural oscillations when processing regular versus irregular speech, as well as the potential influence of rhythmic irregularity on comprehension or comprehension efficiency (Kösem & al., 2018; Li & al., 2019). However, behavioral studies have yielded inconsistent results. Specifically, two studies exploring STM performance with materials presented in regular versus irregular patterns have failed to identify significant performance differences (Gorin, 2020; Lui & Wöstmann, 2022). However, one of these studies (Lui & Wöstmann, 2022) have highlighted effects in two secondary performance measures: reaction time and response confidence.

1.5 Purpose

The lack of consistent findings on the effects of rhythmic structures on STM, highlighted in section 1.4, could originate from three underlying factors. Firstly, the behavioral studies conducted thus far have not distinctly differentiated between internally and externally driven rhythms. As outlined in section 1.2, and supported by neuroimaging research, there is a general consensus that the brain encodes time information through the activation of phase-locking neural oscillations. These oscillations are presumed to be generated by a group of neurons firing synchronously in response to both external physical stimuli and internal cognitive processes (Lakatos, Shah, Knuth, Ulbert, Karmos, & Schroeder, 2005; Giraud & Poeppel, 2012; Paton & Buonomano, 2018; Obleser & Kayser, 2019). The continuous interaction and alignment between internally and externally driven rhythms enables the representation of time. The two STM studies (Gorin, 2020; Lui & Wöstmann, 2022) discussed in the previous section presented participants with external auditory visual stimuli with various temporal regularities, yet the concept of an internal rhythm was absent. Given the brain's capacity to process a wide range of unpredictable auditory stimuli often considered as non-beat-based, such as music and speech, and considering that most of the cues encountered in daily life are not beat-based, the notion of a default internal 'clock' or entrainment rhythm oscillating at a specific rate governing temporal representation is questionable. In Kösem et al. (2018), participants were exposed to a priming phase characterized by a consistent rhythm. They were subsequently assessed on their perception of a target phase, with rhythms that were either consistent or inconsistent with the preceding priming phase. This experimental design facilitates the establishment of an internal rhythm that participants can synchronize with. It suggests the potential for a disruption in fundamental cognitive processes, such as perception, when the externally presented rhythm mismatches with an internally established rhythm.

Secondly, speech is arrhythmic, thus exogenous entrainment during speech processing is unlikely to solely result from acoustic cues. Representations of morpheme, lexical, syntactic, and semantic structures, as well as discourse, and pragmatic inferences cannot be regarded as outcomes driven by sensory inputs from speech, rather, they originate from internally generated predictions (Meyer, Sun, & Martin, 2020). Consequently, the rhythmic predictions formed for actual sentences (as in Kösem & al., 2018; Li & al., 2019) might contrast with those made for lists of digits (as in Gorin, 2020; Lui & Wöstmann, 2022), which could predominantly adhere to a beat-based structure. This distinction could potentially interact with the impact of diverse temporal regularities on cognitive processes in varying ways.

Lastly, the nature of the tasks can influence outcomes due to varying levels of attentional allocation. Comprehension accuracy, achieved through passive perception, might not mirror STM accuracy, which demands active retention, and, in the case of verbal STM repetition, articulation. The allocation of attentional resources to an STM task is higher than that to a perception task. Furthermore, not only do attention levels differ, but metacognition might also vary. It's been proposed that metacognitive monitoring is involved in strategic regulation during memory recall, driven by the assessment and threshold-setting in accordance with task demands (Flavell, 1979; Koriat & Goldsmith, 1996; Goldsmith & Koriat, 2007). This aligns with the concept of STM functioning as a subsystem overseen by the attention-controlling central executive.

The present study aims to explore the role of temporal representation in phonological STM and language aptitude. Specifically, it investigates the impact of rhythmic disruptions on STM repetition tasks and whether the ability to overcome such disruptions correlates with language learning capabilities. The study will assess STM performance using sentences presented at a consistent speech rate compared to sentences with a mid-way change in speech rate. This design allows participants to establish an internal rhythm, which is then disrupted. The study builds upon prior research indicating that disrupting expected speech rhythms hampers speech perception and comprehension (Kösem & al., 2018; Li & al., 2019). We anticipate that STM will also be adversely affected when an internally expected rhythm is disrupted. Furthermore, metacognitive evaluations, such as perceived difficulty in recalling sentences presented with different rhythmic regularities, will be assessed.

1.5.1 Short-Term Memory Tasks

Nonword repetition has been widely employed as a measure of individual phonological STM for phonetically and morphologically plausible words in a given language (Baddeley, Gathercole, & Papagno, 1998; Service, 1992; Service & Craik, 1993). The performance on nonword repetition STM tasks has consistently been shown to predict novel word learning in both native and foreign language acquisition, as well as other language skills like reading and vocabulary knowledge (Gathercole, Willis, Emslie, & Baddeley, 1992; Service & Kohonen, 1992). Conversely, deficits in this ability have been linked to individuals with language disorders (Gathercole & Baddeley, 1990). In this study, we will use pseudo-sentences composed of nonwords (jabberwocky sentences) to evaluate individuals' verbal STM. The rationale behind employing jabberwocky sentences lies in their greater resemblance to linguistic utterances, as they adhere to English syntactical and prosodic structures more closely than individual words. Consequently, compared to isolated words, they are expected to better represent the type of linguistic material encoded into verbal STM in real-word contexts. Further elaboration on the formulation of jabberwocky sentences is provided in Section 2.2.1.

In addition to the English-based jabberwocky memory task, this study incorporates an Urdu memory task aimed at evaluating participants' STM for sentences in a real foreign language that differs significantly from their native English in terms of phonetic, syntactic, and prosodic structures. Because of its dissimilarity to English, the Urdu memory task was anticipated to be more challenging. However, it could offer a more precise reflection of an individual's capacity to acquire a new language. Further elaboration on the formulation of Urdu sentences is provided in Section 2.2.2.

2 Methods

2.1 Participants

A total of forty-two participants were recruited for the study, all of whom were selected from the McMaster University Department of Psychology, Neuroscience & Behaviour (PNB) Research Participation System. Participants received compensation in the form of either one SONA course credit or fifteen dollars cash reimbursement for their participation. Prior to the beginning of the experimental session, each participant completed a written letter of information and a consent form. After completing all components of the experiment, a debriefing letter explaining the research's motives was distributed to each participant. Participants were then asked to provide written re-consent if they wished to remain a part of the study and have their data stored. This study received clearance from the McMaster Research Ethics Board (MREB).

One participant withdrew from the study, and one participant's data was excluded due to recording issues during the data collection process. The remaining sample of 40 participants consisted of 31 females and 9 males, ranging in age from 18 to 24 years (M_{age} = 18.7 years, SD = 1.34 years). All participants were students at McMaster University pursuing undergraduate (n = 41) or graduate (n = 1) education. They all reported being predominantly English speakers. The participants came from diverse racial, ethnic, and language backgrounds (see Table A1-3 in Appendix A).

2.2 Stimuli and Tasks

All auditory stimuli were recorded using Audacity software in a sound-proof recording studio with a Sennheiser ME62 wired electret condenser microphone. The microphone was connected to a Focusrite Scarlett Solo audio interface, which was in turn connected to a laptop without fan (to avoid acoustic interference with the recordings). To avoid plosive sounds, the microphone was placed at approximately 10 cm from the speaker's mouth in a slightly off-center manner (45 degrees). Auditory stimuli were then spliced and trimmed using Audacity. For speech rate and loudness manipulations, Praat (version 6.2.22; Boersma & Weenink, 2009) was employed. For speech rate manipulation, the command is as the following: 1) select 'Manipulation editor'; 2) enter the values (as rates) for the desired percentage in the "Duration manipulation" tab of the manipulation editor. The algorithm used was the Praat-intern PSOLA. For going into the manipulation editor, Praat's standard settings were used. For loudness normalization, the command is as the following: 1) select "Scale to intensity" when selecting multiple sound files; 2) entering the values for the desired loudness (as decibels). Values for speech rate and loudness manipulations are detailed in section 2.2.3.

2.2.1 Jabberwocky Memory Task

A repetition task with meaningless "jabberwocky" sentences was presented to assess phonological STM. Thirty jabberwocky sentences were utilized for the jabberwocky memory task (see List B1 in Appendix B). Out of these thirty sentences, ten were adapted from the jabberwocky sentences used in Ahmed (2022), and an additional twenty sentences were specifically composed for this experiment. The jabberwocky sentences adhered to English phonological and morphological rules. The sentences were structured in a syntactic order of Determiner – Subject – Verb – (Preposition) – Determiner – Object but all words were replaced by pseudowords. A list of jabberwocky function words and their corresponding English counterparts is available in Appendix B (List B2). The length of the jabberwocky sentences ranged from 6 to 8 syllables, and they consisted of 5 to 6 pseudowords each.

Examples of jabberwocky sentences:

ROO MUTHTIN FANED IM GRASHNIT

WOE VOTION PRUSED JISH ROO ENTSIAN

For each jabberwocky sentence, there was a paired English introductory phrase. A total of thirty introductory phrases were composed, with a length of 6 syllables and containing 3 to 6 words each.

Examples of introductory phrase + jabberwocky sentence pairs:

The new rumour is that - ROO MUTHTIN FANED IM GRASHNIT

Yesterday Franklin said - WOE VOTION PRUSED JISH ROO ENTSIAN

All thirty introductory phrases + jabberwocky sentences were recorded by a native speaker of English. To maintain natural speech patterns, the speaker was instructed to utter the introductory phrases and jabberwocky sentences together at their normal speech rate. The introductory phrases had durations ranging from 1024 to 1412 ms (M = 1224.8; SD = 93.3) with a speech rate of 4.25 to 5.86 syllables per second (M = 4.90; SD = 0.382). The jabberwocky sentences had durations ranging from 1677 to 2136 ms (M = 1935.5; SD = 126) with a speech rate of 3.28 to 4.56 syllables per second (M = 3.75; SD = 0.271). The speaker, being less familiar with the nonsense words in the jabberwocky sentences, pronounced them at a slightly slower pace compared to the English introductory phrases. To match the speech rate of both components more closely, the introductory phrases were expanded to 120% of their original speeds using Praat software. Consequently, the new introductory phrases had durations ranging from 1213.3 to 1694.4 ms (M = 1466; SD = 119) with a speech rate of 3.54 to 4.95 syllables per second (M = 4.12; SD = 0.345).

2.2.2 Urdu Memory Task

A second sentence repetition task used material in Urdu. Thirty Urdu sentences were composed for the Urdu memory task (see List B3 in Appendix B). The Urdu sentences varied in length from 5 to 8 syllables and consisted of 5 to 6 words each. Each Urdu sentence was paired with an Urdu introductory phrase. For the task, ten introductory phrases were created, and each introductory phrase was used with three target sentences. The introductory phrases had 6 syllables and contained 3 to 6 words, consistently ending with the word /ke/. While all Urdu introductory phrases and target sentences were constructed to be syntactically grammatical, it is worth noting that some sentences might be perceived as semantically unusual by native speakers.

Examples of introductory phrase + target Urdu sentences:

Hum keh ra.he the KE - NOOR NE BAR.TAN DHO LI.YE

We were saying that – Noor has washed the dishes

Kal ais.e hu.a KE - ZAID DU.KAAN PAR GE.YA

Yesterday it was that - Zaid went to the store

All thirty introductory phrases + target Urdu sentences were recorded by a native Urdu speaker. To maintain the natural flow of the sentences, the speaker was asked to utter the introductory phrases and target sentences together at their normal speech rate. The introductory phrases ranged from 877 to 1454 ms in length (M = 1278; SD = 104) with a speech rate of 4.13 to 6.84 syllables per second (M = 4.73; SD = 0.472). The target sentences ranged from 1006 to 1745 ms in length (M = 1439.4; SD = 156) with a speech rate of 3.51 to 5.64 syllables per second (M = 4.52; SD = 0.44).

2.2.3 Speed Manipulations

The main research question concerned speech rate effects on phonological STM as reflected in sentence repetition tasks. A speed manipulation approach was adopted from Kösem et al. (2018). Following the initial recording, all stimuli were uniformly expanded to 102% of their original speed. This step ensured that all parts of the stimuli received equal synthetic treatment, and thus any consistent artifacts from the speed change treatment would be overlaid onto all stimuli in this experiment. Subsequently, the three speed conditions (normal, slow, fast) for the jabberwocky and target Urdu sentences were created. For this purpose, the stimuli were transformed to 133% and 75% (1/1.33) of their previous speed (after the 102% expansion). Additionally, to maintain consistency in presentation, all stimuli were loudness normalized to 70dB in a final step. As a result, the final set of stimuli included 90 introductory phrases + jabberwocky sentences, each presented in three speed conditions, and 90 introductory phrase + target Urdu sentences, also presented in three speed conditions. Note that the speech rate of the introductory phrases remained constant throughout the task, with only the jabberwocky and target Urdu sentences being presented at three different speech rates. Additionally, it is important to highlight that the reason for keeping the speech rate of the introductory phrases constant throughout the task, while presenting the jabberwocky and target Urdu sentences in three different speed conditions, was to create a clear distinction between rhythmic disruption conditions (fast, slow) and the undisrupted condition (normal). This design choice ensured that any observed effects on memory and perception could be attributed specifically to the manipulation of speech rate in the jabberwocky and target Urdu sentences, while the rate of the introductory phrases remained consistent and served as a stable point of reference for participants. By doing so, we aimed to explore the potential impact of speech rate variations on memory and perception, while keeping other linguistic and acoustic factors controlled across the three speed conditions.

2.3 Experimental Procedure

The experiment comprised three tasks: 1) jabberwocky memory task; 2) Urdu memory task; 3) Foreign Word Learning (FWL) task (not reported here), along with an online demographic and language background questionnaire. The experiment took place in a quiet office space. The tasks were programmed using Psychopy and presented on an Apple iMac computer with a 21.5-inch screen and a resolution of 1920x1080 pixels. Participants listened to the auditory stimuli using a Sony MDRZX110NC over-ear headphones, with wires connected to the computer. The noise canceling feature of the headphones was not activated to ensure that it did not interfere with the auditory stimuli. The volume of the headphones was adjusted to 50% for all participants to maintain a consistent listening experience. Audio recordings were made using a Rode NT-USB microphone and captured using Audacity software, no additional acoustic treatment was performed on the audito recordings. Participants were familiarized with the task procedures through both verbal instructions from the experimenter and written instructions displayed on the screen. Each task was preceded by practice trials, allowing participants to become familiarized with the task procedures and stimulus presentation. The sequence of tasks was as follows: participants first completed the jabberwocky memory task, followed by the Urdu memory task, and finally, the FWL task. After completing all three tasks, participants filled out the demographic and language background questionnaire using Limesurvey, an online survey platform. The experimental session took place in a controlled and quiet environment to minimize potential distractions and ensure data consistency across participants. Instructions and stimulus presentation were standardized to maintain uniformity throughout the experiment. Participants were encouraged to ask questions if they needed clarification during the experimental session.

2.3.1 Jabberwocky Memory Task

The jabberwocky memory task aimed to assess participants' phonological STM capacity for language-like stimuli, without the aid of any semantic references. By using jabberwocky sentences composed of invented words, the task minimized meaning associations and focused solely on the phonological aspects of STM. The inclusion of different speed conditions allowed for the examination of potential effects of speech rate and rhythmic disruption on memory performance during the task.

The task involved the use of thirty jabberwocky sentences, which are sentences composed of nonsensical or invented words, resulting in sentences that may sound meaningful but do not convey any coherent idea. Such sentences were presented at normal, slow, or fast speeds, resulting in a total of 90 trials (30 sentences presented in each speed condition). Each jabberwocky sentence was always preceded by an English introductory phrase, and the speech rate of the English introductory phrases remained constant throughout the task so that the difference in speed between the introductory phrases and target sentences had three values: very small difference for the normal sentences, clear slow-down for the slow sentences, and clear speed-up for the fast sentences. In the jabberwocky memory task, participants were instructed to focus solely on recalling the jabberwocky portion of each sentence and disregard the introductory phrase. Each trial began with the appearance of the word 'LISTEN' on the screen, prompting participants to listen to the auditory stimuli. After each sentence, the word 'REPEAT' appeared, signaling participants to repeat aloud the jabberwocky sentence they had just heard. In case a participant was unable to recall any part of a sentence, they were instructed to respond with the word 'BLANK'. To familiarize participants with the task, they were presented 15 practice trials. These were followed by 75 experimental trials. The scoring of participants' responses was carried out by four native English speakers, with three of the scorers having conversational fluency in French, and one of them fluent in Telugu. Scorers assessed the phonological accuracy of the responses at the consonant, syllable, and word levels. Participants were awarded one point for each correctly recalled consonant, syllable, and word, and no points were given for incorrect responses.

2.3.2 Urdu Memory Task

In the same way as the jabberwocky memory task, the Urdu memory task served to assess participants' STM, but in a foreign language context, where both the phonological and syntactic structures of the sentences differed from their native language. As with the jabberwocky memory task, the sentences were presented at normal, slow, or fast speeds, allowing for the examination of the impact changes in speech rate and rhythmic disruption on memory performance.

The Urdu memory task consisted of thirty Urdu target sentences, presented at normal, slow, or fast speeds, resulting in a total of 90 trials (30 sentences in each speed condition). Each Urdu target sentence was always preceded by an Urdu introductory phrase ending in the sound /ke/, and the speech rate of the introductory phrases remained constant throughout the task. During the Urdu Memory Task, participants were specifically instructed to recall only the portion of the sentence that came after the /ke/ sound and disregard the portion that preceded it. Each trial began with the appearance of the word 'LISTEN' on the screen, prompting participants to listen to the auditory stimuli. After each sentence, the word 'REPEAT' appeared, signaling participants to orally repeat the target sentence they had just heard. In the event that a participant was unable to recall any part of a sentence, they were instructed to respond with the word 'BLANK'. To familiarize participants with the task, they first completed 15 practice trials. These were followed by 75 experimental trials. The scoring of participants' responses was conducted by three native Urdu speakers, all scorers are also native English speakers, and have conversational fluency in Hindi and Punjabi. Scorers assessed the phonological accuracy of the responses at the consonant, syllable, and word levels. Participants received one point for each correctly recalled consonant, syllable, and word, and no points were awarded for incorrect responses.

In both the jabberwocky and Urdu memory tasks, the order of sentence presentation was pseudo-randomized. This means that sentences were presented in random orders, with the same sentence never appearing at different speeds less than two trials apart to avoid potential order effects. The presentation order was also counterbalanced to minimize any order-related biases. Half of the participants (n=20) were presented with one pseudo-randomized order (Version A), while the other half were presented with the inverted order (Version B).

2.3.4 Demographic and Language Background Questionnaire

After completing the three behavioral tasks, participants were asked to fill out a comprehensive Demographic and Language Background Questionnaire. This questionnaire covered demographic information, language proficiency, musical background, and feedback on their experience with the tasks they had just completed. The full list of questionnaire items can be found in Appendix A (Questionnaire A1). Participants completed the questionnaire on the iMac desktop using Limesurvey, an online survey platform. This post-task questionnaire gathered additional information about participants' backgrounds, experiences, and perceptions related to the experiment. The data collected from this questionnaire would be used to better understand potential individual differences and explore any associations between participants' characteristics, feedback, and their performance in the tasks, enabling us to interpret the experimental results in a more comprehensive and contextually informed manner.

2.4 Data Processing

During the data processing phase, 40 sets of participant data from the jabberwocky memory task, along with their corresponding demographic forms, were retained for further analysis. However, for the Urdu memory task, only 30 sets of participant data were included for analysis. The exclusion of 10 sets of participant data from the Urdu memory task was due to the participants’ self-reporting of having knowledge of one or more South Asian languages that shared linguistic similarities with the Urdu language. The decision to exclude these 10 participants was made to ensure the integrity and validity of the study's findings. Participants with prior knowledge of South Asian languages could potentially have an advantage in understanding or recalling the Urdu sentences, which could introduce confounding variables and compromise the accuracy of the results. Additionally, the advantage stemming from their linguistic knowledge could have interacted with the outcome of the speed manipulation used in the Urdu memory task. By excluding these participants, the study aimed to maintain a more homogenous participant sample for the Urdu memory task, thereby enhancing the reliability and generalizability of the findings.

2.4.1 Jabberwocky & Urdu Memory Tasks

Recording of each participant's scores was conducted on individual scoring sheets. These sheets captured the number of correct consonants, syllables, and words recalled for each sentence, as well as identifying the specific consonants, syllables, and words that were correctly recalled. This comprehensive approach provided information not only on the proportion of accuracy for each participant but also insight into the precise points at which participants made mistakes during the recall process. It's important to note that the phonology of the Urdu language includes a number of retroflex, uvular and glottal consonants, as well as nasalized vowels, which are absent in English. This makes it particularly difficult for non-speakers to accurately reproduce these sounds. In scoring responses from the Urdu task, a phonemic approach rather than a phonetic one was utilized, meaning a sound would be scored as correct if the nearest English native phoneme was produced by the participant in response to the non-native Urdu phonemes. This distinction is made because the primary focus of the task is to gauge STM ability, rather than articulation accuracy.

After the scoring on individual scoring sheets, the data were then transferred to a Microsoft Excel spreadsheet. For both the jabberwocky and Urdu memory tasks, the raw scores were converted into proportions representing the correct consonant, syllable, and word recall per sentence. This transformation allowed for a standardized comparison of performance across different sentences and levels of scoring. Two additional columns were included in the spreadsheet to represent the accuracy of recall of phonemes and syllables in serial order. For each sentence, each correctly recalled consonant and syllable was recorded as '1', while each incorrect response was recorded as '0'. This sequential recording of accuracy provided a detailed understanding of participants' recall patterns, enabling the identification of any trends or specific challenges participants encountered during the memory tasks. The data were further processed and analyzed using the Jamovi 2.0.0.0 statistical software package (The jamovi project, 2021).

3 Results

3.1 Inter-Rater Reliability

To assess the inter-rater reliability (IRR) between multiple scorers for the jabberwocky and Urdu scoring tasks, Pearson's *r* correlations were calculated for the consonant level scoring of 75 trials (see Table 1 and 2). The scoring was performed by four native English-speaking scorers for the jabberwocky task and three native Urdu-speaking scorers for the Urdu task. Specifically, each scorer independently scored the same 75 jabberwocky or Urdu responses without any communication or collaboration with the other scorers. This isolation of scoring ensured that each scorer's assessments were independent and unbiased. As the most fine-grained level of scoring among the three, the consonant level provided a detailed examination of participants' accuracy in recalling the words. The inter-rater reliability analysis on this level allowed us to establish the consistency and agreement between multiple scorers in evaluating participants' performance in the jabberwocky and Urdu tasks.

Table 1: IRR correlations on the score proportions of correct consonants for the **jabberwocky memory task**.

		1	2	3	4
1. Scorer 1	Pearsons' *r*	-			
	p-value	-			
2. Scorer 2	Pearson's *r*	0.843	-		
	p-value	<.001	-		

3. Scorer 3	Pearson's *r*	0.858	0.839	-	
	p-value	<.001	<.001	-	
4. Scorer 4	Pearson's *r*	0.825	0.849	0.873	-
	p-value	<.001	<.001	<.001	-

Table 2: IRR correlations on the score proportions of correct consonants for the **Urdu memory task**.

		1	**2**	**3**
1. Scorer 1	Pearsons' *r*	-		
	p-value	-		
2. Scorer 2	Pearson's *r*	0.816	-	
	p-value	<.001	-	
4. Scorer 3	Pearson's *r*	0.863	0.872	-
	p-value	<.001	<.001	-

For the consonant level scores in the jabberwocky sentences, strong positive correlations were observed between any two among the four scorers, and these correlations were found to be statistically significant (Table 1). This indicates a high level of agreement and consistency in the scoring of the jabberwocky sentences between the multiple scorers. The strong positive correlations between their assessments validate the reliability and accuracy of their independent evaluations. Similarly, for the consonant level scoring in the Urdu sentences, conducted by three native Urdu-speaking scorers, strong positive correlations were found between any two among the three scorers, and these correlations were also statistically significant (Table 2). This outcome indicates a high level of agreement and reliability among the multiple scorers in assessing the consonant level scores for the Urdu sentences. Their consistent and statistically significant correlations confirm the reliability and accuracy of their independent scoring.

3.2 Correlational Analysis for Levels of Scoring and Speed Conditions

Pearson's *r* correlations were computed on the scores at each level (consonant, syllable, and word) and across the three speed conditions (normal, fast, and slow). The results revealed strong positive correlations between the scores at the three levels, indicating that participants' repetition accuracy for the jabberwocky and Urdu sentences was strongly correlated across the consonant, to the syllable, and word levels. These correlations were statistically significant. Such correlations indicate that the three levels of scoring are equally valid measures of STM performance on the two memory tasks, and they provide consistent insights into participants' abilities to recall the jabberwocky and Urdu language stimuli. Likewise, for the scores across the three speed conditions, strong positive statistically significant correlations were observed. However, it is important to note that the positive strong correlations across the three speed conditions do not necessarily imply that the speed manipulations did not impact performance. Instead, these correlations indicate that performance on the three speed conditions is similarly affected by individual differences in the task. The results presented in Table 3 provide a comprehensive overview of the significant correlations observed among the various scoring measures.

Table 3: Correlation matrix on the score proportions across three scoring levels and three speed conditions (normal, fast, slow) for the **jabberwocky memory task**. N = 40.

		1	2	3	4	5	6	7	8	9
1. Normal_conProportion	Pearson’s *r*	-								
	p-value	-								
2. Normal_syllProportion	Pearson’s *r*	0.896	-							
	p-value	<.001	-							
3. Normal_wordProportion	Pearson’s *r*	0.844	0.987	-						
	p-value	<.001	<.001	-						
4. Fast_conProportion	Pearson’s *r*	0.972	0.886	0.834	-					
	p-value	<.001	<.001	<.001	-					
5. Fast_syllProportion	Pearson’s r	0.888	0.973	0.956	0.908	-				
	p-value	<.001	<.001	<.001	<.001	-				
6. Fast_wordProportion	Pearson’s *r*	0.858	0.976	0.977	0.869	0.988	-			

	p-value	<.001	<.001	<.001	<.001	<.001	-			
7. Slow_conProportion	Pearson's *r*	0.970	0.870	0.831	0.964	0.871	0.853	-		
	p-value	<.001	<.001	<.001	<.001	<.001	<.001	-		
8. Slow_syllProportion	Pearson's *r*	0.867	0.958	0.955	0.873	0.959	0.967	0.892	-	
	p-value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	-	
9. Slow_wordProportion	Pearson's *r*	0.819	0.944	0.961	0.825	0.936	0.963	0.854	0.988	-
	p-value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	-

Table 4: Correlation matrix on the score proportions across three scoring levels and three speed conditions (normal, fast, slow) for the **Urdu memory task**. N = 30.

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1. Normal_conProportion	Pearson's *r*	-								
	p-value	-								
2. Normal_syllProportion	Pearson's *r*	0.956	-							
	p-value	<.001	-							
3. Normal_wordProportion	Pearson's *r*	0.955	0.994	-						
	p-value	<.001	<.001	-						
4. Fast_conProportion	Pearson's *r*	0.915	0.907	0.903	-					
	p-value	<.001	<.001	<.001	-					
5. Fast_syllProportion	Pearson's *r*	0.879	0.930	0.925	0.967	-				
	p-value	<.001	<.001	<.001	<.001	-				
6. Fast_wordProportion	Pearson's *r*	0.881	0.933	0.930	0.959	0.996	-			
	p-value	<.001	<.001	<.001	<.001	<.001	-			
7. Slow_conProportion	Pearson's *r*	0.926	0.921	0.923	0.919	0.906	0.909	-		
	p-value	<.001	<.001	<.001	<.001	<.001	<.001	-		
8. Slow_syllProportion	Pearson's *r*	0.892	0.933	0.923	0.894	0.924	0.930	0.974	-	
	p-value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	-	
9. Slow_wordProportion	Pearson's *r*	0.890	0.934	0.937	0.890	0.916	0.925	0.975	0.995	-

p-value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	-

3.3 Effect of Speed and Rhythmic Disruption

To assess the effect of rhythmic disruption on STM task performance, a repeated measures ANOVA was conducted using jamovi (version 2.0.0.0). Sentence presentation speed with three levels (normal, fast, and slow) was the repeated measure factor. A similar analysis was performed for all three levels of scoring: consonant, syllable, and word.

3.3.1 Jabberwocky Memory Task Performance

The following tables show the descriptive statistics of the repetition accuracy at three levels of scoring for the jabberwocky memory task (Tables 5 – 7), response accuracy proportions were represented in decimal values, with higher values indicating higher repetition accuracy.

Table 5: Descriptive statistics for the repetition accuracy at **consonant level** in the three speed conditions. N = 40.

	Mean	Median	Standard Deviation	Minimum	Maximum
Normal	0.703	0.721	0.144	0.301	0.942
Fast	0.700	0.718	0.150	0.266	0.936
Slow	0.711	0.706	0.143	0.342	0.933

Table 6: Descriptive statistics for the repetition accuracy at **syllable level** in the three speed conditions. N = 40.

	Mean	Median	Standard Deviation	Minimum	Maximum
Normal	0.594	0.591	0.159	0.308	0.971
Fast	0.585	0.582	0.167	0.258	0.978
Slow	0.591	0.577	0.157	0.286	0.978

Table 7: Descriptive statistics for the repetition accuracy at **word level** in the three speed conditions. N = 40.

	Mean	Median	Standard Deviation	Minimum	Maximum
Normal	0.544	0.541	0.170	0.239	0.881

Fast	0.536	0.526	0.178	0.201	0.873
Slow	0.534	0.500	0.175	0.194	0.858

The repeated measures ANOVAs (see Table 8) yielded no statistically significant differences between the three speed conditions for scores at the consonant level (p = 0.135), syllable level (p = 0.420), or word level (p = 0.346) for the jabberwocky memory task. This indicates that the speed manipulation did not have a significant impact on participants' repetition accuracy at any of these levels.

Table 8: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy for the English-based jabberwocky sentences across three scoring levels. N = 40.

	Sum of Squares	df	Mean Square	*F*	*p*
Consonant - Level	0.00279	2	0.00139	2.05	0.135
Syllable - Level	0.00172	2	0.000858	0.878	0.420
Word - Level	0.00218	2	0.00109	1.08	0.346

3.3.2 Urdu Memory Task Performance

The following tables (9 – 11) show the descriptive statistics of the repetition accuracy at three levels of scoring for the Urdu sentence memory task.

Table 9: Descriptive statistics for repetition accuracy at **consonant level** in the three speed conditions. N = 30.

	Mean	Median	Standard Deviation	Minimum	Maximum
Normal	0.751	0.728	0.127	0.457	0.935
Fast	0.737	0.731	0.126	0.478	0.957
Slow	0.757	0.766	0.119	0.500	0.952

Table 10: Descriptive statistics for repetition accuracy at **syllable level** in the three speed conditions. N = 30.

	Mean	Median	Standard Deviation	Minimum	Maximum
Normal	0.684	0.691	0.144	0.389	0.907
Fast	0.672	0.667	0.150	0.340	0.920
Slow	0.692	0.694	0.146	0.395	0.914

Table 11: Descriptive statistics for repetition accuracy at **word level** in the three speed conditions. N = 30.

	Mean	Median	Standard Deviation	Minimum	Maximum
Normal	0.634	0.646	0.155	0.327	0.867
Fast	0.623	0.611	0.164	0.292	0.903
Slow	0.637	0.633	0.163	0.319	0.903

The repeated measures ANOVA results for the Urdu memory task indicated that there were no statistically significant differences between the three speed conditions at the consonant level ($p = 0.093$), syllable level ($p = 0.174$), and word level ($p = 0.391$). This shows that the speed manipulations did not have a significant impact on participants' repetition accuracy at any of these levels for the Urdu memory task.

Table 12: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy for the Urdu sentences across three scoring levels. N = 30.

	Sum of Squares	df	Mean Square	*F*	*p*
Consonant - Level	0.00621	2	0.00310	2.48	0.093
Syllable - Level	0.00559	2	0.00279	1.81	0.174
Word - Level	0.00346	2	0.00173	0.955	0.391

However, in this truly foreign task, there is a trend for significance at the consonant level. At this level, the slow speed condition seems to be easier to recall than the fast speed ($t(29) = 2.186$; $p = 0.037$).

3.4 Musical and Language Background as Between-Subjects Factors

To assess whether musicianship and bilingual/multilingual status contribute to significant repetition accuracy differences in the three speed conditions, a repeated measures ANOVA was conducted again, with sentence presentation speed at three levels (normal, fast, and slow) as the repeated measure factor, this time including musical and language backgrounds as additional between-subjects control factors. Musicianship and language competence were represented by categorical values, with a value of 1 indicating that a participant reported being at an 'advanced' level in one or more languages/instruments, and a value of 0 indicating 'intermediate', 'beginner', or 'none'. The decision to employ a binary categorization is to maintain an adequate sample size in each category. To ensure the legitimacy of superior musicianship and bilingual status, only participants self-reporting as 'advanced' in either category were considered eligible for musicianship and/or bilingual status, while those reporting 'intermediate', 'beginner', or 'none' levels were not eligible for such qualification. This approach accounts for potential variations in participants' subjective evaluations of their level of proficiency in these areas.

3.4.1 Jabberwocky Memory Task Performance

When controlling for musicianship and bilingual/multilingual status, repeated measures ANOVAs revealed a significant difference between the three speed conditions at consonant level (p = 0.039), see Table 14. Upon conducting post hoc comparisons for the consonant level, a statistically significant difference was found between the fast and slow speed conditions ($t(36)$ = -2.344; p = 0.025, see Table 15), this indicates that participants' overall performance in recalling consonants was significantly better in the slow condition (M = 0.711) compared to the fast condition (M = 0.700). Furthermore, the difference between the normal (M = 0.703) and slow speed conditions was approaching significance ($t(36)$ = -1.836; p = 0.075), see Table 15. This suggests a trend towards overall better repetition accuracy in the slow condition compared to the normal condition. When assessing between-subjects effects, neither musicianship (p = 0.858), bilingual/multilingual status (p = 0.576), nor the interaction of the two factors (p = 0.684), yielded statistical significance.

Table 13: Distribution of participant bi/multilingual status and musicianship in the **jabberwocky memory task**. N = 40. Bi/multilinguals = 13. Musicians = 13.

Language	Musicianship	N
Monolingual	**Non-Musician**	16
	Musician	11
Bi/Multilingual	**Non-Musician**	11
	Musician	2

Table 14: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy at **consonant level**, with musicianship and bi/multilingual status as between-subjects factors. N = 40.

	Sum of Squares	df	Mean Square	*F*	*p*
Speed	0.00465	2	0.00233	3.390	0.039*
Speed*Music	0.000936	2	0.000468	0.682	0.509
Speed*Language	0.00251	2	0.00125	1.826	0.168
Speed*Music*Language	0.00139	2	0.000694	1.012	0.369

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 15: Post hoc comparisons of the effect of three speed conditions on repetition accuracy at **consonant level**, with musicianship and bi/multilingual status as between-subjects factors. N = 40.

Comparison							
Speed		Speed	Mean Difference	SE	df	*t*	*p*
Normal	-	**Fast**	0.00663	0.00773	36.0	0.859	0.396
	-	**Slow**	-0.01376	0.00749	36.0	-1.836	0.075
Fast	-	**Slow**	-0.02039	0.00870	36.0	-2.344	0.025*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

At the syllable level, the repeated measures ANOVA revealed a significant interaction between speed and bilingual/multilingual status on recall performance ($p = 0.047$), see Table 16. This indicates that the impact of the speed manipulation on participants' recall accuracy for syllables differed depending on their bilingual/multilingual status. Upon conducting post hoc comparisons, a difference approaching significance in repetition accuracy was found between the normal and slow speed conditions, specifically among the bilinguals/multilinguals ($t(36) = -1.964$; $p = 0.057$). This suggests a trend towards better performance in the slow condition (M = 0.667) compared to the normal condition (M = 0.648) for the bilingual/multilingual participants. A complete list of post hoc comparisons can be found in Appendix C (Table C1). However, neither musicianship ($p = 0.662$), bilingual/multilingual status ($p = 0.642$), nor the interaction of the two factors ($p = 0.501$), had statistically significant between-subjects effects.

Table 16: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy at **syllable level**, with musicianship and bi/multilingual status as between-subjects factors. N = 40.

	Sum of Squares	df	Mean Square	*F*	*p*
Speed	0.00196	2	0.000980	1.013	0.368
Speed*Music	0.000833	2	0.000417	0.431	0.652
Speed*Language	0.00616	2	0.00308	3.182	0.047*
Speed*Music*Language	0.00178	2	0.000892	0.923	0.402

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

At the word level, the repeated measures ANOVA also revealed a significant interaction between speed and bilingual/multilingual status on recall performance ($p = 0.030$), see Table 17. This again indicates that the impact of the speed manipulation on participants' recall accuracy for syllables differed depending on their bilingual/multilingual status. Upon conducting post hoc comparisons, a significant difference in repetition accuracy was found between the normal and slow speed conditions, specifically among the monolinguals ($t(36) = 2.137$; $p = 0.039$). This suggests a significantly better performance in the normal condition (M = 0.514) compared to the slow condition (M = 0.494) for the monolingual participants. A complete list of post hoc comparisons can be found in Appendix C (Table C2). Again, neither musicianship ($p = 0.574$), bilingual/multilingual status ($p = 0.584$), nor the interaction of the two factors ($p = 0.451$), had statistically significant between-subjects effects.

Table 17: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy at **word level**, with musicianship and bi/multilingual status as between-subject factors. N = 40.

	Sum of Squares	df	Mean Square	*F*	*p*
Speed	0.000637	2	0.000318	0.328	0.722
Speed*Music	0.00210	2	0.00105	1.081	0.345
Speed*Language	0.00715	2	0.00357	3.677	0.030*
Speed*Music*Language	0.00119	2	0.000593	0.610	0.546

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Note that the direction of the effect of speed on recall performance appears to differ between monolingual and bilingual/multilingual participants. For monolingual participants, performance for word recall in the normal condition was significantly better compared to the slow condition. This suggests that monolingual individuals may benefit from the non-disrupted speech rate in terms of recalling whole words in the jabberwocky sentences. On the other hand, the recall of syllables by the bilingual/multilingual participants in the slow condition trends towards being better compared to in the normal condition. This indicates that bilingual/multilingual individuals may not be affected by the disrupted speech rate but instead show improved recall accuracy for individual syllables in the jabberwocky sentences when presented at slower speech rates.

3.4.2 Urdu Memory Task Performance

The repeated measures ANOVA analysis for the Urdu memory task scores, with musicianship and bilingual/multilingual status as the between-subjects factors, did not yield any statistically significant differences between the three speed conditions at any of the three scoring levels (Table 19 – 21). This indicates that the speed manipulation did not have a significant impact on participants' recall accuracy for consonants, syllables, or words in the Urdu memory task, regardless of their musicianship and bilingual/multilingual status. Furthermore, there were no significant interactions found between speed and musicianship or bilingual/multilingual status at any of the scoring levels.

Table 18: Distribution of participant bi/multilingual status and musicianship in the Urdu memory task. N = 30. Bi/multilinguals = 9. Musicians = 12.

Language	Musicianship	N
Monolingual	**Non-Musician**	11
	Musician	10
Bi/Multilingual	**Non-Musician**	7
	Musician	2

Table 19: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy at **consonant level**, with musicianship and bi/multilingual status as between-subjects factors. N = 30.

	Sum of Squares	df	Mean Square	*F*	*p*
Speed	0.00549	2	0.00275	2.227	0.118
Speed*Music	0.00110	2	0.000550	0.446	0.643
Speed*Language	0.00181	2	0.000906	0.735	0.484
Speed*Music*Language	0.00231	2	0.00115	0.936	0.399

Table 20: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy at **syllable level**, with musicianship and bi/multilingual status as between-subjects factors. N = 30.

	Sum of Squares	df	Mean Square	*F*	*p*
Speed	0.00405	2	0.00203	1.302	0.281
Speed*Music	0.00332	2	0.00166	1.067	0.351
Speed*Language	0.000479	2	0.000239	0.154	0.858
Speed*Music*Language	0.000609	2	0.000305	0.196	0.823

Table 21: Repeated measures ANOVA on the effect of three speed conditions on repetition accuracy at **word level**, with musicianship and bi/multilingual status as between-subjects factors. N = 30.

	Sum of Squares	df	Mean Square	*F*	*p*
Speed	0.00178	2	0.000889	0.479	0.622
Speed*Music	0.00131	2	0.000656	0.353	0.704
Speed*Language	0.00105	2	0.000526	0.283	0.754
Speed*Music*Language	0.00255	2	0.00128	0.688	0.507

This suggests that participants' performance in the Urdu memory task was not significantly influenced by their musical background or language competencies when considering the different speech rate conditions.

3.5 Repetition Accuracy of Different Portions of the Target Sentence

As demonstrated in the previous sections, we did not observe significant differences in overall repetition accuracy between different speed conditions for the scores on entire sentences across the two STM tasks. However, conducting a more fine-grained analysis, focusing on recall accuracy at different positions within the sentence (initial, medial, final), could provide valuable insights. This analysis aimed to investigate whether rhythmic disruption between introductory prime and target sentence does indeed have an effect on performance. Particularly the repetition of the initial portion of the target sentence could have been affected, before rate expectations would have been quickly corrected as participants adjusted to the new rhythm.

Furthermore, our motivation for this analysis also stems from the participants' self-perceived difficulty reported in the demographic and language background questionnaire. Participants were asked to indicate which condition they found the most challenging, and the reasons behind their selections. Among the responses, a substantial number (46.81%) identified the slow condition as the most difficult to recall, with 34.02% reporting the fast condition to be most challenging, and 14.89% finding the normal condition most difficult. Interestingly, some participants (4.26%) considered all three conditions to be equally challenging. Among those who found the slow version most difficult, a majority (15 out of 22) specifically mentioned struggling to recall the beginning of the sentence by the time they finished listening to the whole sentence. In light of these self-perceived difficulties, we aimed to achieve two objectives through this analysis: 1) to see if the participants' self-perceived difficulty with the beginning part of the slow condition aligns with their actual performance; and 2) to explore whether there are any significant differences in accuracy at different positions within the sentence that could explain the discrepancy between introspections about the difficulty in the slow and fast conditions, and actual performance, which did not show significant differences across the three speed conditions.

The sequential recordings of accuracy for participant responses were extracted and utilized in this analysis. The procedure for conducting the sequential recording is detailed in section 2.4.1. For the jabberwocky responses, we excluded the first syllable response from the analysis, as the first syllable would always be a jabberwocky function word. These function words were repeated frequently in the stimuli, and to prevent skewing the results towards higher accuracy in the initial position, we decided to exclude them. After excluding the first syllable, the remaining jabberwocky responses and the Urdu responses were grouped as follows: the first two syllables were grouped into the initial position, the last two syllables were grouped into the final position, and whatever syllables were in between (ranging from 1-4 syllables) were grouped into the medial position. Subsequently, we compared the proportion of accuracy for these positions both within and across the three speed conditions. This analysis allowed us to examine if there were any significant differences in accuracy at different positions within the sentence and whether these differences varied across the three speed conditions.

3.5.1 Recall Accuracy for Portions of Jabberwocky Sentences

In this analysis, speed condition and sentence position (initial, medial, final) were used as the repeated measures factors to examine whether speed or the interaction between speed condition and sentence position had an effect on task performance. The repeated measures ANOVA results showed significant effect of sentence position on participants' overall recall accuracy, specifically, participants' overall recall accuracy for syllables was significantly better in the initial position (M = 0.663), followed by the final position (M = 0.540), and poorest for the medial position (M = 0.510). However, no significant effect of speed on performance across different serial positions ($p = 0.207$). Additionally, there was no significant interaction between speed and sentence position on performance ($p = 0.495$), see Tables 22 and 23.

Table 22: Repeated measures ANOVA on the effect of sentence position and speed for the **jabberwocky memory task.** N = 40.

	Sum of Squares	df	Mean Square	*F*	*p*
Position	1.57519	2	0.78759	30.480	<.001***
Speed Condition	0.01175	2	0.00588	1.608	0.207
Position*Speed Condition	0.00882	4	0.00220	0.851	0.495

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 23: Descriptive statistics for proportion accuracy across three sequential sentence positions and three speed conditions in the **jabberwocky memory task**. N = 40.

	Normal Initial	Normal Medial	Normal Final	Fast Initial	Fast Medial	Fast Final	Slow Initial	Slow Medial	Slow Final
Mean	0.662	0.520	0.547	0.664	0.495	0.529	0.662	0.516	0.542
Median	0.700	0.507	0.510	0.700	0.497	0.530	0.680	0.483	0.500
SD	0.174	0.203	0.179	0.180	0.207	0.189	0.164	0.204	0.189
Minimum	0.060	0.133	0.240	0.060	0.133	0.180	0.060	0.127	0.180
Maximum	0.960	0.900	0.880	0.900	0.920	0.940	0.880	0.907	0.920

Contrary to our expectations and participant self-reports, these findings indicate that the speed at which the stimuli were presented did not have a significant impact on participants' performance across different positions within the sentence. Similarly, there was no evidence for an interaction effect between speed and sentence position, suggesting that the pattern of performance across positions did not vary significantly across different speed conditions.

3.5.2 Recall Accuracy for Portions of Urdu Sentences

The same set of analyses as for the English-based jabberwocky sentences, was performed on the Urdu memory task scores. The repeated measures ANOVA results showed a significant effect of sentence position on recall accuracy, participants' overall recall accuracy for syllables was significantly better in the initial position (M = 0.826), followed by the medial position (M = 0.649), and poorest for the final position (M = 0.614). Again, no significant effect of speed on performance across different serial positions (p = 0.144) was detected. Furthermore, there was no significant interaction between speed and sentence position on performance (p = 0.183), see Tables 24 and 25.

Table 24: Repeated measures ANOVA on the effect of sentence position and speed for the **Urdu memory task.** N = 30.

	Sum of Squares	df	Mean Square	F	p
Position	2.3258	2	1.16289	123.68	<.001***

Speed Condition	0.0181	2	0.00903	2.00	0.144
Position*Speed Condition	0.0157	4	0.00393	1.58	0.183

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 25: Descriptive statistics for proportion accuracy across three sequential sentence positions and three speed conditions in the **Urdu memory task**. N = 30.

	Normal Initial	Normal Medial	Normal Final	Fast Initial	Fast Medial	Fast Final	Slow Initial	Slow Medial	Slow Final
Mean	0.829	0.648	0.625	0.827	0.635	0.593	0.823	0.664	0.625
Median	0.860	0.645	0.620	0.880	0.605	0.590	0.840	0.650	0.620
SD	0.118	0.172	0.157	0.125	0.165	0.186	0.118	0.174	0.152
Minimum	0.500	0.227	0.320	0.500	0.230	0.220	0.540	0.303	0.360
Maximum	0.960	0.930	0.900	0.980	0.920	0.900	0.980	0.953	0.900

These findings suggest that the speed at which the Urdu stimuli were presented did not significantly impact participants' performance across different positions within the sentence, and no significant variation in the pattern of performance across positions across different speed conditions could be detected.

3.6 Effect of Speech Rate Condition on Response Speech Rate

Given that no overall significant differences were observed in participant performance between different speed conditions across the two STM tasks, we sought to explore other potential influences of the speech rates of the stimuli on participants' responses. Specifically, we aimed to assess whether participants' speech rates increased or decreased in accordance with the stimulus speech rate, which would confirm that our speed manipulation indeed had some effect on participant responses. Additionally, we aimed to investigate whether participants' response speech rate aligned more with the introductory (prime) speech rate or the stimulus (target) speech rate. This investigation was carried out to determine whether our prime sentences created a rhythmic pattern that participants adhered to when reproducing the information.

To investigate the potential influence of the prime and target sentence speech rates on participants' responses, participants' recall were extracted and compared across the three speed conditions. Specifically, responses that achieved 100% accuracy on the word level were identified and measured. The length of these accurate responses was then divided by the number of syllables expected in the respective sentences to establish the participants' recall speech rates. By focusing on responses with perfect word-level accuracy, the aim was to obtain reliable and precise measurements of participants' speech rates during the recall process. This approach also ensured that any potential inaccuracies caused by missing or added consonants, syllables, or words in the responses were minimized, allowing for a more robust analysis of how speech rate variations in the stimuli might impact participants' own speech rates. For the analysis, only participants with more than ten 100% accurate responses were included to ensure an adequate and reliable sample size. To account for both within-subject variability and between-subject variability, a linear mixed effects model analysis was employed. This statistical approach takes into account the repeated measures within-subject variability and the between-subject variability. In this analysis, target sentence speed condition was used as a factor to examine its effect on participants' speech rates. To ensure the appropriate handling of the clustered nature of the data, both participant and sentence identifiers were used as random cluster variables. This allowed us to control for any potential dependencies between responses from the same participant or associated with the same sentence.

3.6.1 Jabberwocky Responses Speech Rates

Descriptive statistics of speech rates in the different conditions are presented in Table 26 (see also Figure 1). The linear mixed model analysis revealed a significant difference in participants' response speech rate across the different speed conditions ($p < .001$). Post hoc comparisons further elucidated these differences, showing significant differences in response speech rate between the normal and fast speed conditions ($t(33.6) = -2.26$; $p = 0.030$), the slow and fast speed conditions ($t(26.2) = -4.40$; $p < .001$), as well as the slow and normal speed conditions ($t(32.4) = -3.18$; $p = 0.003$). The fastest rate was found in the fast condition and the slowest in the slow condition, speech rates are represented here as syllables per second. Statistical significance of the differences between the slow and fast conditions ($p_{bonferroni}$ = <.001) and between the slow and normal conditions ($p_{bonferroni}$ = 0.010) remained after applying Bonferroni corrections. For a detailed summary of the results, refer to Tables 27 and 28.

Table 26: Descriptive statistics for speech rates (syllables per second) across speed conditions for the **jabberwocky memory task**. N = 318.

	Speed Condition	N	Mean	Median	SD	Minimum	Maximum
Speech Rate	**Normal**	106	3.38	3.36	0.493	2.09	4.44
	Fast	108	3.75	3.79	0.639	2.02	5.30
	Slow	104	3.07	3.05	0.449	1.99	4.10

Table 27: Linear mixed model analysis of the effect of speed conditions on response speech rates for the **jabberwocky memory task**. N = 318.

	F	Num df	Den df	*p*
Speed Conditions	10.1	2	31.7	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 28: Post hoc comparisons of the response speech rates in different speed conditions for the **jabberwocky memory task**. N = 318.

Comparison								
Speed		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Normal	-	**Fast**	-0.262	0.116	33.6	-2.26	0.030*	0.091
Slow	-	**Fast**	-0.612	0.139	26.2	-4.40	<.001***	<.001***
Slow	-	**Normal**	-0.350	0.110	32.4	-3.18	0.003**	0.010*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

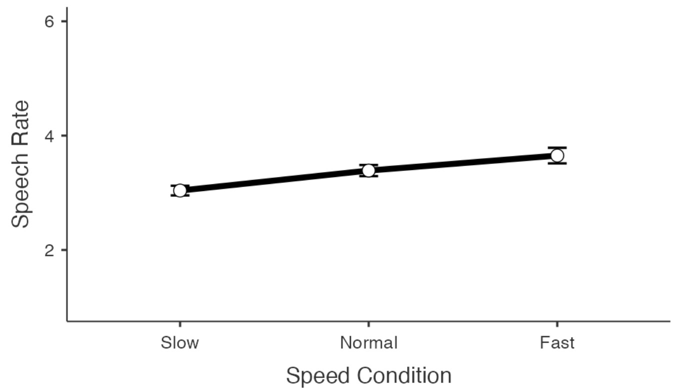


Figure 1: Effect plot of the response speech rates (syllables per second) in different speed conditions for the **jabberwocky memory task**. N = 318.

The results demonstrate a clear trend where participant response speech rate increases as the target sentence speech rate increases. As the target sentence speech rate becomes faster, participants tend to respond with faster speech rates as well. This effect suggests that the speed manipulation in the stimuli had a direct impact on participants' speech rate for reproducing the target sentences.

The next analysis was conducted to determine whether participant responses were more closely aligned with the speech rates of the target sentences or the prime clauses. This examination was carried out for each of the three prime speed conditions: normal, fast, and slow. By comparing the participant response rates with both the target sentence speech rates and the prime speech rates, we sought to identify whether the speed of the target sentences or the prime clauses had a more significant influence on the response speech rates. Results from the normal speed condition showed a significant difference between the speech rates of prime (M = 3.99), target (M = 3.72), and response (M = 3.38) sentences (p = <.001; Table 29. Post hoc comparisons revealed a significant difference in speech rate between any two sentence types (Table 30), with the response speech rates being more closely aligned with the target stimulus speech rates (see Figure 2).

Table 29: Linear mixed model analysis of the difference in speech rates across prime, target, and response sentences for the **normal speed condition**. N = 106.

	F	Num df	Den df	p

Sentence Type	34.1	2	18.0	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 30: Post hoc comparisons of the speech rates for prime, target, and response sentences in the **normal speed condition**. N = 106.

Comparison									
Sentence Type		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Prime	-	**Response**	0.600	0.0853	19.8	7.03	<.001***	<.001***
Prime	-	**Target**	0.276	0.0461	16.3	5.99	<.001***	<.001***
Target	-	**Response**	0.324	0.0839	19.7	3.86	<.001***	0.003**

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

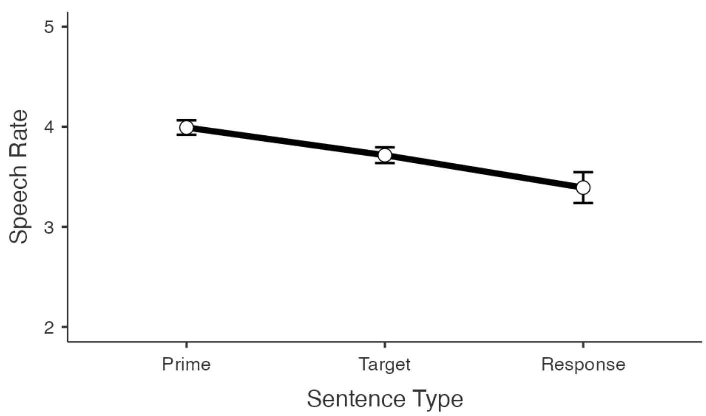


Figure 2: Effect plot on the differences in speech rates across prime, target, and response sentences in the **normal speed condition**. N = 106.

Results from the fast speed condition also showed significant differences between the speech rates of prime (M = 4.03), target (M = 5.02), and response (M = 3.75) sentences (*p* = <.001; Table 31). Post hoc comparisons reveal a significant difference in speech rate between any two sentence types (Table 32), with the response speech rates being more closely aligned with the introductory primes speech rates (see Figure 3).

Table 31: Linear mixed model analysis of the difference in speech rates across prime, target, and response sentences for the **fast speed condition**. N = 108.

	F	Num df	Den df	*p*
Sentence Type	271	2	38.9	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 32: Post hoc comparisons of the speech rates for prime, target, and response sentences in the **fast speed condition**. N = 108.

Comparison									
Sentence Type		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Prime	**-**	**Response**	0.350	0.0881	20.9	3.97	<.001***	0.002**
Prime	**-**	**Target**	-0.987	0.0606	16.8	-16.28	<.001***	<.001***
Target	**-**	**Response**	1.337	0.0842	20.6	15.86	<.001***	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

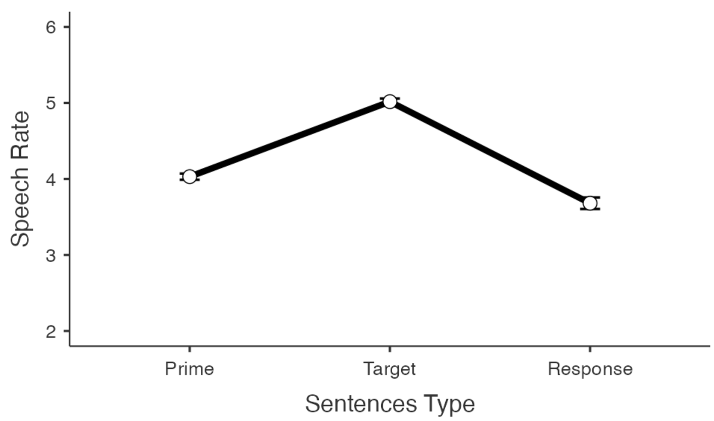


Figure 3: Effect plots on the differences in speech rates across prime, target, and response sentences in the **fast speed condition**. N = 108.

Lastly, results from the slow speed condition (see Figure 4) showed significant differences between the speech rates of prime (M = 4.02), target (M = 2.86), and response (M = 3.07) sentences (p = <.001; Table 33). Post hoc comparisons revealed a significant difference in speech rate between any two sentence types (Table 34), with the response speech rates being more closely aligned with the stimuli speech rates.

Table 33: Linear mixed model analysis of the difference in speech rates across prime, target, and response sentences for the **slow speed condition**. N = 104.

	F	Num df	Den df	*p*
Sentence Type	435	2	26.5	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 34: Post hoc comparisons of the speech rates for prime, target, and response sentences in the **slow speed condition**. N = 104.

Comparison									
Sentence Type		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Prime	-	**Response**	0.980	0.0879	23.18	11.16	<.001***	<.001***
Prime	-	**Target**	1.137	0.0435	8.39	26.14	<.001***	<.001***
Target	-	**Response**	-0.157	0.0929	24.21	-1.69	0.104	0.313

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

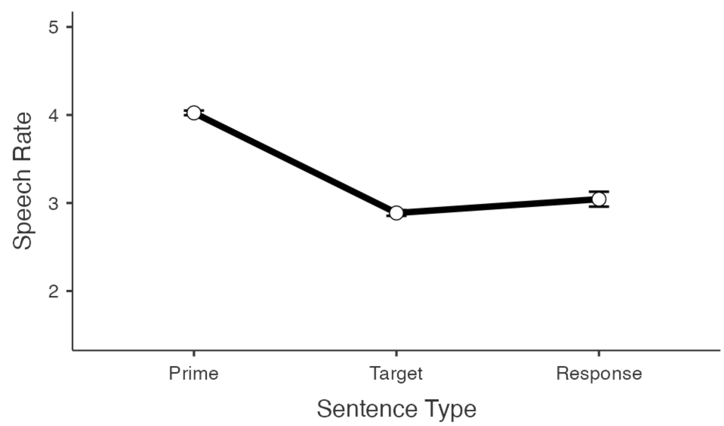


Figure 4: Effect plot on the differences in speech rates across prime, target, and response sentences in the **slow speed condition**. N = 104.

3.5.2 Urdu Responses Speech Rates

Similarly to the findings for jabberwocky responses, linear mixed model analysis of the Urdu responses speech rates also revealed significant differences among participants' response speech rates across the different speed conditions (p <.001; see Tables 35 and 36). Post hoc comparisons further elucidated these differences, showing significant differences in response speech rate between the normal and fast speed conditions ($t(54) = -2.43$; $p = 0.019$), the slow and fast speed conditions ($t(53.4) = -4.49$; p

<.001), as well as the slow and normal speed conditions (*t*(44.9)= -2.56; *p* = 0.014). Statistical significance between the slow and fast conditions ($p_{bonferroni}$ = <.001) and between the slow and normal conditions ($p_{bonferroni}$ = 0.042) remained after applying Bonferroni corrections. For a detailed summary of the results, refer to Tables 36 and 37.

Table 35: Descriptive statistics for speech rates (syllables per second) across speed conditions for the **Urdu memory task**. N = 581.

	Speed Condition	N	Mean	Median	SD	Minimum	Maximum
Speech Rate	**Normal**	203	3.90	3.95	0.605	2.15	5.38
	Fast	185	4.21	4.21	0.771	2.07	5.80
	Slow	193	3.56	3.61	0.597	2.03	5.40

Table 36: Linear mixed model analysis of the effect of speed conditions on response speech rates for the **Urdu memory task**. N = 581.

	F	Num df	Den df	*p*
Speed Conditions	10.4	2	48.7	<.001***

Note. * *p* < .05, ** *p* < .01, *** *p* < .001.

Table 37: Post hoc comparisons of the response speech rates in different speed conditions for the **Urdu memory task**. N = 581.

Comparison								
Speed		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Normal	**-**	**Fast**	-0.344	0.142	54.0	-2.43	0.019*	0.056
Slow	**-**	**Fast**	-0.653	0.146	53.4	-4.49	<.001***	<.001***
Slow	**-**	**Normal**	-0.310	0.121	44.9	-2.56	0.014*	0.042*

Note. * *p* < .05, ** *p* < .01, *** *p* < .001.

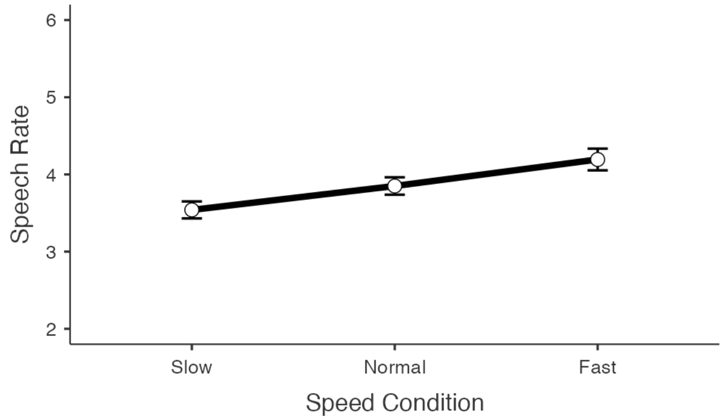


Figure 5: Effect plot of the response speech rates (syllables per second) for different speed conditions for the **Urdu memory task**. N = 581.

Results again demonstrated a clear trend where participant response speech rate increases as the target sentence speech rate increases. This suggests that the speed manipulations employed to the sentence stimuli had a direct impact on participants' recall speech rate. A similar set of analyses as for the English-based jabberwocky sentences was conducted to determine whether participant responses were more closely aligned with the speech rates of the target sentences or those of the prime clauses. This examination was carried out separately for each of the three speed conditions: normal, fast, and slow.

Results from the normal speed condition (see Figure 6) show significant differences between the speech rates of the prime (M = 4.66), target (M = 4.47), and response (M = 3.90) sentences (p = <.001, see Table 38). Post hoc comparisons revealed significant differences in speech rate between the prime clauses and response sentences, and between the target and response sentences (Table 39). The response speech rates being more closely aligned with the target sentence speech rates.

Table 38: Linear mixed model analysis of the difference in speech rates across prime, target, and response sentences for the **normal speed condition**. N = 203.

	F	Num df	Den df	*p*
Sentence Type	47.3	2	24.0	<.001***

Note. * p < .05, ** p < .01, *** p < .001.

Table 39: Post hoc comparisons of the speech rates for prime, target, and response sentences in the **normal speed condition**. N = 203.

Comparison									
Sentence Type		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Prime	-	**Response**	0.806	0.1049	23.6	7.68	<.001***	<.001***
Prime	-	**Target**	0.220	0.1423	23.9	1.54	0.136	0.408
Target	-	**Response**	0.586	0.0976	23.6	6.01	<.001***	<.001***

Note. * *p* < .05, ** *p* < .01, *** *p* < .001.

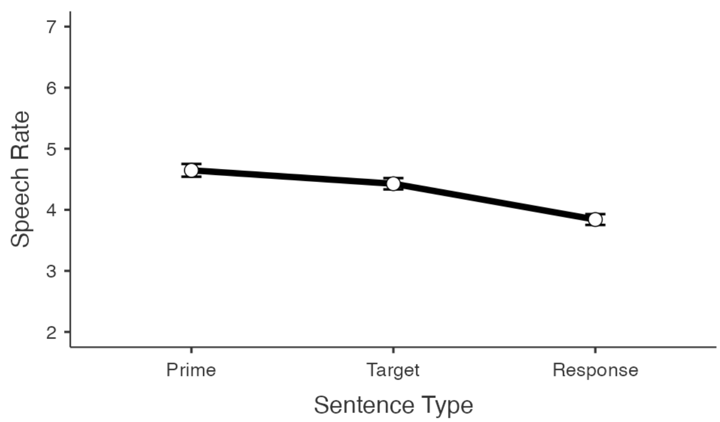


Figure 6: Effect plot of the difference in speech rates across prime, target, and response sentences in the **normal speed condition**. N = 203.

Results from the fast speed condition (see Figure 7) also showed significant differences between the speech rates of prime (M = 4.67), target (M = 5.97), and response (M = 4.21) sentences (*p* = <.001; see Table 40). Post hoc comparisons revealed significant differences in speech rate between any two sentence types (Table 41), with the response speech rates being more closely aligned with the prime phrase speech rates.

Table 40: Linear mixed model analysis of the difference in speech rates across prime, target, and response sentences for the **fast speed condition**. N = 185.

	F	Num df	Den df	*p*
Sentence Type	120	2	22.8	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 41: Post hoc comparisons of the speech rates for prime, target, and response sentences in the **fast speed condition**. N = 185.

Comparison								
Sentence Type		Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$
Prime	-	**Response**	0.474	0.124	22.8	3.82	<.001***	0.003**
Prime	-	**Target**	-1.268	0.168	23.0	-7.53	<.001***	<.001***
Target	-	**Response**	1.742	0.116	22.8	15.02	<.001***	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

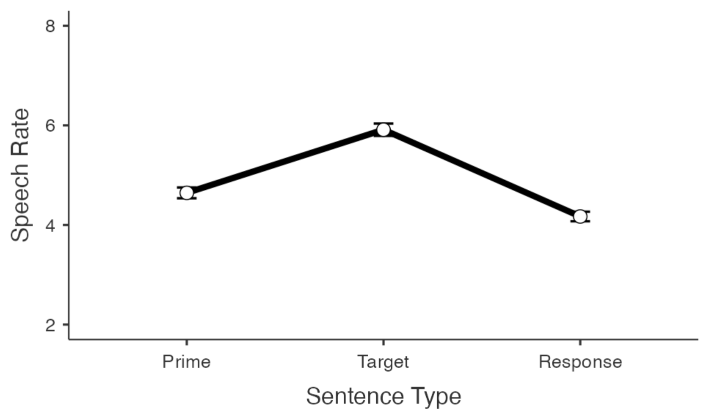


Figure 7: Effect plot of the difference in speech rates across prime, target, and response sentences in the **fast speed condition**. N = 185.

Table 42: Linear mixed model analysis of the difference in speech rates across prime, target, and response sentences for the **slow speed condition**. N = 193.

	F	Num df	Den df	*p*
Sentence Type	52.7	2	21.1	<.001***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 43: Post hoc comparisons of the speech rates for prime, target, and response sentences in the **slow speed condition**. N = 193.

Comparison							
Sentence Type	Speed Difference	Difference	SE	df	*t*	*p*	$p_{bonferroni}$

Prime	**-**	**Response**	1.091	0.1114	21.9	9.79	<.001***	<.001***
Prime	**-**	**Target**	1.322	0.1359	22.0	9.73	<.001***	<.001***
Target	**-**	**Response**	-0.231	0.0798	21.7	-2.89	0.008**	0.025*

Note. * *p* < .05, ** *p* < .01, *** *p* < .001.

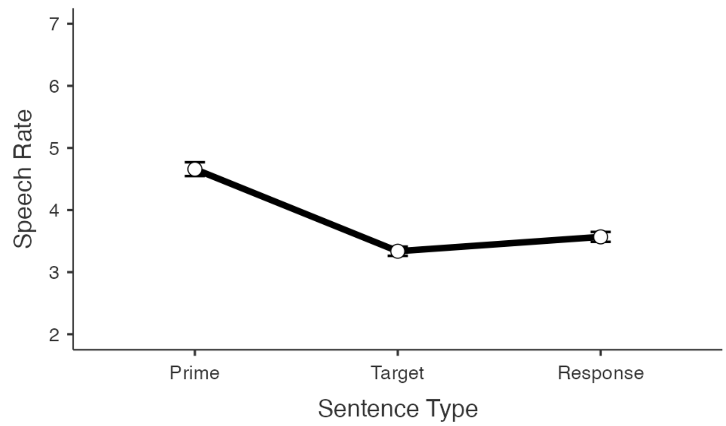


Figure 8: Effect plot of the difference in speech rates across prime, target, and response sentences in the **slow speed condition**. N = 193.

Lastly, results from the slow speed condition (see Figure 8) show significant differences between the speech rates of prime (M = 4.71), target (M = 3.35), and response (M = 3.56) sentences (*p* = <.001; see Table 42), see Table 42. Post hoc comparisons (Table 43) revealed significant differences in speech rate between any two sentence types, with the response speech rates being more closely aligned with the target sentence speech rates.

4 Discussion

This thesis investigated the impact of manipulating the consistency of the spoken speed of heard target sentences. An inconsistency in speed between an introductory phrase and the target part of the sentence disrupted their inherent rhythms, and was expected to affect individuals' verbal STM for those sentences. This investigation was motivated by prior findings indicating impaired perception and comprehension as a result of rhythmic disruptions (Kösem et al., 2018; Li et al., 2019), discussed in sections 1.4 and 1.5. The study took inspiration from the methodology employed by Kösem et al. (2018), where prime phrases are initially presented at their original speed, followed by target sentences presented at different rates: the same rate as prime phrases (normal speed), 133% of prime phrases (slow speed), or 75% of prime phrases (fast speed). Our study incorporated two different STM tasks: the jabberwocky memory task and the Urdu memory task. The jabberwocky task necessitates the temporary retention of a sequence of familiar phonemes, organized according to English morphological and syntactic patterns. Subsequently, participants are tasked with retrieving and repeating this sequence from their STM. A simpler form of task with single items or item pairs as repetition material has been widely employed as a gauge of individual phonological STM. The items have mostly been phonetically and morphologically plausible pseudowords within a given language (Baddeley, Gathercole, & Papagno, 1998; Service, 1992; Service & Craik, 1993). The second task in the experiment, the Urdu memory task, followed a similar rationale as the jabberwocky memory task. It required participants to temporarily hold a sequence of phonemes in their STM, then retrieve and repeat this sequence. However, the Urdu memory task employed an actual foreign language that significantly differed from English both phonologically and prosodically, *i.e.*, with respect to its intonation, stress, and rhythm.

Initially, we hypothesized that STM for sentences at altered speeds, either slower or faster, would be compromised compared to sentences at a normal speed. This assumption was based on the notion that speed manipulation could disrupt participants' established rhythmic attention patterns, potentially affecting STM. However, our results did not reveal any significant differences in repetition accuracy among the three speed conditions across two STM tasks. Additionally, a secondary analysis that compared accuracy across various sentence positions also failed to reveal a discernible pattern of performance variation among the three speech conditions. This suggests that rhythmic disruption might not significantly impact verbal STM accuracy. I will discuss below whether this is the only interpretation. Two potential explanations are considered for the absence of speech manipulation effects on STM in our study. Firstly, the speech manipulations employed in the current study may not have effectively induced rhythmic irregularity. Secondly, the nature of the STM task itself may have compelled participants to allocate more attention to the target parts of the sentences, potentially overriding any possible impact of rhythmic disruption.

4.1. The Speed Manipulation Might Not Have Created Rhythmic Irregularity

As previously mentioned, this study drew inspiration from certain methodological aspects of Kösem et al. (2018), particularly the manipulation of the spoken speed of the target sentence. However, there are differences in the specifics of the methodology. In Kösem et al. (2018), the prime window comprised 12 syllables, while the target window contained only 3 syllables. The primary focus of that study was on the perception of the last syllable within the target window. In contrast, the current study was concerned with assessing verbal STM for longer target sentences. To achieve this objective, we made the decision to shorten the prime phrases and extend the length of the target sentences. In the jabberwocky memory task, the prime phrases consisted of 6 syllables, and the target sentences varied in length from 6 to 8 syllables. Similarly, in the Urdu memory task, the prime phrases also encompassed 6 syllables, while the target sentences ranged from 5 to 8 syllables in length.

One possible reason for the results of the current study is that altering the speed of target sentences did not genuinely result in the creation of irregular rhythm. These suspicions initially arose from the examination of participants' response speech rates in relation to both the prime phrases and target sentence speech rates. The findings revealed that participants' response speech rates not only changed correspondingly with the target speech rates, but they were also more closely aligned with the target speech rates in both the slow and normal conditions. This suggests that the prime phrases had a minor priming influence on speech rate. Only in the fast speech condition, did participants' response speech rates seem to be more aligned with the prime speech rate. We speculate that this might be due to the challenge of reproducing target sentences at a rapid pace while simultaneously maintaining articulatory accuracy.

The apparent failure to induce rhythmic irregularity and disruption by introducing a rhythmic mismatch between the prime phrase and target sentences can be attributed to two potential causes. Firstly, it's important to note that rhythmic irregularity is a comparative concept. In Kösem et al. (2018), a significant portion of the stimulus sentence was presented at an altered speed, while only the final three syllables were presented at the original speed. This created a perceptual contrast between the rhythm of the target portion and the prime portion, leading to a sense of rhythmic irregularity. Conversely, in our study, the prime phrases were of similar length to the target sentences, and in some cases, even shorter. This meant that the sped up and slowed down target phrases might not have been perceived as rhythmically irregular. Additionally, due to the extended target window, participants might have swiftly adjusted to the new speech rates, diminishing the perception of rhythmic disruption.

Secondly, the lack of priming effect from the prime phrases might have arisen from a top-down influence on focusing attention towards the target sentence. As highlighted in the review by Haegens and Golumbic (2018), the synchronization of neural oscillations and behavioral performance with external rhythms can be modulated in a top-down manner to align with cognitive objectives. In studies by Lakatos and her colleagues (Lakatos et al., 2007; 2008; 2009; 2013; 2016), monkeys were engaged in an oddball detection task involving two concurrent rhythmic input streams originating in different sensory modalities, such as auditory and visual stimuli. The monkeys were trained to direct their attention and response to one stream at a time. The outcomes revealed phase synchronization in the sensory cortex corresponding to the attended stream, while weaker synchronization was noticed in the sensory cortex associated with the non-attended modality. Numerous other studies on selective attention have similarly demonstrated heightened responses to rhythmic stimuli when attended to, in comparison to when they are ignored, across diverse sensory modalities (Kim et al., 2006; Bidet-Caulet et al., 2007; Keitel et al., 2017). In our current study, participants were explicitly instructed to disregard the introductory prime phrases. This instruction might have led to a reduced focus on the temporal patterns within these primes, ultimately resulting in a diminished potential for any priming effect to take place.

To conclude, the lack of effective rhythmic irregularity induced by our speed manipulation could be attributed to two significant factors. Firstly, the nature of rhythmic irregularity is relative and hinges on observed perceptual contrasts. Secondly, the absence of pronounced priming effects from the prime phrases could be due to top-down modulation, particularly the allocation of attention mainly to the target sentence.

4.2 Interpretation of Rhythm Effects in STM Tasks in General

As discussed in section 1.5, the potential discrepancies in performance observed between the perception tasks conducted by Kösem et al. (2018) and Li et al. (2019), and the STM tasks conducted by Gorin (2020) and Lui and Wöstmann (2022), may be attributed to the inherent differences in task nature. Specifically, comprehension achieved through passive perception may not necessarily align with STM accuracy, as the latter requires active information retention and, particularly in the case of repetition responses in a verbal STM task, articulation. Moreover, the allocation of attentional resources during an STM task is notably higher compared to a perception task. It is plausible that due to this greater allocation of attentional resources to the STM task, any potential impact of rhythmic irregularity could have been overridden or mitigated.

The Dynamic Attending Theory (Jones, 1976; Jones and Boltz, 1989) posits that attention exhibits intrinsic oscillatory patterns that can be synchronized with external stimuli, allowing for the precise focusing of attention on specific temporal points. In our current study, the attention-demanding nature of the task might paradoxically have enhanced the allocation of processing resources and synchronization to the rhythm of the target sentences and suppressed processing of the introductory phrase primes. This effect of task demands on performance has been observed in other studies, as well. For instance, Morillon et al. (2014) demonstrated that active engagement of the motor system by tapping in synchrony with a reference rhythm improved the processing of on-beat targets and the suppression of off-beat distractors compared to a passive-listening condition. This emphasis on the relevant rhythm through motor engagement could potentially explain why certain STM tasks, which require active verbal rehearsal and repetition, might not show the anticipated effects of rhythmic irregularity. This could have been the case in our present study, and, perhaps, in studies such as Gorin (2020) and Lui and Wöstmann

(2022), where the attention-taxing nature of STM tasks might have concealed the potential effects of rhythmic disruption.

In conclusion, the divergent findings in performance across the mentioned studies underscore the crucial role of task nature in shaping outcomes. The disparities between the perception tasks and the STM tasks likely stem from the fundamental distinctions in the cognitive demands of these tasks. While passive comprehension tasks may not fully parallel the active engagement required for STM tasks, particularly those involving verbal repetition, the increased allocation of attentional resources during STM tasks might play a pivotal role in maintaining repetition accuracy.

4.3 Speed Manipulation Effects on STM and Self-Perceived Difficulty

One of the most intriguing findings in this study is the discrepancy between participants' actual repetition performance and their self-reported perceived difficulty. As part of the demographic and language background questionnaire following the completion of the two STM tasks and the FWL task, participants were asked to identify which condition they found the most challenging and to provide reasons for their choices. Interestingly, a considerable number of participants indicated that the slow condition was the most difficult to recall, followed by the fast condition, and then the normal condition. Among those who struggled with the slow condition, a majority specifically mentioned difficulty in recalling the beginning of the sentence by the time they reached the end. However, despite these perceptions, we did not uncover any significant differences in repetition accuracy across the three speed conditions. Furthermore, when examining repetition accuracy for different sentence positions, we consistently found that the initial portion had the highest recall accuracy, a pattern that remained consistent across all speed conditions.

This finding aligns with the results from Lui and Wöstmann's (2022) study, where the primary measure of participants' recall performance in a digit span task did not exhibit differences across conditions featuring rhythmically regular and irregular distractors. However, a secondary measure, involving confidence ratings, revealed higher levels of confidence when regular-rhythm distractors were presented during the memory encoding phase, as opposed to irregular-rhythm distractors. The impact of distractors in STM tasks on participants' metacognition has also been evident in other research studies, where participants' confidence levels decrease with higher levels of distractors during the retention and memory phases of the STM task (Beasman, Hanczakowski, & Jones, 2014; Kattner & Bryce, 2022). Taken together, these findings suggest that participants' metacognition can be influenced by distractors, and both the current study and Lui and Wöstmann's (2022) study demonstrate its sensitivity to temporal regularity. It is plausible that participants' detection of rhythmic irregularity at the onset of the target sentences triggers an increase in metacognitive awareness. This heightened metacognition could then lead to an enhancement in attentional control, which in turn might help mitigate the difficulties presented by the rhythmically irregular conditions. This cascade of processes suggests a dynamic interplay between metacognition, attention, and rhythmic perception in the context of STM tasks.

4.4. Some Post Hoc Findings

While the present study did not yield significant differences in repetition accuracy among the three speed conditions across the two STM tasks, an intriguing pattern emerged when considering participants' bilingual/multilingual status as additional factors. Particularly in the jabberwocky memory task, several noteworthy distinctions were observed among the speed conditions. Firstly, at the consonant level, repetition accuracy was found to be highest in the slow speed condition, followed by the normal and fast speed conditions. This indicates a potential advantage for participants in terms of repetition accuracy when the target sentences are presented at a slower pace. More intriguingly, bilingual/multilingual individuals exhibited better performance under slow speed conditions at the syllable level, compared to the normal speed conditions. In contrast, at the word level, monolingual participants demonstrated enhanced recall accuracy for whole words in the normal speed condition, compared to the slow speed condition. This finding suggests that the impact of speech rate disruption might differ based on individuals' language background. Monolingual individuals could potentially benefit from a constant speech rate for recalling complete words, while bilingual/multilingual individuals might excel in recalling individual syllables when the speech rate is slower.

The sample size in our current study, with 27 participants in the monolingual group and 13 in the bilingual/multilingual group, may be relatively small for drawing definitive conclusions. However, the nuanced patterns observed in our results provide valuable insights into an area of research that highlights the distinctions in phonological STM and language learning abilities between bilingual/multilingual and monolingual individuals. This finding resonates with previous research such as Yoo and Kaushanskaya's (2012), which demonstrated that while monolinguals excelled at easier levels of tasks like digit-span and nonword repetition, bilinguals performed better at more challenging levels. This study's results further illuminate the intricate interplay between linguistic background, STM performance, and their potential interactions with time representation. While preliminary, these findings suggest that specific task settings may yield varying benefits for monolingual and bilingual/multilingual individuals.

5 Conclusions

This study explored the effect of rhythmic disruptions and phonological STM for spoken sentences. It investigated the impact of manipulated rhythmic structures on verbal STM repetition tasks. The study delved into fundamental mechanisms governing the encoding of time in speech, and their potential influence on the representation of temporal order within STM. Previous studies indicated that rhythmic disruptions, when deviating from established internal rhythms, can negatively influence speech perception and comprehension (Kösem & al., 2018; Li & al., 2019). Our initial hypothesis was based upon these previous findings, proposing that altering the spoken speed in the middle of target sentences would compromise STM accuracy for them due to potential rhythmic disruption. However, this hypothesis was not fully substantiated by our results. Our study did not reveal significant differences in repetition accuracy among the three speed conditions across the two STM tasks. This suggests that the anticipated impact of rhythmic disruption on verbal STM might not be as pronounced as originally hypothesized.

However, the lack of significant differences in overall short-term memory (STM) performance across the three speed conditions can be understood through several experimental details. One key consideration is a likely lack of induced rhythmic irregularity. The study departed from the formerly employed methodology of Kösem et al. (2018), particularly in terms of the balance in length between the initial prime phrases and the subsequent target sentences. The difference in attentional demands between formerly studied speech comprehension and the present recall from STM may also have played a role. The alignment of participants' response speech rates with heard target speech rates hinted at the presence of a minor priming influence from the prime phrases.

The absence of genuine rhythmic irregularity could stem from two factors. Firstly, rhythmic irregularity is inherently relative, and the study's design, differing from that of Kösem et al. (2018), may not have created a salient contrast in rhythm perception between the prime phrases and target sentences. Secondly, a pronounced attentional focus on the target part of the sentences could override potential priming effects from introductory phrase speed. Furthermore, the nature of the task might have influenced performance irrespective of rhythmic irregularity. The divergence in performance outcomes between perception tasks and STM tasks can be attributed to the distinct cognitive demands of these tasks. The attention-taxing nature of STM tasks, including verbal repetition, might obscure the potential impacts of rhythmic disruptions, as observed across various speech STM studies.

Interestingly, delving further into participants' bilingual/multilingual status yielded intriguing insights. Particularly in the jabberwocky memory task, distinct patterns emerged among speed conditions and language backgrounds. Bilingual/multilingual individuals exhibited enhanced performance under slow speed conditions when scored at the syllable level, whereas monolinguals displayed improved recall accuracy for whole words in the normal speed condition. These differential effects based on linguistic background emphasize the interplay between language aptitude, speech rate manipulation, and STM performance. While the sample size of the study may temper conclusive findings, these patterns provide a foundation for future research endeavors delving into the interaction of linguistic factors with STM processes. Moreover, the allocation of attentional resources and metacognitive monitoring during STM tasks might be crucial for understanding the observed inconsistencies in the effects of rhythmic structures on STM. The central executive, functioning as a STM subsystem overseeing attention-control, might be influenced by metacognitive evaluations and has a role in strategic regulation. The study's assessment of metacognitive evaluations, including participants' perceived difficulty in recalling sentences presented with varying rhythmic regularities, introduces a layer of complexity to the interaction between attention, metacognition, and rhythmic disruptions. The observed discrepancies between perceived difficulty and actual performance add to the growing body of evidence suggesting that metacognition is shaped by various factors beyond mere performance.

In conclusion, while the manipulated speed conditions did not elicit significant disruptions in verbal STM accuracy as initially hypothesized, the study's outcomes underscore the potential roles of linguistic background, attentional allocation, and metacognition in shaping STM performance, inviting further investigation into these complex interactions.

References

Ahmed, F. (2022). Does rhythmic ability help with learning foreign words. Thesis, McMaster University.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence and J.T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory*. (Vol. 2, pp. 89–195). New York: Academic Press.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *In The Psychology of Learning and Motivation*. (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 302(1110), 311-324.

Baddeley, A.D. (2007). Working Memory, Thought and Action (Oxford: Oxford University Press.)

Baddeley, A. (2010). Working memory. *Current Biology*, 20(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Beaman, C. P., Hanczakowski, M., & Jones, D. M. (2014). The effects of distraction on metacognition and metacognition on distraction: Evidence from recognition memory. *Frontiers in Psychology*, 5, 1–13.

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P.-E., Giard, M.-H., & Bertrand, O. (2007). Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. *Journal of Neuroscience*, 27, 9252–9261.

Boersma, P. & Weenink, D. (2009). Praat: doing phonetics by computer (Version 6.2.22)

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language,* 31, 429–460.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review*, 106, 551–581.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, 55(4), 627–652. https://doi.org/10.1016/j.jml.2006.08.005

Brown, G. D., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107(1), 127.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing. A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* 11, 671–684.

DeBorba, E. E. (2020). The investigation of temporal order in language learning using behavioural tasks and MMN. MSc thesis, McMaster University.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection?. *Journal of Memory and Language*, 29(3), 336-360.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, 28(5), 887-898. https://doi.org/10.1037/0012-1649.28.5.887

Gilbert, R.A., Hitch, G.J., & Hartley, T. (2017). Temporal precision and the capacity of auditory–verbal short-term memory, *The Quarterly Journal of Experimental Psychology*, 70:12, 2403-2418, DOI: 10.1080/17470218.2016.1239749

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517.

Goldsmith, M. & Koriat, A. (2007). The strategic regulation of memory accuracy and informativeness. *Psychology of Learning and Motivation - Advances in Research and Theory*, 48, 1–60.

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, 220(1), 51+.

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

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

Gorin, S. (2020). The influence of rhythm on short-term memory for serial order. *Quarterly Journal of Experimental Psychology*, 73(12), 2071–2092. https://doi.org/10.1177/1747021820941358

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and biobehavioral reviews*, *86*, 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology*, 87, 135-178.

The jamovi project (2021). jamovi. (Version 2.0) [Computer Software]. Retrieved from https://www.jamovi.org.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323–335.

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

Kattner, F. & Bryce, D. (2022). Attentional control and metacognitive monitoring of the effects of different types of task-irrelevant sound on serial recall. *Journal of Experimental Psychology: Human Perception and Performance*, 48, 139–158.

Keitel, C., Thut, G., & Gross, J. (2017). Visual cortex responses reflect temporal structure of continuous quasi-rhythmic sensory stimulation. *NeuroImage*, 146, 58–70.

Kim, Y. J., Grabowecky, M., Paller, K. A., Muthu, K., & Suzuki, S. (2006). Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nature Neuroscience*, 10, 117–125.

Koriat, A. & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review,* 103, 490–517.

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology*, 28(18). https://doi.org/10.1016/j.cub.2018.07.023

Laasonen, M.; Virsu, V.; Oinonen, S.; Sandbacka, M.; Salakari, A.; Service, E. (2012). Phonological and sensory short-term memory are correlates and both affected in developmental dyslexia. *Reading and Writing 2012*, 25, 2247–2273.

Lakatos, P., Barczak, A., Neymotin, S. A., McGinnis, T., Ross, D., Javitt, D. C., & O'Connell, M. N. (2016). Global dynamics of selective attention and its lapses in primary auditory cortex. *Nature neuroscience*, *19*(12), 1707–1717. https://doi.org/10.1038/nn.4386

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53, 279–292.

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

Lakatos, P., Musacchia, G., O'Connell Monica, N., Falchier Arnaud, Y., Javitt Daniel, C., & Schroeder Charles, E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, 77, 750–761.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The leading sense: Supramodal control of neurophysiological context by attention. *Neuron*, 64, 419–430.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, 94(3), 1904–1911.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: new data and a model. *Advances in Research and Theory*, 1–48. https://doi.org/10.1016/s0079-7421(08)00001-7

Li, X., Shao, X., Xia, J., & Xu, X. (2019). The cognitive and neural oscillatory mechanisms underlying the facilitating effect of rhythm regularity on speech comprehension. *Journal of Neurolinguistics*, 49, 155–167. https://doi.org/10.1016/j.jneuroling.2018.05.004

Lui, T. K.-Y., & Wöstmann, M. (2022). Effects of temporally regular versus irregular distractors on goal-directed cognition and behavior. *Scientific Reports*, 12(1), 10020–10020. https://doi.org/10.1038/s41598-022-13211-3

Lopez Ricote, M. d. L. A. (2020). Serial order in language learning in bilinguals. MSc thesis, McMaster University.

Meyer, L., Sun, Y., & Martin, A. E. (2020). Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, 35(9), 1089–1099. https://doi.org/10.1080/23273798.2019.1693050

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

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926.

Papagno, C., Comi, A., Riva, M., Bizzi, A., Vernice, M., Casarotti, A., Fava, E. and Bello, L. (2017). Mapping the brain network of the phonological loop. *Human Brain Mapping*, 38: 3011-3024. https://doi.org/10.1002/hbm.23569

Paton, J.J., & Buonomano, D. V. (2018). The neural basis of timing: distributed mechanisms for diverse functions. *Neuron*, 98(4):687-705. doi: 10.1016/j.neuron.2018.03.045.

Service, E., & Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language*, 32(5), 608–623.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology Section A,* 45(1), 21-50.

Service, E., DeBorba, E., Lopez-Cormier, A., Horzum, M., & Pape, D. (2022). Short-term memory for auditory temporal patterns and meaningless sentences predicts learning of foreign word forms. *Brain Sciences*, 12(5), 549. MDPI AG. Retrieved from http://dx.doi.org/10.3390/brainsci12050549

Service, E, & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition?. *Applied Psycholinguistics*, 16(02), 155-172.

Teki, S., & Griffiths, T. D. (2014). Working memory for time intervals in auditory rhythmic sequences. *Frontier in Psychology*. 5:1329. doi:10.3389/fpsyg.2014.01329

Teki, S., & Griffiths, T. D. (2016). Brain bases of working memory for time intervals in rhythmic sequences. *Frontier in Neuroscience*. 10:239. doi: 10.3389/fnins.2016.00239

Teki, S., Grube, M., & Griffiths, T. D. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontier in Integrative Neuroscience*, 5:90. doi:10.3389/fnint.2011.00090

Tierney, A., White-Schwoch, T., MacLean, J., & Kraus, N. (2017). Individual differences in rhythm skills: links with neural consistency and linguistic ability. *Journal of Cognitive Neuroscience*, 29(5), 855–868. https://doi.org/10.1162/jocn_a_01092

Yoo, J., & Kaushanskaya, M. (2012). Phonological memory in bilinguals and monolinguals. *Memory & Cognition*, 40, 1314–1330. https://doi.org/10.3758/s13421-012-0237-x

APPENDIX A

Questionnaire A1: Demographic and Language Background Questionnaire used for the experiment.

Section 1: Demographic & Language Background

Gender: __________

Age: ________

Ethnicity: _____________

Education (eg. Undergraduate, Master's, PhD, Post-Doc, etc.): __________

Length of residence in English-speaking country/countries (years): ________

What is/are your native language(s)? ________

Is English a dominant language of yours? Meaning it is the language with which you have the greatest proficiency and/or uses the most, AND that you have completed at least all of high school in English. [YES/NO]

Please list the all languages you know in order of dominance (high to low), including English:

Language 1	Language 2	Language 3	Language 4	Language 5	Language 6

Please indicate your level of proficiency in each language for listening, speaking, reading, and writing. Indicate your proficiencies as B (beginner), I (intermediate), or A (advanced).

	Language 1	Language 2	Language 3	Language 4	Language 5	Language 6
Listening						
Speaking						
Reading						
Writing						

Section 2: Musical Background

Do you have any experience with any musical instrument training (including voice)? [YES/NO]

If you answered 'Yes' to the previous question, please indicate the type of musical instrument training (including voice) that you had or currently have experience with. Indicate the years of experience, and your abilities achieved in each as B (beginner), I (intermediate), or A (advanced). If you answered 'No' to the previous question, disregard the remainder of the questions.

	Instrument 1	Instrument 2	Instrument 3	Instrument 4	Instrument 5	Instrument 6
Type of Instrument						
Years of Experience						
Ability Achieved						

Section 3: Task Feedback

For the Jabberwocky & Urdu Tasks (Tasks 1 & 2), you might have noticed that sentences were presented at either slower, regular, or faster speed. Which of the condition(s) did you find the most difficult to recall? You may make multiple selections. Please identify in the comment section the reason(s) for your choice.

Slower: ______________________

Normal: ______________________

Faster: _______________________

About the same: ________________

Other comment: ________________________

For the Foreign Word Learning Task (Task 3), what strategies did you use to improve your memory for the English-Finnish word pairs?

____

Table A1: Descriptive statistics of participants' self-report ethnicity. N =40.

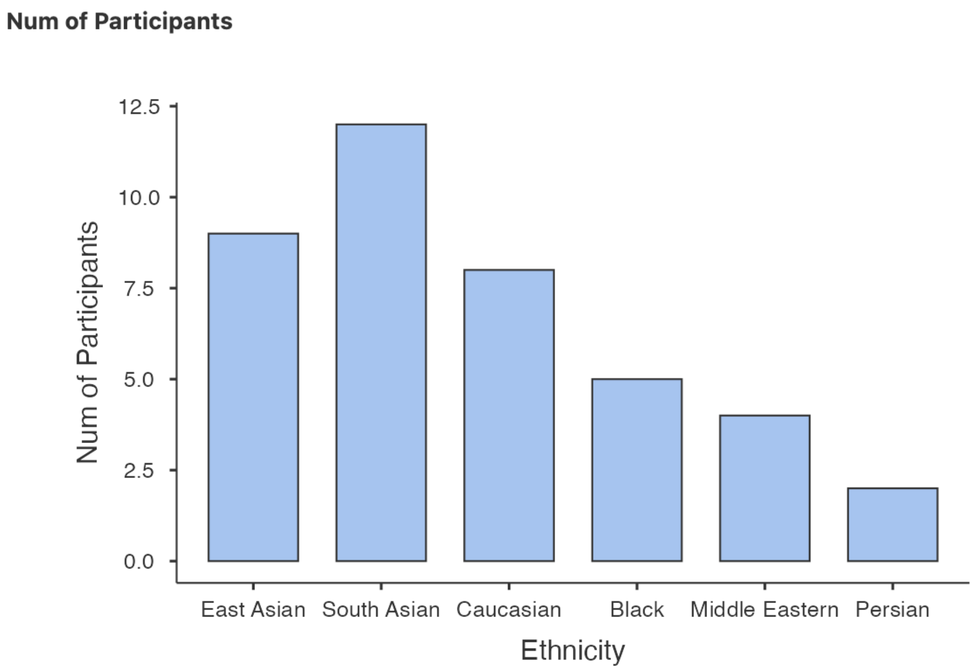


Table A2: Descriptive statistics for participants' self-report bi/multilingual status. Note that only participants who reported to be 'Advanced' in speaking in language(s) other than English were considered bilingual/multilinguals. N =40.

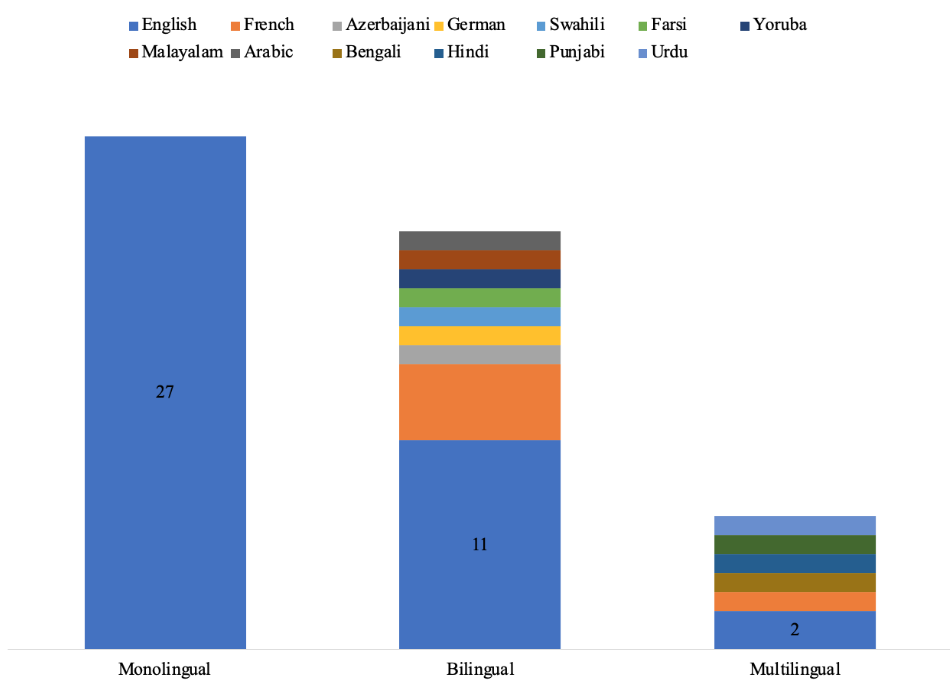
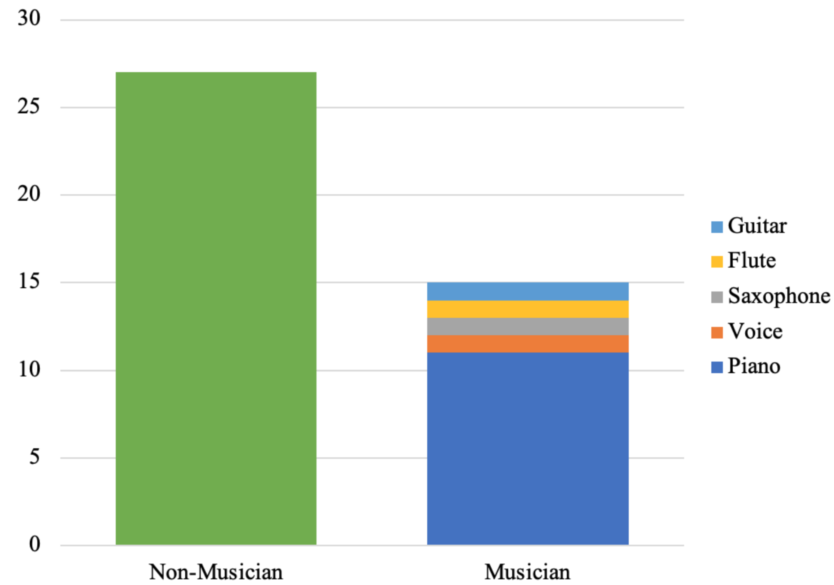


Table A3: Descriptive statistics for participants' self-report musicianship. Note that only participants who reported to be 'Advanced' in an instrument(s) were considered musicians. N =40.



APPENDIX B

List B1: Stimuli used in the jabberwocky memory task. Sentences 1, 4, 7, 10, 13 were used in the practice trials, the rest of the sentences were used in the experimental trials.

	Introductory Phrases	Jabberwocky Sentences
1.	The new rumor is that…	ROO MUTHTIN FANED IM GRASHNIT
2.	The new story is that…	WOE UPLING JIDED MO NAFF
3.	The new legend is that…	KAY HUSS EMENTED IM KILPH
4.	Yesterday Mabel said…	KADE GILPERNS KARNAYED ROO WUPS
5.	Yesterday Morris said…	ROO BEACHLORN SWEENS KAY NEDRIL
6.	Yesterday Franklin said…	WOE VOTION PRUSED JISH ROO ENTSIAN
7.	The name of the book is…	OH MOZZ TOMASHED MO FLOOKMUN
8.	The name of the song is…	OH DREG PRILED OT ROO BLANTIS
9.	The name of the game is…	PEEB TAFFERS SURFEWED ROO ZYPT
10.	Sam read a novel called…	KADE ADLOOT CONDLES KAY DUT
11.	Sam read a paper called…	ROO FLOB SUBSITTED IM OH GAST
12.	Sam read a report called…	KADE SMORKET NOOLED OH GUZDIN
13.	I was surprised she said…	KAY SIPPERT SEFT JISH ROO GAUM
14.	I was happy she said…	WOE HORIDGE JEPLES ROO VIGHT
15.	I was confused she said…	PEEB FLOTTERS PEFF CHAG LAUMSES
16.	Ben told me the tale of…	MO JELK MEMBLED ROO BLEARNATES
17.	Fred told me the tale of…	KADE TRANTOE DRIMES JISH OH PILK
18.	Kat told me the tale of…	ROO FLOGIN JANED IM OH KINTO
19.	The strange passenger yelled…	KAY DUCTORM PRICALLS IM THORK
20.	The old passenger yelled…	OH KALP SMIRRED MO KAPATTER
21.	The young passenger yelled…	OH JOWLER FRINED JISH OH VULKIT
22.	The happy children screamed…	WOE GEPLER SEVITS OT ROO QUAWN
23.	The happy students screamed…	PEEB ODDWAY FEETANED JISH HING
24.	The happy toddler screamed…	OH PONDLE BOYED IM ROO TROMPERS
25.	The writer quickly typed…	KAY BAXTOY SPLOPED IM OH WOAM

26.	The teacher quickly typed…	ROO NOLK WINTED ROO PRILES
27.	The student quickly typed…	KADE CALP GRAMUSHED MO FLOYANS
28.	The sign on the road reads…	MO SAMPRENG COLBS IM ROPOUTS
29.	The sign on the wall reads…	PEEB STORDINS RIMPLE JISH MIPTERS
30.	The sign on the board reads…	MO HEDDIES HAYVLED CHAG UFAL

List B2: List of jabberwocky function words and their corresponding English counterparts.

1. ROO - THE
2. OH - A
3. KADE - THESE
4. PEEB - SOME
5. MO - THOSE
6. WOE - THIS
7. KAY - THAT
8. IM - TO
9. CHAG - THEIR
10. OT - ON
11. JISH - WITH

List B3: Stimuli used in the Urdu memory task. Sentences 1, 4, 7, 10, 13 were used in the practice trials, the rest of the sentences were used in the experimental trials..

	Introductory Phrases	Urdu Sentences
1.	Hum keh rahe the ke… *We were saying that…*	Noor ne bartan dho liye *Noor has washed the dishes*
2.	Hum keh rahe the ke… *We were saying that…*	Uss ne ghar ki safaai ki hai *He/she cleaned the house*
3.	Hum keh rahe the ke… *We were saying that…*	Iss bachi ka naam ye he *This girl's name is this*
4.	Idhar ye hua ke… *Here, this happened that…*	Mein pani pi rahi thi *I (fem.) was drinking water*
5.	Idhar ye hua ke… *Here, this happened that…*	Mujhe sardi lag gayi *I caught a cold*
6.	Idhar ye hua ke… *Here, this happened that…*	Wo gaana ga rahi hai *She is singing*
7.	Kal aise hua ke… *Yesterday is was that…*	Zaid dukaan par geya *Zaid went to the store*
8.	Kal aise hua ke… *Yesterday is was that…*	Ye kitab meri he *This book is mine*

9.	Kal aise hua ke… *Yesterday is was that…*	Uss ne meri kursi li *He/she took my chair*
10.	Usko batana ke… *Tell him/her that…*	Maryam ne chirya pakri *Maryam caught a bird*
11.	Usko batana ke… *Tell him/her that…*	Usko nazla hai *He has a runny nose*
12.	Usko batana ke… *Tell him/her that…*	Abu ne gaari chalai thi *Dad drove a car*
13.	Kal ye baat hui ke… *Yesterday the thing that happened was that…*	Vo paise gin rahi thi *She was counting the money*
14.	Kal ye baat hui ke… *Yesterday the thing that happened was that…*	Vo tasveer baniyy gi *She will draw a picture*
15.	Kal ye baat hui ke… *Yesterday the thing that happened was that…*	Vo billi bhaag gayi *That cat run away*
16.	Mein udhar kal jaa ke… *I'm going there tomorrow and…*	Ye mera kalam he *This is my pen*
17.	Mein udhar kal jaa ke… *I'm going there tomorrow and…*	Mein chaand ko dekh ti hun *I (fem.) watch the moon*
18.	Mein udhar kal jaa ke… *I'm going there tomorrow and…*	Humari dosti hai *We are friends*
19.	Aise hua kal ke… *What happened yesterday was…*	Mein khaana kha rahi thi *I (fem.) was eating food*
20.	Aise hua kal ke… *What happened yesterday was…*	Vo angrezi bolta hai *He speaks English*
21.	Aise hua kal ke… *What happened yesterday was…*	Bache ludo jeet gaye *The kids won their game of Ludo*
22.	Vo chahta hai ab ke… *Now he wants that…*	Adan ka kaam theek ho *It should be Adan's joke*
23.	Vo chahta hai ab ke… *Now he wants that…*	Ami roti paka lein *Mom finishes making roti*
24.	Vo chahta hai ab ke… *Now he wants that…*	Ami aam bana lein *Mom makes mango*
25.	Ab mein chahti hun ke… *Now I want that…*	Sab kuch theek ho jaye *Everything is okay*
26.	Ab mein chahti hun ke… *Now I want that…*	Adil ki gaari ruk jaye *Adil's car stopped*
27.	Ab mein chahti hun ke…	Hum khane per jayein

	Now I want that…	*We will go out for dinner*
28.	Ab ye kal ho ga ke… *Now this will happen tomorrow that…*	Mein nahi jaa sakta *I can't go*
29.	Ab ye kal ho ga ke… *Now this will happen tomorrow that…*	Kaam khatam ho jai ga *Work will be finished*
30.	Ab ye kal ho ga ke… *Now this will happen tomorrow that…*	Aaj aam khareed lo *Buy mangoes today*

List B4: Stimuli used in the FWL task. List 1 was used for practice trials, List 2, 3, 4 were used for experimental trials.

List 1	**List 2**	**List 3**	**List 4**
CHIN - OMAISUUS	FLAG - UUPUMA	BOAT - EILINEN	CRAB - KOLLAASI
GANG - IKUINEN	KING - AITAUS	SILK - KUUSISTO	SCAR - AAPINEN
PIPE - KIIPELI	MEAT - UKAASI	ROCK - AAVISTUS	FEET - UUSINTO
LEAF - HAITUVA	SHOP - HAAVISTO	DOOR - OSOITIN	PLUM - HALAISTU
TANK - HUOMINEN	RICE - OUDOSTI	POND - KUVAAMO	NOSE - AJOITUS
WIFE - KALUUNA	WOOL - ILMAISU	HAND - KEITTAJA	GIRL - ETEINEN

APPENDIX C

Table C1: Post hoc comparisons of the effect of the interaction between speed conditions and bi/multilingual status at syllable level for the **jabberwocky memory task**. N = 40.

Post Hoc Comparisons - Speed ✻ Language

Comparison									
Speed	Language		Speed	Language	Mean Difference	SE	df	t	p
Normal	Monolingual	-	Normal	Bi/Multilingual	-0.05367	0.06533	36.0	-0.821	0.417
		-	Fast	Monolingual	0.01230	0.00772	36.0	1.593	0.120
		-	Fast	Bi/Multilingual	-0.05953	0.06812	36.0	-0.874	0.388
		-	Slow	Monolingual	0.01397	0.00868	36.0	1.609	0.116
		-	Slow	Bi/Multilingual	-0.08713	0.06327	36.0	-1.377	0.177
	Bi/Multilingual	-	Fast	Monolingual	0.06597	0.06607	36.0	0.998	0.325
		-	Fast	Bi/Multilingual	-0.00587	0.01516	36.0	-0.387	0.701
		-	Slow	Monolingual	0.06764	0.06480	36.0	1.044	0.304
		-	Slow	Bi/Multilingual	-0.03347	0.01704	36.0	-1.964	0.057
Fast	Monolingual	-	Fast	Bi/Multilingual	-0.07183	0.06882	36.0	-1.044	0.304
		-	Slow	Monolingual	0.00167	0.00936	36.0	0.178	0.859
		-	Slow	Bi/Multilingual	-0.09943	0.06403	36.0	-1.553	0.129
	Bi/Multilingual	-	Slow	Monolingual	0.07350	0.06761	36.0	1.087	0.284
		-	Slow	Bi/Multilingual	-0.02760	0.01837	36.0	-1.503	0.142
Slow	Monolingual	-	Slow	Bi/Multilingual	-0.10110	0.06272	36.0	-1.612	0.116

Table C2: Post hoc comparisons of the effect of the interaction between speed conditions and bi/multilingual status at word level for the **jabberwocky memory task**. N = 40.

Post Hoc Comparisons - Speed ✻ Language

Comparison									
Speed	Language		Speed	Language	Mean Difference	SE	df	t	p
Normal	Monolingual	-	Normal	Bi/Multilingual	-0.05468	0.06885	36.0	-0.794	0.432
		-	Fast	Monolingual	0.00666	0.00762	36.0	0.874	0.388
		-	Fast	Bi/Multilingual	-0.05468	0.07247	36.0	-0.755	0.455
		-	Slow	Monolingual	0.01946	0.00911	36.0	2.137	0.039
		-	Slow	Bi/Multilingual	-0.08283	0.06895	36.0	-1.201	0.237
	Bi/Multilingual	-	Fast	Monolingual	0.06133	0.06981	36.0	0.879	0.385
		-	Fast	Bi/Multilingual	9.09e-10	0.01496	36.0	6.08e-8	1.000
		-	Slow	Monolingual	0.07414	0.06888	36.0	1.076	0.289
		-	Slow	Bi/Multilingual	-0.02815	0.01787	36.0	-1.575	0.124
Fast	Monolingual	-	Fast	Bi/Multilingual	-0.06133	0.07338	36.0	-0.836	0.409
		-	Slow	Monolingual	0.01281	0.00909	36.0	1.408	0.168
		-	Slow	Bi/Multilingual	-0.08949	0.06991	36.0	-1.280	0.209
	Bi/Multilingual	-	Slow	Monolingual	0.07414	0.07249	36.0	1.023	0.313
		-	Slow	Bi/Multilingual	-0.02815	0.01784	36.0	-1.578	0.123
Slow	Monolingual	-	Slow	Bi/Multilingual	-0.10229	0.06898	36.0	-1.483	0.147