UNPACKING THE RELATIONSHIP BETWEEN TEMPORAL STRUCTURES AND LANGUAGE: AN EXAMINATION OF HOW VARIOUS RHYTHMIC TASKS RELATE TO LANGUAGE SKILLS

UNPACKING THE RELATIONSHIP BETWEEN TEMPORAL STRUCTURES AND LANGUAGE: AN EXAMINATION OF HOW VARIOUS RHYTHMIC TASKS RELATE TO LANGUAGE SKILLS

By TROY ALEXANDER FREIBURGER, B.A., GDIP., M.A.

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

McMaster University

McMaster University

MASTER OF SCIENCE (2024)

Hamilton, Ontario (Department of Linguistics and Languages)

TITLE: Unpacking the Relationship Between Temporal Structures and Language: An Examination of How Various Rhythmic Tasks Relate to Language Skills

AUTHOR: Troy Alexander Freiburger, B.A., GDIP., M.A. (McMaster University)

SUPERVISOR: Dr. Elisabet Service

NUMBER OF PAGES: x 62

Lay Abstract

While rhythm is most commonly associated with music, it is also important to our use of language. Rhythm is a complex phenomenon that involves a number of different cognitive processes and abilities. The most obvious of these are the ability to perceive and produce rhythm, and prior research has shown that these abilities vary both in and across individuals. People's capacity for remembering rhythms is another important aspect of people's rhythmical ability. However, it is unclear which aspects of rhythm are the most important to language. This experiment administered a number of different rhythmic tasks in order to determine which aspects of rhythm were most related to language. It was found that people's ability to remember rhythms was the greatest predictor of their ability to remember language-like sentence stimuli.

Abstract

The ability to represent time is essential for many aspects of human cognition and is especially critical for how people structure language and speech. Concerning language, people's ability to represent time relates to prosody (particularly, to the rhythmic aspects of language). Past research has shown a robust connection between people's rhythmic ability and their language skills; however, the nature of this relationship remains unclear. This unclarity, at least in part, results from the multidimensionality of people's rhythmic ability. The composite nature of rhythmic ability is evident from the broad variety of different rhythmic tasks that are found in the literature as well as the interindividual variation in people's performance on them. In order to clarify which aspects of people's rhythmic ability are related to language skills, in the present study a number of different rhythm tasks as well as a short-term memory task for non-sense sentences were administered. Participants' performance on the various rhythm tasks was used to predict performance on the non-sense sentence repetition task. It was found that the tapping memory task (but none of the various rhythmic production or perception tasks) was able to predict people's performance on the non-sense repetition task. In this task, participants were asked to tap from memory a sequence of auditorily presented short or long tones. The task showed a variable ability to predict performance depending on the size of the language units that were analyzed. Prediction was better for those units of language that included more rhythmic and temporal information. These findings suggest that the tapping memory task is distinct from the other rhythmic tasks administered in the present study (which either assessed people's ability to produce or perceive rhythms), and that the tapping memory task assesses a different rhythmical ability – which is here referred to as auditory sequential memory.

Acknowledgments

I would like to thank my supervisor Dr. Elisabet Service for all of her comments and support in the completion of this thesis. I would also like to thank my committee members, Dr. Cannon and Dr. Pape for all of their comments and feedback on this thesis. Another thanks to Christian Brodbeck for your time in acting as the external reviewer for this thesis. I would also like to acknowledge the other members of the Memory, Language, and Brain Lab, and Fiza Ahmed in particular, for their time, support, and assistance over the course of our degrees. I too would like to express my appreciation to Tiffany Vandusen for volunteering as an RA and to Jasmine Kwon for their time in volunteering to help with this thesis. I sincerely thank Charles-Etienne Benoit for all of his time and assistance with the BAASTA protocol, and for his support on the project. I also thank the BAASTA team for their collaboration and Nick Foster for his time in processing the BAASTA data. Finally, I would like to thank my family, friends, and my partner for their support during the completion of this thesis.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ASAP - Action Simulation for Auditory Prediction Hypothesis

BAASTA – Battery for the Assessment of Auditory Sensorimotor and Timing Abilities

BAT – Beat Alignment Test

MREB – McMaster Research and Ethics Board

PATH – Precise Auditory Timing Hypothesis

STM – Short-Term Memory

WM – Working Memory

Declaration of Academic Achievement

I declare that the work undertaken in this thesis is original was conducted by myself in consultation with my supervisor Dr. Elisabet Service. The thesis was written by myself and incorporated the feedback of my supervisor and the committee members Dr. Cannon and Dr. Pape.

Introduction

The brain's ability to represent time is an important aspect of cognition which, while shared by many different taxa for movement control, is especially important for humans due to its relation to language. The ability to represent and structure patterns in time is critical to (or else may be considered to be a predisposition for) our production and comprehension of language and speech. Within the empirical literature, the connection between temporal structures and language is studied apropos rhythm and prosody. Prosody, in linguistics, refers to the suprasegmental aspects of language manifested in speech such as loudness, pitch, and duration which operate beyond the level of a single phoneme (segment) and underlie rhythm. Prosody is often used to add emotional nuances to speech (such as an imperative tone), but also may affect the semantic meaning of an utterance – such as in English where a rising pitch accent is used to ask a question. Arguably the most important application of prosody to language is its use in speech segmentation.

One critical aspect of prosody that is seemingly ubiquitous among the languages of the world is rhythm. Rhythm is such a distinctive and pervasive element of language that a ‘rhythm class hypothesis’ has been suggested which maintains that groups (classes) of related languages can be identified on the basis of their rhythmical prosodic characteristics. One often studied distinction is whether languages are stress or syllable-timed, that is whether the timing between all syllables or main stressed syllables is constant in a language (Grabe & Low, 2002; Ramus et al., 2003; White & Mattys, 2007; Arvaniti & Rodriquez, 2013). It has been found that newborns are capable of discriminating between different languages according to their rhythmical elements. This has been interpreted as evidence that rhythm perception may play a role in language acquisition (Mehler et al., 1988; Nazzi et al., 1998; Soderstrom et al., 2003). Other studies have similarly found that prosodic and rhythmical cues are utilized in language discrimination and recognition in four (Bosch & Sebastián-Gallés, 1997) and seven (Gervain & Werker, 2013) month-old infants. A still more concrete example of rhythm's importance to language comes from its involvement in the speech segmentation of listeners, a process that turns the otherwise overlapping acoustic signals into recognizable (which is to say phonemically discrete) units in the listener's perception. A good example of this is how rhythmical cues play an important role in distinguishing whether one has heard 'recognize speech' as opposed to 'wreck a nice beech' (Myers et al., 2019). Prosodic perception has also been related to higher-order aspects of language, with prior research having found that people can use prosody to discover hierarchical syntactic structures in speech (Langus et al., 2012).

The empirical study of time and rhythm is often complicated by the vague sense with which the words are often used and the lack of an explicit definition for these terms. For the purposes of this thesis, a definition that has been offered in the literature will be used. The term *rhythm* will be used in a general sense to refer to a pattern of events in time (Wade, 2004; McAuley, 2010; Ravignani & Madison, 2017). This definition has the benefit of being broad enough to also account for empirical evidence which suggests that other species are capable of perceiving rhythm (Merchant & Honing, 2014; Ravignani & Madison, 2017; Honing et al., 2017). It is also critical to explicitly distinguish between *rhythmical ability* (which is to say the ability of an organism to perceive, produce, structure, or represent predictable patterns in time) and a rhythm as any temporal structure which can be perceived. The term temporal acuity will here be taken, in the broader sense, to denote the ability of an organism to accurately process, produce, remember and structure/represent phenomena in time.

Rhythmical Abilities in Relation to Language Skills

The empirical literature has quite robustly demonstrated a connection between rhythmical abilities (or a higher temporal acuity) and linguistic ability (Thomson & Goswami, 2008; Corriveau & Goswami, 2009; Strait, Hornickel & Kraus, 2011; Huss, Verney, Fosker, Mead & Goswami, 2011; Strait et al., 2012; Tierney & Kraus 2013; Tierney & Kraus 2014; Woodruff et al., 2014; Gordon et al., 2015; Tierney et al., 2017; Kachlicka et al., 2019). This connection nevertheless remains complicated by the fact that both rhythmical and linguistic abilities are themselves multi-faceted and further rely on a number of different cognitive abilities. Consequently, it is unclear exactly which aspects of rhythmical abilities relate to linguistic abilities. While some rather specific findings are available, such as the fact that reading ability is correlated with the ability to entrain to a metronome (Thomson and Goswami, 2008; Goswami, 2011; Woodruff et al., 2014), it is unclear precisely why this is the case and further, the extent to which phonological memory is a factor. Another specific finding, discussed by Tierney and Kraus, is that children with reading impairments also tend to have difficulties with timing integration in that their categorical perception of plosives is often abnormal (Tierney & Kraus, 2014).

More recent research has tried to identify which aspects of rhythmical abilities are related to people's language skills. Some researchers have suggested that a driving factor of rhythmical ability is temporal synchronization – understood as the ability to synchronize time between the motor and auditory/perceptual system, or else as the ability to time or align motor movements to sounds (Tierney et al., 2017; Rimmele et al., 2022; Luo & Lu, 2023). Tierney et al., (2017) suggested that performance on rhythm tasks can be reduced to two factors, rhythm sequencing, and rhythm synchronization. They took the former to reflect "the ability to perceive and reproduce rhythmic sequences" and the latter to reflect "the ability to tap consistently in time to stimuli, a process that relies on auditory motor timing integration" (Tierney et al., 13, 2017). While both of these factors were found to be related to general rhythmical ability, only rhythmical sequencing abilities were found to be related to language skills insofar as they were linked to verbal memory and reading (Tierney et al., 2017). In contrast, rhythmical synchronization abilities were only related to nonverbal auditory temporal processing (Tierney et al., 2017). However, the researchers' distinction between rhythmical sequencing and rhythmical synchronization does not preserve a clear distinction between people's ability to produce and perceive rhythms. This may pose a complication in light of other research which has shown that poor rhythm perception does not necessarily entail poor rhythm production (Dalla Bella et al., 2017). It was found that those with a great difficulty in perceiving rhythm were nonetheless able to produce an accurately timed beat to a regular rhythm (Dalla Bella et al., 2017). Moreover, a comparative case study has shown that even beat deaf individuals, while having the most extreme form of perceptual deficits for rhythmical stimuli, are nonetheless able to tap along to a beat and produce a steady rhythm (Begel et al., 2017).

Individual Variance in Rhythmic and Language Skills

While it is clear that rhythmical ability is in some way connected to people's capacity for language, to understand the nature of this relationship it is important to better understand the interindividual variance in people's rhythmical skills. It is quite clear from the literature that there is a large amount of individual variance in people's rhythmical abilities, or temporal acuity (Tierney et al., 2017; Dalla Bella et al., 2017; Bégel et al., 2017; Lizcano-Cortés et al., 2022; Fiveash et al., 2022). This variance has led some researchers to suggest that people should be categorically classified as either high or low auditory-motor synchronizers (Lizcano-Cortés et al., 2022). However, the picture is complicated by the fact that, in addition to the variance across the rhythmical ability of different individuals, there is also variance in the performance of individuals across different kinds of timing/rhythm tasks. Most broadly rhythm tasks are usually classified as either perceptual or productive (with examples including rhythm perception, beat alignment, on the one hand and isochronous tapping tasks on the other). However, many tasks involve both rhythmic perception and production. To further complicate the matter, other research has hypothesized that beat perception in auditory regions of the brain may, even in the absence of movement, still involve 'action-like processes' within the supplementary motor area. These are thought to facilitate temporal predictions of upcoming beat times (Cannon & Patel, 2021). In this case, it is unclear whether such a discrete distinction between production and perception is reflective of the neural pathways/processes involved in the completion of rhymical tasks. Researchers have also identified a third dimension of rhythmical ability, which has been referred to as ‘sequence memory-based rhythm processing’ (Tierney & Kraus, 2015; Fiveash et al., 2022). While prior research had shown a large degree of interindividual variance in people’s performance on rhythm production and perception tasks (Bégel et al., 2017), recent research has found that interindividual variance in people’s rhythmical ability also extends to their performance on sequence memory-based rhythm tasks (Fiveash et al., 2022). In particular, while some individuals performed well across all kinds of rhythm tasks (production, perception, and sequence memory-based) others showed a selective impairment in one of these areas (Fiveash et al., 2022). It may therefore be best to view rhythmical ability as a chimera composed of a number of different rhythmical, motor, memory, and or auditory abilities.

It has been suggested that attaining a better understanding of what drives the differences in individuals' rhythm skills may also explain some of the variance in individuals' capacity for using language (Tierney, et al., 2017). Due to the importance that rhythmical prosody has in speech and language, this is a promising avenue of research. A finding that accurate vowel perception and grammatical judgments were related to lower psychoacoustic thresholds, better auditory-motor integration, and more consistent frequency-following responses to sound (Kachlicka et al., 2019) led researchers to suggest that imprecise auditory perception could impair the perception of certain prosodic features marking phrase boundaries. Slight changes in duration and pitch are examples of such syntactic cues (Kachlicka et al., 2019). Therefore, the researchers argued that individual differences in auditory processing may be driving the inter-individual variance in both people's rhythmical and language abilities. A similarity can be seen between the words of a language and the notes of a musical score insofar as both are ‘chunked’ into smaller parts or phrases. A deficit in the ability to detect those cues that signal the end and beginning of phrases – which in language are often conveyed by prosody – may in part, therefore, explain some of this individual variance. While the ways in which such a deficit may relate to people’s ability to process rhythm are apparent, how such a deficit may relate (or not relate) to either people’s ability to produce rhythm or their performance on sequence memory-based tasks is less clear. In light of the multidimensionality of rhythmical ability, which is composed of at least 3 distinct components (Tierney & Kraus, 2015; Dalla Bella et al., 2017; Bonacina et al., 2019; Fiveash et al., 2022), it is critical for research to identify how the various aspects of rhythmical ability relate to people’s language skills. Moreover, it is important to determine if the observed interindividual variance of people’s performance across the different kinds of rhythmical tasks may serve to explain or else predict people's language skills.

Musicality, Synchronization, and Entrainment

The link between language skills and auditory-motor integration is particularly interesting in light of the connection between precise auditory-motor coupling and people's rhythmical ability (Tierney et al., 2017; Dalla Bella et al., 2017; Bégel et al., 2017; Kachlicka et al., 2019; Cannon & Patel, 2021; Lizcano-Cortés et al., 2022). Musical ability is also connected to precise auditory-motor coupling as musicians must have very precise timing and synchronization of their movements to the rhythms they play. Those with increased musical training have been seen to possess better timing abilities, such as an increased ability to synchronize to a rhythm (Bailey & Penhune, 2010; Krause et al., 2010). If precise auditory-motor coupling is also found to be important for language, then one might hypothesize that training other activities – such as the production of music – that require this skill may help facilitate people's fluency in the rhythmic and temporal aspects of language. Some support for this notion is suggested in findings that adults who exhibit a greater rhythmical ability are also more fluent in speech (Saito et al., 2018). Furthermore, adults with better auditory-motor integration have shown an increase in vowel detection and in the accuracy of their grammatical judgments (Kachlicka et al., 2019). A number of direct experimental findings have further shown that musical training relates to language learning outcomes and to the perception of foreign speech stimuli (Slevc & Miyake, 2006; Martinez-Montes et al., 2013; Swaminathan & Gopinath, 2013; Cooper & Ashley, 2017; Dittinger et al., 2018).

As a conclusion based on the recent work on rhythmical skills, researchers have proposed the *precise auditory timing hypothesis* (PATH) (Tierney & Kraus, 2014) as an explanation of why increased timing and rhythmical abilities (such as those exhibited by trained musicians) have been found to be correlated with performance on phonological and other language tasks. According to the PATH hypothesis, “entrainment practice is the core mechanism underlying enhanced phonological abilities in musicians” (Tierney & Kraus, 1, 2014). Tierney and Kraus defined entrainment as “the process of moving to a repeated auditory signal such that there is a consistent relationship between the timing of one's movements and the timing of sound onsets" (Tierney & Kraus, 1, 2014)[1]. The researchers argued that there is an overlap in the neural pathways responsible for processing time for musical linguistic stimuli and that such an overlap gives reason to suppose that cross-domain facilitation could occur. PATH offers the prediction that those with musical training in entraining to rhythms will exhibit increased phonological skills, and one may also predict a priming effect should a task requiring entrainment occur shortly before a language task.

Taken in this sense, entrainment is most closely related to the rhythmical synchronization factor which Tierney et al., (2017) identified in a later paper. However, and contrary to what would be predicted by PATH, rhythmical synchronization was not found to be related to language skills (although rhythmical sequencing was), but rather, it was only related to nonverbal auditory-temporal processing (Tierney et al., 2017). The researchers suggested that this result may have been due to the age of participants and suggested that a longitudinal study may be required. Another possible complication may have resulted from the variance in rhythmical synchronization abilities which are found

---

[1] This definition is very likely too strict for two reasons. Firstly, it requires that the auditory stimulus be repeated and therefore entrainment to a rhythm necessarily requires the rhythm be external – it requires that movements be aligned with the continued onsets of an auditory signal. Similarly, this definition does not count internal simulation, or else generation, of accurately timed continuations to an external auditory signal/rhythm after it has ended as an example of entrainment. In relation to the present study, only the tapping to music and tones would count as a demonstration of entrainment by this definition. More problematic, however, is that this definition requires – perhaps arbitrarily – that a physical movement take place on time with an external auditory signal. Thus, tapping one's foot to a song would count as entrainment but following along or simulating the rhythm or song mentally, without the presence of a motor action, would not count as entrainment by this definition. This definition does not cover the possibility that a rhythm is perceived or maintained internally in the absence of movement. This is precisely what the action simulation for auditory prediction (ASAP) hypothesis maintains (Cannon & Patel, 2021), and such mentally simulated motor actions should in all likelihood be considered as examples of entrainment when they are in time with a rhythm.

even between musicians. With respect to auditory synchronization, musicians may not be a homogenous group. Previous research has found clear differences in the performance of drummers, pianists, and singers on synchronization and temporal discrimination tasks (Krause et al., 2010).

Finally, and most critically to the present paper, it is not necessarily the case that the evidence demonstrating a connection between musical training and language learning (Slevc & Miyake, 2006; Martinez-Montes et al., 2013; Swaminathan & Gopinath, 2013; Cooper & Ashley, 2017; Dittinger et al., 2018), means that musical training will increase people's capacity for language. One alternative explanation may be that those abilities that are important to musicality utilize or are the same abilities that are important to language, and that to be a musician one must already have a high aptitude for these abilities. Further research is therefore required into the efficacy of training the timing/rhythmical skills which are relevant to both musical and language skills. It is unclear to what extent these abilities are themselves plastic, inasmuch as it is unclear what other cognitive capacities they are dependent upon. Therefore, it is unclear if entrainment practice will increase language capacity in the manner that PATH suggests. While a very promising avenue of research, it is possible, for example, that individual differences in auditory processing – such as some individuals having lower psychoacoustic thresholds – are driving many of these effects, as some researchers have suggested (Kachlicka et al., 2019).

Language Learning Impairments and Rhythmic or Temporal Ability

Little serves to so clearly illustrate the relationship between people's language capacity and temporal acuity as the literature concerning those with language learning impairments. The literature has consistently found that deficits in rhythmic perception and production are concomitant with language learning impairments such as dyslexia (Laasonen et al., 2001; Laasonen et al., 2002; Thomson et al., 2006; Thomson & Goswami, 2008; Goswami et al., 2010; Goswami, 2018; Canette et al., 2020; Fiveash et al., 2020), and in those with specific language impairments (Corriveau et al., 2007; Corriveau & Goswami, 2009; Cumming et al., 2015). Research has also shown a similar connection between dyslexia, phonological memory, and auditory awareness (Thomson et al., 2006; Laasonen et al., 2012). Deficits in other prosodic aspects of language, such as the perception of syllable stress, have also been linked to dyslexia (Goswami et al., 2013).

The research cited above shows a clear relationship between language learning deficits and rhythmical ability. These findings have led some researchers (Fujii & Wan, 2014; Canette et al., 2020) to propose that exposure to musical stimuli (such as regular rhythmic primes) may be helpful to people with language learning disabilities. The basis for this musical training program comes from the fact that those with musical training have been reported to be better at perceiving temporal regularities (Rimmele et al., 2022; Luo & Li, 2023), and that temporal regularities (such as prosodic rhythm) are an important aspect of language. Consequently, if it should be the case that those with language learning deficits such as dyslexia have a general impairment with the production and or perception of temporal regularities, then training such people with musical/rhythmical stimuli would likely be beneficial. Experimental support for this notion is found in prior research, which has demonstrated that musical training is beneficial to the phonological and reading skills of dyslexic children (Overy, 2000; Flaugnacco, 2015). However, and due to the real-world applications of such an idea, it is critical to more fully understand the nature of the relationship between language learning disabilities and deficits in rhythmical abilities before advocating for a musical training program for such a population.

Perhaps the greatest challenge in understanding this relationship comes from the great variety of timing tasks that are used to assess rhythmical abilities. In the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA) alone, there are eight distinct kinds of timing tasks. While previous research has shown that certain kinds of timing tasks, like rhythmic tapping of sequences from memory, are associated with both language learning and memory (Service et al., 2022), it remains unclear which of the other broad variety of timing tasks from the literature are similarly related to people's language capacity. Because temporal acuity is not a unitary ability but is instead reliant upon different cognitive abilities that show a great degree of individual variance[2], further research is required in order to determine which kinds of timing tasks (and therefore which facets of temporal acuity) are reliably related to people's language capacity.

Phonological Short-Term Memory

---

[2] As was discussed in the earlier subsection: Individual Variance in Rhythmic and Language Skills.

Of the many different cognitive abilities that are important to people’s language capacity, phonological short-term memory (STM) can be argued to be among the most critical. The ability to accurately store the phonological characteristics of a speech signal offers a decisive advantage to language learning and processing. A key advance in the literature on STM was the seminal proposal of working memory by Baddeley and Hitch (1974). Working memory was then conceptualized as a kind of temporary information store that served to explain how information was transferred from the short-term to long-term memory systems (Baddeley & Hitch, 1974). In its original version, their theory of working memory (WM) was multidimensional. WM was composed of the following three parts: a) the central executive, b) the phonological loop, and c) the visuospatial sketchpad (Baddeley & Hitch, 1974). In later work, a fourth component was added, the episodic buffer (Baddeley, 2000). The central executive (titularly named) served as the central component of this theory and acted to control, direct, and consolidate the information held in STM (Baddeley & Hitch, 1974; Baddeley, 1983; Baddeley, 2000; Baddeley, 2007; Baddeley, 2010). The other three components served as subsidiary systems that ‘held’ a specific kind of information according to their name. Respectively, the phonological loop held representations of phonological/auditory stimuli, the visuospatial sketchpad of visual stimuli, and the later introduced episodic buffer held representations of whole experienced events (Baddeley & Hitch, 1974; Baddeley, 1983; Baddeley, 2000; Baddeley, 2007; Baddeley, 2010).

The phonological loop is the most critical of these sub-systems to the present paper due to the phonological/auditory nature of language and rhythm. According to

Baddeley, “The loop is assumed to comprise two components, a phonological input store and an articulatory rehearsal process involving subvocal speech” (Baddeley, p. 78, 1983). The role as much as the utility of this rehearsal process to memory is made apparent when one is tasked to remember something like a phone number or an address for immediate recall. In this case, one very often consciously plays back (rehearses) the number or addresses it repeatedly in one’s mind. Baddeley and Hitch’s theory of working memory offers a good explanation of this occurrence, whereby the central executive, in order to retain the information, consciously directs attention to the phonological loop and rehearses a representation of the auditory stimulus. While it is clear that the phonological loop is highly relevant to language, the subsystem's involvement in the retention of non-linguistic auditory stimuli (such as the beats/tones that make up a rhythm) is less clear.

The literature has shown a clear relationship between people’s phonological STM capacity and language skills such as people’s reading ability (Brady, 1986; Rapala & Brady, 1990; Adams & Gathercole, 1995; Conti-Ramsden & Durkin, 2007; de Abreu et al., 2011; Vulchanova et al., 2014). However, while early research has shown that phonetic processes are linked to verbal memory span, they were not found to be related to non-verbal memory (Brady, 1986). It is therefore unclear if the phonological loop is domain-specific (a module) to language-like stimuli, or if it may be extended/used for the retention of non-linguistic auditory information. However, it has been suggested that the phonological loop may be relevant to people’s memory for music at short intervals (Baddeley, 2017). Obtaining a clearer understanding of how the phonological loop is or is not related to the retention of information for non-linguistic stimuli may be of importance to understanding how rhythmical abilities and language skills are related. Moreover, a relationship between deficits in phonological STM and language learning impairments is seen to exist within the literature (Gathercole & Baddeley, 1989; Gathercole & Baddeley, 1990; Rapala & Brady, 1990; Montgomery, 1995; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Conti-Ramsden & Durkin, 2007). This relationship mirrors that discussed in the last section, that between language learning impairments and rhythmical abilities. One possible explanation for this may be that language learning and rhythmical tasks both utilize the phonological loop in an important way.

The Present Research Goal and Purpose

While the literature suggests that some relationship exists between temporal acuity and people's language skills (see subsections **Rhythmical Abilities in Relation to Language Skills** and **Language Learning Impairments and Rhythmic or Temporal Ability**), the nature of this relationship remains unclear. A broad range of different timing tasks has been used to assess people's rhythmical ability. Therefore, due to both the inter-individual variance in people's performance across these tasks, and to the composite nature of this cognitive ability itself (see subsection **Individual Variance in Rhythmic and Language Skills**), it is unclear how these tasks relate to either one another or to people’s language skills. Moreover, researchers differ in their views of the architecture of rhythmic ability, with some breaking it down into rhythmical synchronization and rhythmical sequencing factors (Tierney et al., 2017), while others favor a perception versus production distinction (Dalla Bella et al., 2017).

The purpose of the present thesis is to better clarify how a previously used rhythm STM memory task with tapping responses to binary sequences of long and short stimuli (Service & al., 2022) relates to various kinds of timing tasks and to people's phonological STM and word learning skills. It is unclear how sequence tapping tasks (such as that used in the present study) relate to rhythm perception and production tasks such as those administered in the BAASTA protocol, or if they are instead more reliant upon another cognitive skill entirely – such as STM for order. In either case, whether such tasks relate directly to people’s language skills or through mediation of other cognitive processes is also unclear. The goal of the present study is to examine how closely related the various aspects/components of people's temporal acuity are (as demonstrated by participants' performance across a number of different timing/rhythm tasks), and which aspects of rhythmic ability are most relevant to people’s language skills. In a more specific sense, the goal of this thesis is to identify which timing tasks are the most predictive of people's language skills.

Methods

Participants

The participants for this study were drawn from either the Linguistics Research Participant System or the Psychology, Neuroscience, and Behavioral Research Participant System at McMaster University, both administered using SONA software. A total of 67 participants participated in this study (8 males). Three participants were later excluded as outliers. Anyone with prior knowledge of the Finnish language was not eligible to participate in this study. Only those participants with self-reported normal hearing and for whom English was a dominant language were selected for participation in the study. Because at the time of the experiment two other research studies were also being run which utilized similar stimuli, those SONA participants that had participated in either of those studies were excluded. A consent form was signed in person prior to participation in the study. The study was cleared by the McMaster Research Ethics Board (MREB).

Tasks and Stimuli

BAASTA Tasks

The following five tasks were drawn from the BAASTA battery: 1) Unpaced tapping, 2) Beat Alignment Test, 3) Paced tapping (tones), 4) Anisochrony detection (music), and 5) Paced tapping (music). Three of these tasks (1, 3, and 5) are production tasks that assess a participant's ability to produce a rhythmic/temporal structure by tapping with their finger. Two of these tasks (2, and 4) are perception tasks which assess a participant's ability to discern whether a temporal structure is rhythmically regular or not. The Beat Alignment Task was performed in full (for a total of 72 experimental stimuli) and included slow, medium, and fast rhythms. The Paced tapping task was also performed for tones presented at slow, medium, and fast frequencies at the same pitch.

After a brief calibration of tapping registration and audio equipment, the BAASTA tasks were administered in the following order: 1) Unpaced tapping, 2) Beat Alignment Test, 3) Paced tapping, 4) Anisochrony detection, and 5) Paced tapping. This order was chosen and preserved so that participants alternated between production and perception tasks in turn. The experimenter administered the BAASTA by navigating through the various tasks on the tablet. Prior to each task, participants received written instructions on the tablet and completed practice trials for the task. An additional training phase (in which examples were presented) was completed for the Beat Alignment Test and Anisochrony detection tasks as per the BAASTA protocols. For a further description of the nature of the BAASTA tasks and stimuli included in this experiment, see Dalla Bella et al., (2017).

1) Unpaced Tapping (Dalla Bella et al., 2017)

For this production task, participants were instructed to generate and maintain a regular rhythm at their own pace by tapping their finger on a Galaxy tablet for a period of 60 seconds. Because this task assessed the participants' ability to self-generate a regular rhythm, no stimulus was presented; instead, the regularity of the tapped beat was measured.

2) Beat Alignement Test (BAT) (Dalla Bella et al., 2017)

Four different samples of computer-generated musical riffs were used as the stimuli in this task. Two of these samples were derived from Bach’s ‘Badinerie’ and the other two came from Rossini’s ‘William Tell Overture’. These were played at three different tempos ranging from fast to slow based on the interval between the beats of the piece. The fast tempo had an inter-beat interval of 450 ms, the regular tempo an interval of 600 ms, and the slow tempo an interval of 750 ms. All generated samples were composed of 20 beats played at the respective intervals where a beat was equal to a quarter note. An isochronous sequence of tones was then superimposed over the computer-generated musical piece which either aligned with the beat or was unaligned with the piece. This superimposed sequence was a triangle timbre that began at the 7th beat and continued until the end of the stimulus. There were two ways in which the superimposed sequence could be unaligned with the musical piece. In one case, the 'relative phase' was changed such that the superimposed tones were presented either 33% earlier or later than the inter-beat interval while maintaining the same tempo. For the other kind of unalignment, the superimposed tones were played at a tempo that was either 10% faster or slower than the musical piece. Participants were asked whether the superimposed sequence was aligned or unaligned with the beat of the musical piece.

3) Paced Tapping (Tones) (Dalla Bella et al., 2017)

This task assessed the ability of participants to follow and maintain a steady beat by tapping along to a regular tone sequence. The stimuli for this task were an isochronous series of 60 piano tones played at an interval of either 450, 600, or 750 ms and at a frequency of 1319 Hz. Participants were instructed to tap along to the tones as closely as they were able.

4) Anisochrony Detection (Music) (Dalla Bella et al., 2017)

This task assessed the ability of participants to detect an irregularity in time – in this case, a shift in the time at which one beat of a musical stimulus is presented. The musical stimulus for this task was a computer-generated 2-bar excerpt of Bach's 'Badinerie' played at the timbre of a piano at an interval of 600 ms. In the case of irregular (anisochronous) stimuli, a 'local time-shift' occurred at the 5th beat of the excerpt which ranged up to as much as a 30% displacement or 180 ms. The MLP algorithm of BAASTA controlled the presentation of these anisochronous tones – see Dalla Bella et al., (2017) for a further description.

5) Paced Tapping (Music) (Dalla Bella et al., 2017)

This task was similar to the paced tapping to tones task described above but instead utilized musical stimuli. Participants were instructed to tap along to the beat of a computer generated excerpt from a musical piece (again either from Bach's 'Badinerie' or Rossini's 'William Tell Overture'). The piece was played for 64 beats at an interval of 600ms (Dalla Bella et al., 2017).

STM for Sentences Consisting of Pseudowords

This task was used to assess participant's phonological short-term memory for non-sense language stimuli. The non-sense (jabberwocky) sentences used were constructed using the phonological and syntactic features of Canadian English. Participants were instructed to listen attentively to the non-sense sentences as they would later have to repeat them out loud. After the presentation of a fixation mark ('+') the word "listen" appeared on the screen and the Jabberwocky sentence was presented. After the conclusion of the stimuli, a fixation mark ('+') reappeared followed by the word "repeat". Participants received three practice trials after which they were presented with experimental trials (no sentences were repeated). This task was administered via a local PsychoPy script on a laptop computer.

A total of 20 non-sense (jabberwocky) sentences were used as trial stimuli for this study. These sentences were drawn from an existing bank of 40 jabberwocky sentences from Dr. Service's Language, Memory & Brain Lab. The stimuli were recorded using Audacity software. Sentences contained between 5 and 6 "non-words" and contained between 7 and 9 syllables. Unlike some jabberwocky sentence stimuli that are used in the literature, those used in this study had both the content and function words replaced with nonsense words. See Figure 5 for the complete list of all the jabberwocky trial stimuli used in this study.

Foreign Word Learning

Finnish was used as the foreign language for this experiment's word-learning task. The Finnish language offered a number of benefits to the present study, foremost of which was how distinct its phonology is from English and many other languages. This task assessed the ability of participants to learn and remember the meaning of words in a foreign language. Consequently, the task required participants to remember both the semantic meaning as well as the phonological features of a given word. The task was run using a local PsychoPy script.

Participants were verbally and visually instructed that their task was to learn the Finnish words as they would later be asked to recall them when cued with the English words. In this task, the words were presented in isolation (i.e., not within a sentence context) first a familiar English word, and then a matched word in Finnish. Prior to the presentation of the stimulus, participants were presented with a fixation mark (+) which changed to a fixation cross (X) while stimuli were being presented. A word in English was heard alongside this fixation cross and after one second the word was pronounced in

Finnish. The next word would then be presented after a 2.5 s delay. In this task, participants were presented with four blocks of six word pairs (for a total of 24 different word pairs). Participants were first presented with a practice block that contained six English words each matched with a Finnish word. After all six of the word pairs had been presented, participants were given a recollection task in which they would hear one of the six words in English and would have to recall and phonate the corresponding Finnish word. Once the participants were familiar with the task, the trials began. Participants completed three blocks of word-learning trials, each of which contained six words (for a total of 18 Finnish words tested six at a time). Within each block, all six of the words were presented sequentially and then participants were given the recall task. This was repeated five times for each trial block with the participants being exposed to each Finnish word for a total of five times. In order to avoid any order effects, the stimuli within each block were split such that half of the participants received them in a different order (see Table 1 for a visual representation of the counterbalanced task structure).

The stimuli for this task were taken from a set of priorly used stimuli from Dr. Service's Language, Memory & Brain Lab. Stimuli were arranged into 4 blocks each composed of 6 words, with the first serving as practice stimuli. These stimuli were recorded using Audacity. In addition to the Finnish words, the equivalent English word/translation was used as stimuli during the learning phase of each trial block. See Table 2 for the complete list of the English and Finnish word stimuli used in this study.

There were two different between-subjects conditions for the foreign word learning task. In the first of these conditions ("Sequential") participants were asked to recall the words in the same order that they were presented in the learning phase. For example, if during the learning phase, the order of presentation had been: flag, king, meat, shop, rice, wool, the recall phase would preserve this same order. Although the words themselves were presented in a meaningless list format, this condition allowed for participants to use the serial order in which the words were presented as a cue to learning. In the second condition ("Random"), participants were asked to recall the words in a pseudorandom order (shown in Figure 1).

Tapping Memory Task

The tapping memory task assessed participants' phonological short-term memory for non-linguistic temporally structured tonal stimuli. Participants were presented with a series of long and short tone tokens and were asked to repeat the sequence back in the same order. After three practice sequences had been completed, participants were presented with ten sequences of seven tokens of the 527 Hz tone. Each sequence had three of one length and two of the other. The long tones lasted 800 ms and the short tones lasted 200 ms with 200-ms inter-stimulus-intervals (ISIs). Three practice sequences were first presented. After them, 10 trial tone sequences were presented. For a visual representation of the tone sequences used see Figure 2.

The word "listen" appeared on the screen prior to the presentation of the sequence of short and long tokens. While the tone sequence was playing "listen" was replaced with a fixation mark ('+'). When the tone had concluded, the word "repeat" appeared, indicating that participants were to reproduce the tonal sequence using the spacebar. This procedure was repeated three times during the practice trials and then ten times during the experimental phase. The tapping memory task was administered on a laptop computer through an online PsychoPy script using the website Pavlovia interface.

Materials

A Samsung Galaxy Tablet A 2016, Android version 8.1.0, was used to administer the BAASTA tasks. BAASTA version 0.6.0 was downloaded onto this tablet. Audacity software for sound processing was run on a MacOS Monterey Version 12.7.4; iMac Retina 5k 27-inch late 2015 and was used to record participants' responses for the Finnish word learning and Jabberwocky repetition tasks. Responses were recorded using a Rode Microphone NT-USB Versatile Studio-Quality USB Microphone. A pop filter was used and the microphone was positioned slightly to the left of the participant. Both the local as well as the online PsychoPy scripts were run on a HP Laptop, 15-dw3xxx, x64-based PC running Microsoft Windows 11 Home. The spacebar key of this laptop was used to record the responses of all participants on the tapping memory task. The online PsychoPy script was run using the Pavlovia interface.

Procedure

For this experiment, participants completed the following four different tasks: a foreign word learning task, a subset of 5 tasks from the BAASTA protocol, a tapping recall task, and a jabberwocky memory task. These tasks were subdivided into two blocks with one block being the BAASTA tasks and the other block encompassing the foreign word learning, tapping, and jabberwocky memory tasks. Each of these blocks lasted roughly 25-30 minutes and participants were prompted to take a short break in between the blocks. Which block a participant first received was randomized in order to avoid any order effects. Within the BAASTA block, the five tasks were administered in the order outlined above (see the subsection Tasks and Stimuli), while within the other block the order was always as follows: 1) the foreign word learning task, 2) the Jabberwocky memory task, and 3) the tapping memory task.

After a consent form was signed, participants were asked about their musical background – whether or not they would self-identify as musical. They then moved on to one of the two blocks, receiving both written and verbal instructions before each task. Participants were instructed to ask questions at any point and were told that they may ask for a break at any point in the experiment. The experiment occurred in a quiet lab space in the presence of a researcher. It was necessary for an experimenter to be present in the room with the participant due to the nature of some of the tasks (i.e., the BAASTA requires a researcher to administer). The experimenter sat beside (but out of the sightline of) the participant and remained silent throughout the trials. The auditory stimuli were presented at a fixed volume through the computer’s speakers. After completing the experiment, participants were given a short debrief and shown out.

Data Processing

Tapping Memory Task

A script automatically recorded participants' responses to the tapping memory task into an Excel sheet. Participants' raw data was then turned into a proportion score in a three-step process. First, the responses were converted into milliseconds. Then, those responses over 350ms were labeled “long” with those under 350ms being labeled “short.” A proportion score was then generated as the proportion of the first 7 recorded taps that matched the corresponding element of the correct sequence (see Figure 2). Following the precedent of prior experiments, any additional taps beyond the count of 7 were disregarded.

STM for Sentences Consisting of Pseudowords

Participants' responses for the non-sense sentence repetition task were scored by a single experimenter. Participants were scored according to the number of syllables, consonants, and words that they correctly repeated from the nonsense sentence stimuli. Proportion scores were then generated for each participant.

BAASTA Tasks

The data for the BAASTA tasks was processed by the BAASTA data processing team. An Excel spreadsheet was provided containing the raw scores of participants. These were then converted into Z-scores using the norm tables published in Dalla Bella et al. (2024).

Foreign Word Learning

The foreign word learning data has not yet been processed and scored. This is due to the lack of native Finnish speakers to assist in the scoring. This data will be scored and incorporated at another date.

Results

Descriptive Statistics

Tapping Memory Task

The descriptive statistics of participants' performance on the Tapping Memory Task are presented in the following table. Scores are given as a proportion of correctly identified taps (long or short). The data of one participant was removed as an outlier as their performance on the Tapping Memory Task fell outside of two and a half standard deviations from the mean.

Table 1: Descriptives for Tapping Memory Task: Proportion of Correct Taps

	Mean	Median	Number	Standard Deviation	Minimum	Maximum
Tapping Memory Task	0.751	0.760	62	0.081	0.570	0.910

Non-Sense Sentence Repetition Task

The following table provides the descriptive statistics for participants' performance on the non-sense sentence repetition task. Scores indicate the proportion of correctly identified units of language at the consonant, syllable, and word levels. The data of two participants were removed as outliers since their scores at one or more of these levels of analysis exceeded two and a half standard deviations from the mean.

Table 2: Descriptives for Non-Sense Sentences Task: Proportion of Correct Language Units.

	Mean	Median	Number	Standard Deviation	Minimum	Maximum
Consonant	0.663	0.660	62	0.110	0.43	0.920
Syllable	0.755	0.760	62	0.104	0.50	0.970
Word	0.407	0.400	62	0.123	0.19	0.700

BAASTA Tasks

The recorded BAASTA responses were scored by the BAASTA team (Dalla Bella & al., 2024). In the following table, the descriptive statistics for the various BAASTA tasks are provided. These descriptives come from the participants' raw scores, as opposed to the converted Z-scores[3] which were used for all the subsequent analyses. For the BAT, these raw scores represented d-prime values. Raw scores for the mean of tapping to pieces, metronome synchronization accuracy, and the unpaced tapping tasks all represent the collected variance of participants' inter-tap intervals. Raw scores in the anisochrony detection task give an indication of the threshold at which an anisochronous beat was able to be detected by a participant. Although some participants here exceed the two-and-a-half standard deviations from the mean threshold, no additional outliers were removed. This is because the BAASTA test processing team has already cleaned up and processed this data into a workable form (a process that may have already included excluding the greatest outliers). The synchronization accuracy scores to the following three tasks: a) paced tapping to a 450-ms inter-onset interval (IOI), b) paced tapping to a 600-ms IOI, and c) paced tapping to a 750-ms IOI, were averaged into one variable hereafter referred to as 'Metronome Synchronization Accuracy'. Similarly, participants' synchronization accuracy scores to paced tapping to the two musical pieces (Bach's 'Badinerie' and

---

[3] From these, Z scores were derived from the published BAASTA norm tables (Dalla Bella et al., 2024) using the means for all age groups.

Rossini's 'William Tell Overture') were averaged into one variable hereafter referred to as 'Music Synchronization Accuracy'.

Table 3: Descriptives for BAASTA Tasks: (Raw Scores).

	Mean	Median	Number	Standard Deviation	Minimum	Maximum
Beat Alignment Test	3.510	3.820	63	0.944	1.140	4.420
Music Synchronization Accuracy	0.055	0.040	63	0.059	0.022	0.380
Metronome Synchronization Accuracy	0.042	0.042	63	0.009	0.028	0.077
Unpaced Tapping	0.049	0.049	64	0.023	0.026	0.132
Anisochrony Detection	10.20	8.620	59	6.510	3.200	33.500

Correlational Analysis

Table 4 provides the results of an omnibus correlational analysis performed for all the various tasks. For this analysis, Spearman's ρ correlation coefficients were calculated because several of the BAASTA tasks were not normally distributed.

Significant correlations were observed between the beat alignment test and each of the other BAASTA tasks. It was the only BAASTA task to correlate significantly with all the others. Significant correlations were found between people's music synchronization accuracy, a production task, and: a) the beat alignment test, a perception task, $\rho(60) = -0.320$, $p = 0.010$; as well as b) their metronome synchronization accuracy, a production task, $\rho(60) = 0.461$, $p < 0.001$. Significant correlations were observed between the metronome synchronization accuracy, a production task, and: a) the beat alignment test, a perception test, $\rho(60) = 0.460$, $p < 0.001$; b) their music synchronization accuracy, a production task, $\rho(60) = 0.461$, $p < 0.001$; and c) unpaced tapping, a production task, $\rho(61) = 0.475$, $p < 0.001$. Significant correlations were further found between the unpaced tapping task, a production task, and: a) the beat alignment test, a perception task, $\rho(61) = -0.400$, $p = 0.001$; and b) people's metronome synchronization accuracy, a production task, $\rho(61) = 0.475$, $p < 0.001$. The anisochrony detection task, a perception task, was significantly correlated only with the beat alignment test, a perception task $\rho(56) = -0.290$, $p = 0.026$.

The connection between rhythm variables and verbal short-term memory was of special interest. Significant correlations were found within the non-sense sentence repetition task across each level of analysis, with participants' proportion of correctly repeated consonants, syllables, and words each being strongly correlated with one another at a $p < .001$. Significant correlations were replicated between the tapping memory task and: a) the proportion of correctly repeated syllables for non-sense sentences, $\rho(58) = 0.319$, $p = 0.013$; b) the proportion of correctly repeated words for non-sense sentences, $\rho(58) = 0.321$, $p = 0.012$ although significance was not quite reached for correctly repeated consonants, $\rho(58) = 0.253$, $p = 0.051$. The tapping memory task further correlated with the perceptual beat alignment test $\rho(59) = 0.302$, $p = 0.018$.

Table 4: Omnibus Correlation Matrix for All Tasks.

		1	2	3	4	5	6	7	8	9
1: Tapping Memory Task	Spearman's rho	—								
	df	—								
	p-value	—								
2: JABB Consonant Proportion	Spearman's rho	0.253	—							
	df	58	—							
	p-value	0.051	—							
3: JABB Syllable Proportion	Spearman's rho	0.319	0.915	—						
	df	58	60	—						

	p-value	**0.013***	**< .001****	—						
4: JABB Word Proportion	Spearman's rho	0.321	0.920	0.845	—					
	df	58	60	60	—					
	p-value	**0.012***	**< .001****	**< .001****	—					
5: Beat Alignment Test	Spearman's rho	0.302	0.137	0.149	0.200	—				
	df	59	59	59	59	—				
	p-value	**0.018***	0.293	0.251	0.123	—				
6: Music Synchronization Accuracy	Spearman's rho	-0.200	-0.120	-0.170	-0.190	-0.320	—			
	df	59	59	59	59	60	—			
	p-value	0.114	0.377	0.190	0.145	**0.010***	—			
7: Metronome Synchronization Accuracy	Spearman's rho	-0.150	-0.040	-0.120	-0.030	-0.460	0.461	—		
	df	59	59	59	59	60	60	—		
	p-value	0.252	0.777	0.367	0.811	**< .001****	**< .001****	—		
8: Unpaced Tapping	Spearman's rho	-0.010	0.032	-0.050	0.004	-0.400	0.224	0.475	—	
	df	60	60	60	60	61	61	61	—	
	p-value	0.940	0.805	0.705	0.974	**0.001***	0.078	**< .001****	—	
9: Anisochrony Detection	Spearman's rho	-0.160	-0.090	-0.160	-0.130	-0.290	0.228	0.218	0.260	—
	df	56	55	55	55	56	56	56	57	—
	p-value	0.235	0.485	0.229	0.327	**0.026***	0.086	0.100	0.050	—

*Note: * p = < .05, ** p = < .001.*

Exploratory Factor Analysis

A central research question in the present thesis is how the tapping memory task relates to rhythmic skills as tested by the BAASTA battery of tasks. An exploratory factor analysis was performed to try to separate out different aspects of rhythm skills. The following table contains the results of this exploratory factor analysis for the tapping memory and BAASTA rhythm tasks. A two-factor model was used for this analysis. The first factor consists of the Tapping Memory Task, the Beat Alignment Test, the Music Synchronization Accuracy, the Metronome Synchronization Accuracy, and the

Anisochrony Detection tasks. The second factor included the Unpaced Tapping and the Metronome Synchronization Accuracy tasks.

Table 5: An Exploratory Factor Analysis for the Tapping Memory and BAASTA Tasks.

	Factor Loadings		
	1	2	Uniqueness
Tapping Memory Task	-0.402		0.86181
Beat Alignment Test	-0.790		0.35047
Unpaced Tapping		1.005	-0.00314
Music Synchronization Accuracy	0.586		0.70457
Metronome Synchronization Accuracy	0.436	0.400	0.51343
Anisochrony Detection	0.349		0.81897

Note. 'Minimum residual' extraction method was used in combination with 'oblimin' rotation

Regression Analyses

Using these factors, linear regression analyses were run to test if any combination of these factors could be used to predict performance on the non-sense sentence repetition task. First, regression statistics are reported for the Tapping Memory Task on its own. Then, in order to see if the other tasks would increase the portion of variance explained, additional regression analyses were run. Three linear regression analyses were then run using factor 1 as a potential predictor of non-sense word repetition at the consonant, syllable, and word levels. Then, three more linear regression analyses were run using factor 2.

Tapping Memory Task Alone

The following three linear regression analyses were performed to assess the ability of the tapping memory task alone to predict people's performance on the non-sense repetition task. The three tables which follow show the results of these analyses. At the consonant level, a significant result was found, $t = 2.45$, $p = 0.017$. A significant result was also found at the syllable level, $t = 2.79$, $p = 0.007$. The word level also showed a significant result, $t = 3.108$, $p = 0.003$. The Tapping Memory Task explained some portion of verbal memory variance at the consonant ($R^2 = 0.094$), syllable ($R^2 = 0.118$), and word ($R^2 = 0.143$) levels.

Table 6: Multiple Regression Analysis, Tapping Memory Task Predicting the Proportion of Correctly Recalled Consonants for Non-sense Sentence Repetition Task

Predictor	Estimate	SE	t	p
Intercept	0.349	0.129	2.71	0.009
Tapping Memory Task	0.418	0.170	2.45	0.017

Model	R	R²
1	0.307	0.0941

Table 7: Multiple Regression Analysis, Tapping Memory Task Predicting the Proportion of Correctly Recalled Syllables for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.424	0.119	3.57	< .001
Tapping Memory Task	0.439	0.158	2.79	0.007

Model	R	R²
1	0.344	0.118

Table 8: Multiple Regression Analysis, Tapping Memory Task Predicting the Proportion of Correctly Recalled Words for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	-0.0214	0.139	-0.154	0.878
Tapping Memory Task	0.5715	0.184	3.108	0.003

Model	R	R^2
1	0.378	0.143

Factor 1

What follows in this subsection are the results of three more linear regression tests that used factor 1 to try and predict participants' performance in the non-sense sentence repetition task at the consonant, syllable, and word levels. The models accounted for 8.1 %, 8.9 %, and 15 % of the variation of correctly recalled language units at the consonant, syllable, and word levels, respectively. At the word level, the tapping memory task, but no other components of factor 1, had an estimated positive regression coefficient which was significant, $t = 2.205$, $p = 0.032$. This suggests that the predictive power of the models is driven by the Tapping Memory Task (and even then significantly so at only the word level), with none of the other variables serving as a strong predictor of people's performance on the non-sense sentence repetition task.

Table 9: Linear Regression Analysis, Factor 1 Predicting the Proportion of Correctly Recalled Consonants for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.41375	0.15697	2.6359	0.011
Tapping Memory Task	0.34672	0.21109	1.6425	0.107
Beat Alignment Test	-0.00177	0.02303	-0.0769	0.939
Music Synchronization Accuracy	-0.00336	0.00838	-0.4009	0.690
Metronome Synchronization Accuracy	0.01644	0.02578	0.6379	0.527
Anisochrony Detection	-0.01208	0.01745	-0.6924	0.492

Model	R	R^2
1	0.284	0.0806

Table 10: Multiple Regression Analysis, Factor 1 Predicting the Proportion of Correctly Recalled Syllables for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.48397	0.14520	3.333	0.002
Tapping Memory Task	0.36117	0.19527	1.850	0.071
Beat Alignment Test	-0.00404	0.02130	-0.190	0.850
Music Synchronization Accuracy	-8.98e−4	0.00775	-0.116	0.908
Metronome Synchronization Accuracy	-0.00376	0.02384	-0.158	0.875
Anisochrony Detection	-0.01058	0.01614	-0.655	0.516

Model	R	R^2
1	0.298	0.0890

Table 11: Multiple Regression Analysis, Factor 1 Predicting the Proportion of Correctly Recalled Words for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.05025	0.16771	0.2997	0.766
Tapping Memory Task	0.49723	0.22553	2.2048	0.032
Beat Alignment Test	-0.00144	0.02460	-0.0586	0.954
Music Synchronization Accuracy	-0.00867	0.00895	-0.9678	0.338
Metronome Synchronization Accuracy	0.02045	0.02754	0.7427	0.461
Anisochrony Detection	-0.01482	0.01864	-0.7947	0.431

Model	R	R^2
1	0.387	0.150

Factor 2

The following tables give the results of three linear regression analyses that used factor 2 to try to predict participants' performance in the non-sense sentence repetition task at the consonant, syllable, and word levels. Each of the three models' ability to account for variance in verbal STM was close to zero. None of the contributions of the two factor components were significant, with factor 2 by itself showing no ability to predict people's performance on the non-sense sentence repetition task at either the consonant, syllable, or word levels.

Table 12: Multiple Regression Analysis, Factor 2 Predicting the Proportion of Correctly Recalled Consonants for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.66169	0.0188	35.11314	< .001
Unpaced Tapping	-7.67e−5	0.0425	-0.00180	0.999

Table 12: Multiple Regression Analysis, Factor 2 Predicting the Proportion of Correctly Recalled Consonants for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Metronome Synchronization Accuracy	0.00244	0.0255	0.09574	0.924

Model	R	R^2
1	0.0150	2.26e-4

Table 13: Multiple Regression Analysis, Factor 2 Predicting the Proportion of Correctly Recalled Syllables for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.74545	0.0178	41.793	< .001
Unpaced Tapping	-0.00885	0.0402	-0.220	0.827
Metronome Synchronization Accuracy	-0.01028	0.0241	-0.426	0.672

Model	R	R^2
1	0.0914	0.00836

Table 14: Multiple Regression Analysis, Factor 2 Predicting the Proportion of Correctly Recalled Words for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p
Intercept	0.40318	0.0211	19.0881	< .001
Unpaced Tapping	0.00157	0.0476	0.0329	0.974
Metronome Synchronization Accuracy	-0.00438	0.0286	-0.1534	0.879

Table 14: Multiple Regression Analysis, Factor 2 Predicting the Proportion of Correctly Recalled Words for Non-sense Sentence Repetition Task.

Predictor	Estimate	SE	t	p

Model	R	R^2
1	0.0218	4.77e-4

At the consonant and syllable levels, the strength of a model using the tapping memory task alone was higher than either factor 1 or 2 alone. The tapping memory task explained 9.4% of the variance at the consonant level, 12% at the syllable level, and 14.3% at the word level compared to factor 1 which explained 8% of the variance at the consonant level, 8.9% at the syllable level, and 15% at the word level. In contrast, factor 2 showed no ability to predict performance on the non-sense sentence repetition task. These results are consistent with the findings of the earlier omnibus correlation matrix (see Table 4), where the only task to be significantly associated with any level of the non-sense sentence repetition task was the Tapping Memory Task. When taken together, these results show that the tapping memory task served as the strongest predictor of people's performance on the non-sense sentence repetition task

Discussion

The various tasks used and reported on in the present study can be heuristically divided into at least three groups. One of these groups, that pertaining to the non-sense sentence repetition task, concerns people's phonological STM for linguistic stimuli. While this task gave an indication of people's ability to recall language-like stimuli in the short term, it did not include any semantic element insofar as these were jabberwocky stimuli. A second group, which consists of the various BAASTA tasks included in this study, gave an indication of participants' ability to perceive and produce rhythms. One question asked by the present study is whether a third group, which consists of the tapping memory task, exists distinct from this second. This is to say, whether the tapping memory task appears to be assessing something different from or part of that which is assessed by the BAASTA rhythm tasks. One may expect *prima facie* that the tapping memory task, while also a rhythm task, would appear distinct (though not unrelated) from the other BAASTA tasks because it is a memory task. The BAASTA tasks assess people's ability to perceive and produce rhythm and neither assess nor place any pressure upon people's STM capacity. In contrast, the tapping memory task (while also including an aspect of rhythm perception and production – the latter insofar as participants had to tap back the sequence), assessed participants' ability to remember and repeat a sequence of short and long tones. An important question therefore concerns how the tapping memory task relates to people's phonological STM for language (reflected in the non-sense sentence repetition task) and to other rhythm tasks focused on rhythmic perception and production (seen by the BAASTA tasks).

Omnibus Correlation Matrix

The results of the omnibus correlation matrix (see Table 4) suggest that the tapping memory task is distinct, though not unrelated, to the BAASTA tasks both with respect to themselves and to the non-sense sentence repetition task. The tapping memory task was found to be significantly correlated with participants' performance on the non-sense sentence repetition task at the syllable and word levels. In comparison, none of the BAASTA tasks were significantly correlated with the non-sense repetition task at any level. This is likely because the tapping memory task gives an indication of people's STM capacity for the temporal structure of auditory stimuli, whereas the BAASTA tasks give an indication of people's ability to perceive or produce rhythms – not their ability to hold them in memory. It is unsurprising then that performance on this STM task was found to correlate with the non-sense repetition task (an assessment of phonological STM). What is interesting, however, is that this suggests a relation between people's STM for lists of linguistic and non-linguistic stimuli. This relation may be taken to suggest that those cognitive abilities (such as the phonological loop) that facilitate the accurate recall of linguistic stimuli in STM may also be used for the retention of other auditory stimuli such as music, as has been suggested in the literature (Baddeley, 2017).

The results also demonstrated a great cohesion between the various BAASTA tasks and the levels of the non-sense repetition task. Unsurprisingly people's recall of the non-sense sentence stimuli at the consonant, syllable, and word levels were found to be significantly related to one another. The relations found between the various BAASTA tasks were more complicated but, nevertheless, showed robust correlations between the various rhythmic production and perception tasks. Those BAASTA tasks that related to the production of rhythms (music synchronization accuracy, metronome synchronization accuracy, and unpaced tapping) were generally closely correlated, with the metronome synchronization accuracy being significantly correlated with all the other production tasks. On the other hand, the two rhythmic perception tasks (the BAT and the anisochrony detection) were also significantly correlated with one another. The BAT itself was significantly correlated with each of the BAASTA tasks. Taken together these results show that the BAASTA rhythm tasks are highly correlated with one another and, most importantly, that the production and perception tasks are highly correlated with the other production and perception tasks of the battery.

The tapping memory task, however, was only found to be correlated with the BAT, with no other significant correlations for the other BAASTA tasks being observed. Insofar as the BAT was seen to be correlated with every other rhythm task from the BAASTA, it is unsurprising that it was significantly correlated with the tapping memory task. The lack of a correlation between the tapping memory task and any of the other rhythmic production and perception tasks, however, suggests that the memory task is likely assessing a different cognitive ability. Were it to indicate (or rely upon in a critical manner) people's ability to perceive or produce rhythms, a correlation should have been observed between it and the other tasks from the BAASTA battery. But if it instead showed another cognitive ability entirely, in other words, another rhythmic ability beyond just production and perception, then these results should be expected. The results suggest that the tapping memory task is distinct from the BAASTA tasks, and, most importantly, that it taps into another aspect of people's rhythmical ability. I suggest that this other rhythmical ability is the ability to remember/retain the order/sequence (the temporal order) in which stimuli were encountered in time. A significant correlation to the BAT is still to be expected insofar as that test is very likely the most robust and encompassing of the entire BAASTA battery, giving the best single indication of people's rhythmical ability (Dalla Bella et al., 2017), such that any task which failed to be associated with it may give rise to question if that task was a rhythm task at all.

Regression Analyses

The results of the further regression analyses support the suggestions made in the prior subsection. For these analyses, two rhythm ability factors found in an exploratory factor analysis were used to predict participants' performance on the non-sense repetition task. Their predictive power was less than that of the tapping memory task alone. The regression results performed on the tasks comprising factor 1 revealed that the only significant predictor of people's performance on the non-sense sentence repetition task was the tapping memory task and, even then, only at the word level (see Table 11). The results showed that factor 2 had practically no ability to predict performance on the non-sense sentence repetition task at any level as the variance explained by it was close to zero. The results revealed that none of the rhythmical production or perception tasks served to predict people's performance on the language task. This lack of a significant prediction, coupled with the presence of one for the tapping memory task, supports the idea that the tapping memory task is distinct from the other rhythm tasks of BAASTA. The results showed that the tapping memory task, by itself, served as a better predictor of participants' performance on the non-sense repetition task than either factor 1 or factor 2, in nearly every case. The tapping memory task by itself significantly predicted performance at the consonant (see Table 6), syllable (see Table 7), and word (see Table 8) levels. This suggests not only that the tapping memory task assesses a cognitive ability that is clearly distinct from the other rhythm tasks, but more, that that ability is more related to people's ability to hold phonological stimuli in their STM. This is, however, not to say that people's rhythmical ability is not related to their language ability, a relation which has been well substantiated in the literature (Mehler et al., 1988; Bosch & Sebastián-Gallés, 1997; Nazzi et al., 1998; Soderstrom et al., 2003; Thomson & Goswami, 2008; Corriveau & Goswami, 2009; Strait, Hornickel & Kraus, 2011; Huss, Verney, Fosker, Mead & Goswami, 2011; Strait et al., 2012; Tierney & Kraus 2013; Gervain & Werker, 2013; Tierney & Kraus 2014; Woodruff et al., 2014; Gordon et al., 2015; Tierney et al., 2017; Kachlicka et al., 2019). Instead, the limitations of the conclusions that are to be drawn from the present study's results are discussed in a later subsection.

Differences at the Consonant, Syllable, and Word Levels

While the fact that the tapping memory task reflects a different cognitive ability than rhythmic perception or production is indicated both by the results of the omnibus correlation matrix and the various regression analyses, the critical question remains: What might this different cognitive ability be? It is possible to suggest an answer to this question from the results of the present study insofar as the tapping memory task tended to be differently related to the various levels of the non-sense repetition task.

Within the omnibus correlation matrix (see Table 4), the tapping memory task was found to be significantly correlated to the syllable and word, but not the consonant, levels of non-sense sentence repetition. Moreover, within the regression analyses anytime that the tapping memory task was incorporated into the model it served to significantly predict performance on the non-sense sentence repetition task at the word level. It did not, however, significantly predict performance at the consonant or syllable levels in the context of the variables loading on factor 1 (see Table 9 and Table 10). Furthermore, while the tapping memory task did significantly predict performance at the consonant and syllable levels on its own (see Table 6 and Table 7), the strength of its predictive association (R^2) was higher when predicting the proportion of variance at the word level (see Table 8). Nevertheless, the tapping memory task did a numerically better job of predicting participants' performance at the syllable level than it did at the consonant level (see Table 6 and Table 7). This trend is further supported by the presence of a significant correlation on the omnibus correlation matrix between the tapping memory task and the non-sense sentence repetition at the syllable level (see Table 4) while no significant correlation at the consonant level was observed. In summary, a hierarchy appears to emerge whereby the tapping memory task has a greater ability to predict performance on the non-sense repetition task in the following order: word > syllable > consonant. These results suggest that the tapping memory task might be differently related to (in the sense of being more strongly associated with) performance at the various language units.

One reason why the different levels of the non-sense sentence repetition task would interact differently with the tapping memory task is that they carry varying levels of rhythmical or else temporal material. The syllable proportion scores gave an indication of the number of vowels a participant was able to recall which, therefore, may lead one to think that more of the rhythmical information of the non-sense sentence was carried by the syllable cores than by the consonants alone. Insofar as the tapping memory task is also a rhythm memory task, it is sensible that such a task should be a better predictor of more rhythmically dense units of language – such as is suggested by the slightly differing results at the syllable and consonant level.

The word level of the non-sense sentence repetition task would also have been the most sensitive to temporal deviations in the sequence/order of the stimuli. This is because correct responses were only given when the whole word, with all of its various phonemes, was recalled in the exact order of the stimuli. In contrast, participants could correctly recall syllables and consonants in isolation. For example, with the non-sense word 'EMENTED', participants could potentially receive a proportion correct for answers such as 'EMEN', 'TED', 'NT', or even DEMENTED at the consonant and syllable levels; however, they would need to repeat the whole word 'EMENTED' exactly in its correct order/sequence. Consequently, the word level of the non-sense sentence repetition task (compared to the consonant and syllable levels) was the most sensitive to the sequential order of the stimulus and, therefore, also gave the most information about how well a participant could recall the sequential order of the sentence. The tapping memory task was similarly sensitive to the sequential order of participants' responses, since they had to repeat back the sequences of short and long stimuli in the same order in which they had occurred. These aspects of the tasks design, as well as the results obtained across the various levels of the non-sense sentence repetition task, suggest that the tapping memory task may assess a different cognitive ability than those of the other BAASTA rhythm tasks. This ability may be called 'Sequential Memory', understood as the ability to accurately remember the sequence/order (the temporal structure) in which some stimuli/input were presented in time. This is not to say that Sequential Memory is the only ability demonstrated or assessed by the tapping memory task, however, such an ability would likely be among those most critical to performance in the task. The fact that the tapping memory task had the strongest ability to predict performance at the word level of the non-sense sentence repetition task (which was the only level at which the temporal order/sequence of the repetition mattered) supports this conclusion. This finding relates to, or else supports, the multidimensional view of rhythm found in Fiveash et al (2022), who identified 3 dimensions to rhythmic competence as follows: a) rhythm production, b) rhythm perception, and c) 'sequence memory-based rhythm processing'. This identification was made on the basis of prior research from the literature (Tierney & Kraus, 2015; Dalla Bella et al., 2017; Bonacina et al., 2019), which distinguishes between rhythmic production tasks, rhythmic perception tasks, and sequence memory-based rhythm tasks.

Limitations and Future Directions

The key limitation of the present study is that the non-sense repetition task did not incorporate the learning of any higher-order (semantic or else syntactic) elements of language. Consequently, the observed failure of the BAASTA tasks to predict language ability in the non-sense repetition task should not be taken as an indication that rhythm perception and production skills are not related to language ability. Rather, the results of the present study indicate that the relationship between rhythm and language skills is complex and that many kinds of rhythm tasks (such as those from the BAASTA) may have a relationship primarily to the higher-order features of language. An important future direction of study, therefore, is to analyze the results of the foreign word learning task (which includes a semantic component) that was included in the study in order to discern how these various tasks relate to language learning at the word form level.

Conclusions for the Understanding of Rhythm and Phonological STM

The present experiment sought to clarify how a previously used rhythm STM memory task (the tapping memory task) relates to other rhythm tasks from the literature and how these various tasks relate to people's phonological STM and word learning skills. While the literature overwhelmingly suggests that a connection exists between people's rhythmical ability and their language capacity (Thomson & Goswami, 2008; Corriveau & Goswami, 2009; Strait, Hornickel & Kraus, 2011; Huss, Verney, Fosker, Mead & Goswami, 2011; Strait et al., 2012; Tierney & Kraus 2013; Tierney & Kraus 2014; Woodruff et al., 2014; Gordon et al., 2015; Tierney et al., 2017; Kachlicka et al., 2019), both of these abilities are multi-faceted. Consequently, both rhythmical ability and linguistic ability are composed of, and reliant upon, a number of other cognitive abilities. In order to clarify the relationship between these abilities, participants' performance on a number of different timing tasks assessing their ability to produce and perceive (Dalla Bella et al., 2017) rhythm, as well as their STM for rhythmical sequences (Service & al., 2022), were compared and used to predict their performance on a non-sense sentence repetition task

Contrary to prior evidence which suggests that temporal synchronization is the driving factor of rhythmical ability (Tierney et al., 2017; Rimmele et al., 2022; Luo & Lu, 2023), the results of the present study did not find that rhythmical synchronizing tasks (the tapping to 2 pieces and the tapping to 3 paces tasks) were the most highly correlated with the various rhythm tasks. Instead, the BAT was the only task to be correlated with all of the other rhythm tasks included in the present study (the tapping memory task included). This is interesting insofar as the BAT included no (external) component of production – as in it participants made a single discrimination judgement of whether the superimposed beat was aligned or not aligned with the rhythm of a song. This is perhaps surprising in light of the evidence suggesting that rhythmical abilities are very closely connected to the precision of people's auditory-motor coupling (Tierney et al., 2017; Dalla Bella et al., 2017; Bégel et al., 2017; Kachlicka et al., 2019; Lizcano-Cortés et al., 2022). One may therefore expect that those tasks that most directly assess people's auditory-motor coupling (their ability to time movements to a rhythm) would be the most critical. However, the results of the present study found instead that a perceptual discrimination task was the most closely correlated with all of the other rhythm tasks. It is possible that (even in the absence of any external auditory-motor coupling or movement), during the BAT simulated action like processes occurred within the supplementary motor area as the ASAP hypothesis suggests (Cannon & Patel, 2021). Were this to be the case an auditory-motor coupling may still have been involved internally.

With respect to language, it was only performance on the tapping memory task that served to predict participant's performance on the non-sense sentence repetition in the present study. Performance on any of the rhythmical perception or production tasks from the BAASTA did not significantly predict performance on the non-sense sentence repetition task. The key difference between the tapping memory task and the other rhythm tasks is that it assesses temporal STM. The greater ability of the tapping memory task to predict performance on the non-sense repetition task may suggest the existence of a relationship between phonological STM capacity and language skills (Brady, 1986; Rapala & Brady, 1990; Adams & Gathercole, 1995; Conti-Ramsden & Durkin, 2007; de Abreu et al., 2011; Vulchanova et al., 2014). However, because the language task analyzed in the present study did not assess people's capacity for the higher-level aspects of language, no firm conclusion on this can be drawn until the results of the Finnish word learning task are analyzed.

The most interesting finding from the present study is the unique relationship between the tapping memory task and the various language units of the non-sense sentence repetition task. It showed a stronger ability to predict performance at the syllable level than at the consonant level, and an even higher predictive ability at the word level. In other words, the predictive strength of the tapping memory task improved as the language units included more rhythmical information, with the greatest predictability coming at the word level where the temporal sequence/order of the sentence was the most critical. However, as the trends and regression coefficients were roughly similar and were not directly compared, these findings should be taken as preliminary. Nevertheless, one conclusion that might be drawn from this is that the tapping memory task may indicate a different kind of rhythmical ability than the rhythmical production and perception tasks of BAASTA (from those included in the present study). This suggested ability has here been called Sequential Memory which is understood to be the ability to accurately recall the sequence/order in which stimuli occurred in time. Moreover, this would suggest that the tapping memory task may, following the prior literature (Tierney & Kraus, 2015; Dalla Bella et al., 2017; Bonacina et al., 2019; Fiveash et al., 2022), be classified as a sequence memory-based rhythm task. The increased ability of the tapping memory task to predict performance on the non-sense repetition task (with respect to the other rhythm tasks) may suggest that the ability to accurately remember the sequence/order in which stimuli have occurred is an important aspect of the relation between rhythmical abilities and language skills.

References

Adams, A. M., & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech, Language, and Hearing Research*, *38*(2), 403-414.

Arvaniti, A., & Rodriquez, T. (2013). The role of rhythm class, speaking rate, and F 0 in language discrimination. *Laboratory Phonology*, *4*(1), 7-38.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OuP Oxford.

Baddeley, A. (2010). Working memory. *Current biology*, *20*(4), R136-R140.

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second language research*, *33*(3), 299-311.

Baddeley, A. (2020). Working memory. In *Memory* (pp. 71-111). Routledge.

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

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of memory and Language*, *61*(3), 438-456.

Bailey, J., & Penhune, V. (2010). Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. Exp Brain Res, 204, 91-101.

Bégel, V., Benoit, C. E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). “Lost in time” but still moving to the beat. *Neuropsychologia*, *94*, 129-138.

Bonacina, S., Krizman, J., White-Schwoch, T., Nicol, T., & Kraus, N. (2019). How rhythmic skills relate and develop in school-age children. *Global Pediatric Health*, *6*, 2333794X19852045.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33-69.

Brady, S. (1986). Short-term memory, phonological processing, and reading ability. *Annals of dyslexia*, *36*(1), 138-153.

Canette, L. H., Fiveash, A., Krzonowski, J., Corneyllie, A., Lalitte, P., Thompson, D., ... & Tillmann, B. (2020). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia*, *138*, 107324

Conti-Ramsden, G., & Durkin, K. (2007). Phonological short-term memory, language and literacy: developmental relationships in early adolescence in young people with SLI. *Journal of Child Psychology and Psychiatry*, *48*(2), 147-156.

Cooper, A., Wang, Y., & Ashley, R. (2017). Thai rate-varied vowel length perception and the impact of musical experience. Language and Speech, 60, 65-84.

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *cortex*, *45*(1), 119-130.

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *cortex*, *45*(1), 119-130.

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and specific language impairment: A new look at an old hypothesis.

Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with specific language impairments. *Frontiers in human neuroscience*, *9*, 672.

Dalla Bella, S., Farrugia, N., Benoit, C. E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities. *Behavior Research Methods*, *49*, 1128-1145.

Dalla Bella, S., Foster, N. E., Laflamme, H., Zagala, A., Melissa, K., Komeilipoor, N., ... & Kotz, S. A. (2024). Mobile version of the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA): Implementation and adult norms. *Behavior Research Methods*, 1-20.

de Abreu, P. M. J. E., Gathercole, S. E., & Martin, R. (2011). Disentangling the relationship between working memory and language: The roles of short-term storage and cognitive control. *Learning and Individual Differences*, *21*(5), 569-574.

Dittinger, E., D'Imperio, M., & Besson, M. (2018). Enhanced neural and behavioral processing of a non-native phonemic contrast in professional musicians. European Journal of Neuroscience, 47, 1504-1516.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136-1146.

Fiveash, A., Schön, D., Canette, L. H., Morillon, B., Bedoin, N., & Tillmann, B. (2020). A stimulus-brain coupling analysis of regular and irregular rhythms in adults with dyslexia and controls. *Brain and Cognition*, *140*, 105531.

Fiveash, A., Bella, S. D., Bigand, E., Gordon, R. L., & Tillmann, B. (2022). You got rhythm, or more: The multidimensionality of rhythmic abilities. *Attention, Perception, & Psychophysics*, *84*(4), 1370-1392.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., Schon, € D., 2015. Music training increases phonological awareness and reading skills in developmental dyslexia: a randomized control trial. PLoS One 10 (9). https://doi.org/10.1371/journal.pone.0138715.

Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers in human neuroscience*, *8*, 777.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of memory and language*, *28*(2), 200-213.

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.

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature communications*, *4*(1), 1490.

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental science*, *18*(4), 635-644.

Goswami, U. (2018). A neural basis for phonological awareness? An oscillatory temporal-sampling perspective. *Current directions in psychological science*, *27*(1), 56-63.

Goswami, U., Mead, N., Fosker, T., Huss, M., Barnes, L., & Leong, V. (2013). Impaired perception of syllable stress in children with dyslexia: A longitudinal study. *Journal of memory and language*, *69*(1), 1-17.

Goswami, U., Mead, N., Fosker, T., Huss, M., Barnes, L., & Leong, V. (2013). Impaired perception of syllable stress in children with dyslexia: A longitudinal study. *Journal of memory and language*, *69*(1), 1-17.

Goswami, U., Wang, H., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2010). Language-universal sensory deficits in developmental dyslexia: english, spanish, and chinese. Journal of Cognitive Neuroscience, 23, 325-337. Goswami, U., Mead, N., Fosker, T., Huss, M., Barnes, L., & Leong, V. (2013).

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, *7*(515-546), 1-16.

Honing, H. et al. (2018) Rhesus monkeys (Macaca mulatta) sense isochrony in rhythm, but not the beat: additional support for the gradual audiomotor evolution hypothesis. Front. Neurosci.

Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex*, *47*(6), 674-689.

Kachlicka, M., Saito, K., & Tierney, A. (2019). Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain and language*, *192*, 15-24.

Krause, V., Pollok, B., & Schnitzler, A. (2010). Perception in action: the impact of sensory information on sensorimotor synchronization in musicians and non-musicians. Acta Psychologica, 133, 28-37.

Laasonen, M., Service, E., & Virsu, V. (2002). Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. *Brain and language*, *80*(3), 340-354.

Laasonen, M., Service, E., & Virsu, V. J. (2001). Temporal order and processing acuity of visual, auditory, and tactile perception in developmentally dyslexic young adults. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(4), 394-410.

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*, *25*, 2247-2273.

Langus, A., Marchetto, E., Bion, R. A. H., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech?. *Journal of Memory and Language*, *66*(1), 285-306.

Lizcano-Cortés, F., Gómez-Varela, I., Mares, C., Wallisch, P., Orpella, J., Poeppel, D., ... & Assaneo, M. F. (2022). Speech-to-Speech Synchronization protocol to classify human participants as high or low auditory-motor synchronizers. *STAR protocols*, *3*(2), 101248.

Luo, L., & Lu, L. (2023). Studying rhythm processing in speech through the lens of auditory-motor synchronization. *Frontiers in Neuroscience*, *17*, 1146298.

Luo, L., & Lu, L. (2023). Studying rhythm processing in speech through the lens of auditory-motor synchronization. *Frontiers in Neuroscience*, *17*, 1146298.

Martinez-Montes, E., Hernández-Pérez, H., Chobert, J., Morgado-Rodriguez, L., Suárez-Murias, C., Valdés-Sosa, P., & Besson, M. (2013). Musical expertise and foreign speech perception. Frontiers in Systems Neuroscience, 7, 84

McAuley, J. D. (2010). Tempo and Rhythm Music Perception. Berlin: Springer, 165–199. doi: 10.1007/978-1-4419-6114-3_6

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & AmielTison, C. (1988). A precursor of language acquisition in young infants. Cognition, 29, 143–178

Merchant, H. and Honing, H. (2014) Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. Front. Neurosci. 7, 274

Montgomery, J. W. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech, Language, and Hearing Research*, *38*(1), 187-199.

Myers, B. R., Lense, M. D., & Gordon, R. L. (2019). Pushing the envelope: Developments in neural entrainment to speech and the biological underpinnings of prosody perception. *Brain sciences*, *9*(3), 70.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. Journal of Experimental Psychology: Human Perception and Performance, 24, 756–766.

Overy, K. (2000). Dyslexia, temporal processing and music: The potential of music as an early learning aid for dyslexic children. *Psychology of music*, *28*(2), 218-229.

Ramus, F., Dupoux, E., & Mehler, J. (2003, August). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th international congress of phonetic sciences* (Vol. 3, pp. 337-342).

Rapala, M. M., & Brady, S. (1990). Reading ability and short-term memory: The role of phonological processing. *Reading and writing*, *2*, 1-25.

Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in psychology*, *8*, 1820.

Rimmele, J. M., Kern, P., Lubinus, C., Frieler, K., Poeppel, D., & Assaneo, M. F. (2022). Musical sophistication and speech auditory-motor coupling: easy tests for quick answers. *Frontiers in neuroscience*, *15*, 764342.

Rimmele, J. M., Kern, P., Lubinus, C., Frieler, K., Poeppel, D., & Assaneo, M. F. (2022). Musical sophistication and speech auditory-motor coupling: easy tests for quick answers. *Frontiers in neuroscience*, *15*, 764342.

Saito, K., Sun, H., & Tierney, A. (2018). Explicit and implicit aptitude effects on second language speech learning: scrutinizing segmental and suprasegmental sensitivity and performance via behavioural and neurophysiological measures. Bilingualism: Language and Cognition. doi: 10.1017/S1366728918000895

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.

Slevc, R., & Miyake, A. (2006). Individual differences in second-language proficiency: does musical ability matter? Psychological Science, 17, 675-681

Soderstrom, M., Seidl, A., Nelson, D. G. K., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249-267.

Strait, D. L., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral and Brain Functions*, *7*, 1-11.

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and language*, *123*(3), 191-201.

Swaminathan, S., & Gopinath, J. (2013). Music training and second-language English comprehension and vocabulary skills in Indian children. Psychological Studies, 58, 164-170.

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris*, *102*(1-3), 120-129.

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris*, *102*(1-3), 120-129.

Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of research in reading*, *29*(3), 334-348.

Tierney, A. T., & Kraus, N. (2013-A). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and language*, *124*(3), 225-231.

Tierney, A., & Kraus, N. (2013-B). The ability to move to a beat is linked to the consistency of neural responses to sound. *Journal of Neuroscience*, *33*(38), 14981-14988.

Tierney, A., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in human neuroscience*, *8*, 949.

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.

Vulchanova, M., Foyn, C. H., Nilsen, R. A., & Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year-old children. *Learning and Individual Differences*, *35*, 87-95.

Wade, B. C. (2004). Thinking Musically: Experiencing Music, Expressing Culture. New York, NY: Oxford University Press.

Weismer, S. E., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, *43*(4), 865-878.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*(4), 501-522.

White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*(4), 665-679.

Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences*, *111*(40), 14559-14564.

Appendix

Figure 1. Visualization of Counter-Balanced Order of Conditions/stimuli Repeated After Every Eight Participants.

Recall Order	Stimuli Randomization	Order of Task Presentation
Random	A	PsychoPy First
Sequential	A	PsychoPy First
Random	B	PsychoPy First
Sequential	B	PsychoPy First
Random	A	BAASTA First
Sequential	A	BAASTA First
Random	B	BAASTA First
Sequential	B	BAASTA First

Figure 2. Visualization of the Tone Stimulus Sequences.

1 · · · _ _ · ·
2 _ _ · _ · · ·
3 · _ · _ · · _
4 _ _ _ · _ · ·
5 _ · · · _ _ ·
6 _ _ · · · · _
7 · · _ · _ _ _
8 _ _ · · _ _ _
9 · _ _ · _ · _
10 _ · _ · · · ·

Figure 3. Finnish Words Used by Trial Block.

Block	English Word	Finnish Word
1	shop	haavisto
1	rice	oudosti
1	wool	ilmaisu
1	flag	uupuma
1	king	aitaus
1	meat	ukaasi
2	door	osoitin

2	pond	kuvaamo
2	hand	keittaja
2	boat	eilinen
2	silk	kuusisto
2	rock	aavistus
3	plum	halaistu
3	nose	ajoitus
3	girl	eteinen
3	crab	kollaasi
3	scar	aapinen
3	feet	uusinto

Figure 4. Finnish Word Stimuli by Order Randomization.

Stimuli Order A	Stimuli Order B
uupuma	haavisto
aitaus	oudosti
ukaasi	ilmaisu
haavisto	uupuma
oudosti	aitaus
ilmaisu	ukaasi
eilinen	osoitin
kuusisto	kuvaamo
aavistus	keittaja
osoitin	eilinen
kuvaamo	kuusisto
keittaja	aavistus
kollaasi	halaistu
aapinen	ajoitus
uusinto	eteinen
halaistu	kollaasi
ajoitus	aapinen
eteinen	uusinto

Figure 5. Jabberwocky Sentence Stimuli (in order of presentation).

ROO MUTHTIN FANED IM GRASHNIT
WOE UPLING JIDED MO NAFF
KAY HUSS EMENTED IM KILPH
KADE GILPERNS KARNAYED ROO WUPS
ROO BEACHLORN SWEENS KAY NEDRIL
WOE VOTION PRUSED JISH OH ENTSIAN
OH MUZE TOMASHED MO FLOOKMUN
OH DREG PRILED OT ROO BLANTIS
PEEB TAFFERS SURFEWED ROO ZYPT
KADE ADLOOT CONDLES KAY DOOT
ROO GLOB SUBSITTED IM OH GAST
KADE SMORKET NOOLED OH GUZDIN
KAY SIPPERT SEFT JISH ROO GAUM
WOE HORIDGE JEPLES IM VIGHT
PEEB FLOTTERS PEFT CHAG LAUMSES
MO JOLK MEMBLED ROO BLEARNATES
KADE TRANTOE DRIMES JISH OH PILK
ROO FLOGIN JANED IM OH KINTO
KAY DUCTORM PRICALLS IM THORK
OH KALP SMIRRED MO GAPATTER