SERIAL ORDER IN LANGUAGE LEARNING IN BILINGUALS

SERIAL ORDER IN LANGUAGE LEARNING IN BILINGUALS

By MARIA DE LOS ANGELES LOPEZ RICOTE, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements

for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2020)

Hamilton, Ontario (Cognitive Science of Language)

TITLE: SERIAL ORDER IN LANGUAGE LEARNING IN BILINGUALS

AUTHOR: Maria de los Angeles Lopez Ricote, B.Sc. (McMaster University)

SUPERVISOR: Dr. Elisabet Service

NUMBER OF PAGES: xii, 79

ABSTRACT

The current thesis has two aims to further the understanding of the cognitive underpinnings that are involved in word-learning and language processing. One aim is to understand how individuals are able to make temporary phonological and serial order representations of new words in language and non-language domains. The second aim is to investigate whether the mechanisms involved in maintaining temporary nonverbal and serial order representations are related to verbal short-term abilities. We created four behavioural tasks to determine the processing of phonological short-term memory information, nonverbal short-term memory information, serial order short-term memory information and rhythmic short-term memory. We used 30 adult Spanish-English bilinguals as the target population to also investigate whether individual language abilities in two spoken languages affect the learning of words in a foreign language with distinct prosody and phonology. The first correlational analysis revealed that performance on a serially ordered verbal short-term memory task that involved a language of unfamiliar prosody and phonology was strongly predicted by performance on two serially ordered verbal short-term memory tasks that involved languages of familiar prosody and phonology. A second correlational analysis showed that tasks that tapped into individuals' memory for serial order in the verbal, nonverbal and rhythmic domains were weakly associated with one another. In a third correlational analysis, it was shown that individuals' lexical knowledge of Spanish was not a predictor of their performance on a measure of their serially ordered verbal short-term memory abilities. Multiple linear regression analyses found that none of the tasks that were used to measure individuals' abilities for processing serial order information in the verbal, nonverbal and rhythmic domain were strong predictors for foreign-word learning. Overall, the results show promising findings for the tasks that tapped into serial order short-term memory for verbal information. However, they also suggest

that the nonverbal and rhythm tasks may not be reliable measures of the constructs we were hoping to study. Future work should adjust the tasks to ensure we are properly tapping into individuals' serial order abilities in the nonverbal and rhythm domains.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. Elisabet Service, for her guidance and support through each stage of the process. I would also like to thank my thesis committee member, Dr. Daniel Pape, for his valuable comments, help and assistance at various stages of this project. I would like to acknowledge Chelsea Whitwell, Fareeha Rana and Melda Coskun for their valuable support, assistance, and feedback throughout the entirety of the project. I would like to also express my gratitude to Chia-Yu for her guidance and support all these years, and the ARiEAL Research Centre who made this research possible by financially supporting it and assisting me with the recruitment and compensation of the study subjects. To the undergraduate research assistants at the LMB lab, who were involved in assisting with the piloting and scoring of the tasks for this project, I am extremely grateful. I would like to express an immeasurable amount of gratitude to my colleague, research partner and new life-long friend, Erin DeBorba, who was instrumental in the development, execution and writing of this thesis. Lastly, I would like to thank my mother, Maria-Eugenia Ricote, who was also instrumental in the development, creation, piloting and scoring of the Spanish stimuli throughout the experiment.

I want to also take this time to acknowledge and thank the rest of my immediate family, friends and fiancé for their assistance with various stages of the project such as the recruitment of participants in the GTHA community, the piloting of the experiment and their revisions and feedback of my many thesis drafts. Without your support and unconditional love, I would have not been able to complete this project in the midst of a world pandemic. I am eternally grateful to each of you.

TABLE OF CONTENTS

TABLE CAPTIONS

FIGURE CAPTIONS

LIST OF ABBREVIATIONS AND SYMBOLS

WM = working memory

STM = short-term memory

LTM = long-term memory

NWR = nonword repetition

SLI = specific language impairment

TD = typically developing

ITL = iambic trochaic law

ΔC = standard deviation of consonantal interval

%V = percentage of vocalic space

FWL = foreign-word learning

TRA = temporal rhythm accuracy

TOJ = temporal order judgement

DECLARATION OF ACADEMIC ACHIEVEMENT

The data presented in this thesis were obtained in an experiment carried out by myself, in collaboration with Dr. Elisabet Service and Erin Deborba in the LMB Lab. I played a major role in the preparation and execution of the experiment, and the data analysis and interpretation are entirely my own work. Any contributions from colleagues in the collaboration, such as diagrams or calculations, are explicitly referenced in the text. I am aware of and understand the McMaster's policy on plagiarism and I certify that this thesis is my own work, except where indicated by referencing, and the work presented in it has not been submitted in support of another degree or qualification from this or any other university or institute of learning.

INTRODUCTION

The current thesis has two aims to further understanding of the cognitive underpinnings that are involved in word-learning and language processing. One aim is to understand how individuals are able to make temporary phonological and serial order representations of new words in language and non-language domains. The second aim is to investigate whether the mechanisms involved in maintaining temporary nonverbal and serial order representations are related to verbal short-term abilities. We used four tasks to determine the processing of phonological short-term memory information, nonverbal short-term memory information, serial order short-term memory information and rhythmic short-term memory. In the working memory literature, it has been widely established that tasks that assess serial *verbal* short-term memory abilities are good predictors of language aptitude (Attout, Grégoire & Majerus, 2020; Baddeley, Gathercole & Papagno, 1998; Gathercole, 2006; Gathercole, Hitch, Service & Martin, 1997; Majerus & Boukebza, 2013). However, the literature remains unclear on whether serial *nonverbal* short-term memory abilities can also predict language aptitude outcomes (Lahti-Nuuttila et al., submitted). In addition, we use a bilingual sample as the target population to investigate whether individual language abilities in two spoken languages affect the learning of words in a foreign language with distinct prosody and phonology.

In order to better understand language learning and processing, we examine a few of the acoustic cues that individuals exploit when exposed to speech. First, we investigate the relationship between performance in two pseudoword sentence-repetition tasks that are supported by familiar prosody and phonology and a foreign sentence-repetition task that requires representation of both unfamiliar prosody and phonology. Then, we explore the relationship among all the verbal and non-verbal STM tasks. We also investigate the relationship between performance on the Spanish proficiency test, a measure of lexical knowledge in a language of both familiar prosody and phonology, and performance in a Spanish pseudoword sentence-repetition task. Lastly, we ask whether the newly created measures of potential individual language learning aptitudes, such as performance in sentence-repetition tasks, a temporal rhythm accuracy task and a temporal order judgment task, are able to predict learning of foreign words in a foreign-word learning task.

I will first present a general overview of the theoretical framework of working memory, followed by a review of the literature of phonological, visuospatial, serial order and rhythmic information processing and encoding. I will end the introduction with a detailed explanation of the reasoning behind the tasks and hypotheses in the present thesis. The next chapters present the methods and results of the thesis experiments. The last chapter aims to explain the findings and highlight the implications, and directions for future research.

Background

Working Memory as a Multi-Store System. Working memory (WM) is commonly defined as the systems involved in keeping information in the mind while performing complex cognitive tasks such as reasoning, comprehension, and learning (Baddeley, 2010). Since its conceptualization by Miller, Galanter, and Pribram (1960), WM has been an area of principal interest in the field of cognitive psychology (Baddeley, 2010). The term has been used in various models to describe a system that is necessary for learning, the retrieval of old information, and performance in various cognitive tasks (Baddeley, 1992). The concept of WM was initially further developed in the Atkinson and Shiffrin (1968) multi-store memory model of short-term memory (STM), which described a temporary store for small amounts of information (Atkinson & Shiffrin, 1968). Baddeley and Hitch (1974) considered the possibility that STM served as a general WM to support complex cognitive activities. However, they soon discovered that Atkinson and Shiffrin's (1968) concept of a single information processing channel with a series of successive stages did not accommodate experimental evidence from normal adults, children, and neuropsychological patients (Baddeley, 1992).

Figure 1: The original WM model proposed by Baddeley and Hitch (1974). Figure originally published in Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley and Hitch (1974) abandoned the Atkinson and Shiffrin (1968) model and instead proposed a three-component model of WM that consisted of three distinct but interacting subsystems (**Figure 1**) (Baddeley, 2010). The *central executive* is responsible for attentional control and coordination of information from two slave systems, both of which are limited in capacity and used for short-term storage (Baddeley & Hitch, 1974; Baddeley, 1992). One slave system, the *phonological loop*, is assumed to maintain verbal and phonological information. The other slave system, the *visuospatial sketchpad*, is assumed to maintain visuo-spatial information (Baddeley & Hitch, 1974; Baddeley, 2010). A later development of the multicomponent model introduced a fourth component of WM, the *episodic buffer* (**Figure 2**) (Baddeley, 2000, 2010). This is assumed to temporarily store information and is responsible for conjoining information from the slave systems and long-term memory (LTM) (Baddeley, 2000, 2010), thus performing the storage functions originally proposed for the central executive. The Baddeley and Hitch (1974) model re-defined WM as a multicomponent system that concurrently stores information in the mind while performing complex cognitive tasks such as reasoning, comprehension, and learning (Baddeley, 2010). The conceptual framework of this thesis will be the WM model of Baddeley and Hitch. For the purposes of this thesis, STM will be defined as the process of temporary storage of information, whereas WM will include the processing of stored information.

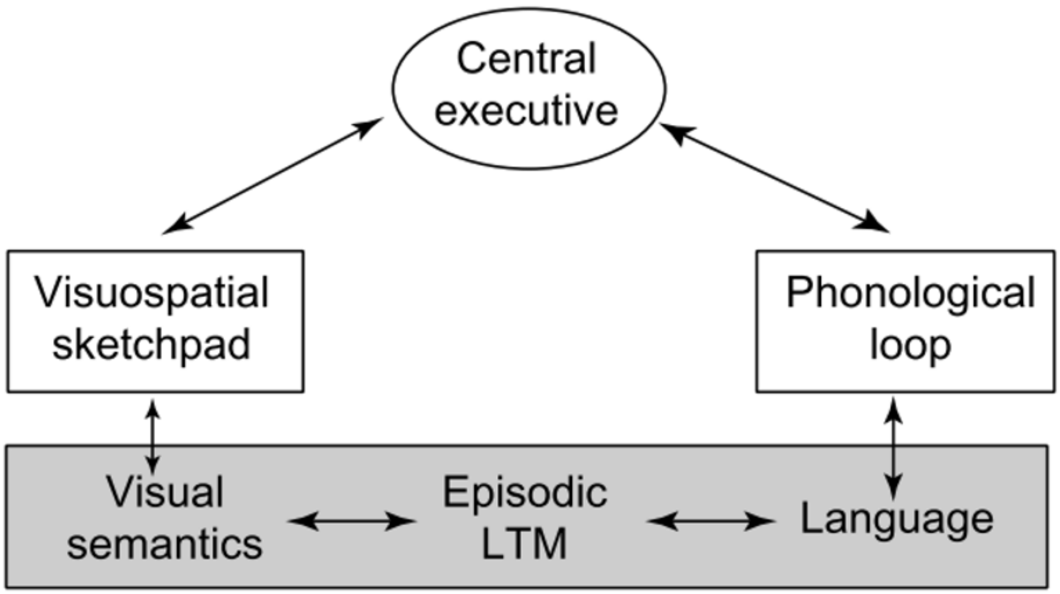

Figure 2: A further development of the WM model proposed by Baddeley and Hitch (1974). Figure originally published in Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

The Slave Systems as Language Learning Devices. The phonological loop is the one component of WM that has been the most extensively studied. This component is comprised of a phonological short-term store and an articulatory rehearsal process that are thought to be necessary for understanding memory of verbal information (Baddeley, 1992). The phonological store is responsible for holding acoustic and speech-based information in their temporal order for a limited amount of time (Baddeley, 1992). In contrast, the articulatory rehearsal process is responsible for maintaining material within the phonological store via subvocal (or internal speech) repetition (Baddeley, 1992). The same process is also thought to be used to register visually presented information (i.e. the names of pictures, written words, etc.) into the phonological store (Baddeley, 1992). There are a number of sources of evidence in the WM literature that provide strong support for the different assumptions describing the concept of the phonological loop (Baddeley, 2012).

The first type of evidence, known as the *acoustic similarity effect* (Baddeley, 2012; Conrad, 1964), supports the assumption that also visually presented verbal information is coded phonologically. It has been found that it is more difficult to remember a list of visually presented words that sound similar in comparison to a list of words that have dissimilar sounds (Baddeley, 2012; Conrad, 1964).

The second type of evidence supports the assumption that verbal information is stored via subvocal rehearsal (Baddeley, 2012). This is known as the *word length effect*, the finding that memory span for words is inversely related to the spoken duration of the words (Baddeley, 1992, 2010, 2012). Thus, longer words take more time to rehearse, and as a result, they are more likely to decay and be recalled poorly (Baddeley, 1992, 2012). Although the time-related claim has been contested, the alternative explanations (e.g., Jones, Hughes, Macken, 2006) also rely on phonology and do not affect the general WM architecture.

The third type of evidence corroborates the assumption that verbal information is stored via subvocal rehearsal. This is based on experimental findings using so called *articulatory suppression* (Baddeley, 2012). It has been shown that having individuals repeatedly say an irrelevant sound, for instance, a single syllable or word such as “la” or “the,” leads to poorer recall of a concurrently presented list of words, independently from the length of the words (Baddeley, 1992, 2012; Baddeley, Lewis, & Vallar, 1984). Additionally, having participants employ articulatory suppression during stimulus presentation and list recall removes the phonological similarity effect for visually presented stimuli, but not for auditorily presented stimuli (Baddeley, 2012; Baddeley et al., 1984). Thus, to encode visual stimuli, participants appear to need to subvocalize the information for it to register in the phonological store. Employing articulatory suppression is assumed to disrupt the subvocal rehearsal and subsequent storage of the visual stimuli (Baddeley, 1992, 2012).

A fourth type of evidence provides support for the assumption that spoken or verbal material is stored by subvocal rehearsal. This evidence is based on the *irrelevant speech effect*, the finding that concurrent presentation of any distracting speech material, whether it comprises nonsense syllables or an unfamiliar language, leads to poorer recall of lists of visually presented words (Baddeley, 1992). During these tasks, disruptive verbal information is thought to gain obligatory access to the phonological store and interfere with the subvocal rehearsal process that concomitantly encodes the visually presented stimuli in the task (Baddeley, 1992; 2012).

Finally, a fifth type of evidence can be drawn from neuropsychological patients with impaired phonological short-term memory. These individuals appear unable to use the articulatory rehearsal process when encoding visual information, show no word length effect, nor experience a disruption from articulatory suppression (Baddeley, 2012).

Work on the phonological loop has not only focused on testing its assumptions, but also on further understanding its fundamental function. Various studies have posited that the phonological loop reflects a capacity to form long-term representations of novel verbal information, a key component of language learning (Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1995; Gathercole, Willis, Baddeley, & Emslie, 1994). As such, the phonological loop provides a temporary storage of unfamiliar phonological forms, while more permanent memory representations are built (Baddeley et al., 1998; Gathercole et al., 1994). Over the years, several studies exploring how the phonological loop is involved in language learning have adapted the nonword repetition paradigm to test their predictions (e.g., Gathercole, Service, Hitch, Adams, & Martin, 1999; Gathercole et al., 1994; Gupta, 2003; Gupta, Lipinski, Abbs, & Lin, 2005; Schraeyen, Van Der Elst, Geudens, Ghesquière, & Sandra, 2019; Service, 1992; Szewczyk, Marecka, Chiat, & Wodniecka, 2018).

The nonword repetition (NWR) task has been widely used to study first language development and second language learning in children. In this task participants are asked to repeat back pseudowords based on the phonology and phonotactics of their native or some other language. Accuracy at the phonological level (from phonemes to utterances) is scored. NWR studies have provided insight into the relationship between word learning and the phonological loop (Gathercole, 2006; Gathercole et al., 1994). NWR provides a measure of the accuracy with which subjects can repeat unfamiliar spoken forms and assesses individuals' coding and retrieval of phonological information in their verbal STM (Gathercole et al., 1994; Schraeyen et al., 2019). This is grounded on the assumption that when repeating the unfamiliar word, subjects are unable to rely solely on lexical knowledge (Gathercole et al., 1999; Szewczyk et al., 2018). Instead, subjects must rely more heavily on their representations of the nonwords in the phonological loop (Baddeley et al., 1998). The use of the NWR paradigm has also allowed for the understanding of the knowledge of children's native languages and the role of the phonological loop in vocabulary learning. That is, children appear to use knowledge of their own language to support the repetition of the novel speech form (Gathercole et al., 1999, 1994; Szewczyk et al., 2018). As a result, it has been established that subjects are more accurate when repeating nonwords that show a high degree of phonotactic similarity with forms in their native language (Baddeley et al., 1998; Gathercole & Baddeley, 1995; Gathercole et al., 1994).

These findings are further corroborated by adult studies, where adults use their existing language knowledge when performing an NWR task and to facilitate learning of novel verbal forms (Baddeley et al., 1998). Nonword repetition has been a central task in enabling researchers to understand the role of the phonological loop in language learning. However, the extent to which both adults and children appear to draw from long-term knowledge of their own native language to perform NWR in a foreign language remained unclear. Papagno and Vallar (1992) found that when individuals learn unfamiliar vocabulary from a foreign language, they are unable to use existing language knowledge to the same extent as they would use knowledge of their own native language to learn new words in their L1. The authors found evidence for this effect by demonstrating that indeed subjects were vulnerable to the acoustic similarity effect and exhibited slower learning when presented to nonwords compared to learning new words (Papagno & Vallar, 1992). Thus, in removing individuals' ability to rely on their knowledge of their native language, they are forced to be more reliant on the phonological loop system because they are unable to use any pre-existing lexical representations available in their long-term storage.

The visuospatial sketchpad is similar to the phonological loop as it is also postulated to be a limited-capacity STM storage system. However, it is mainly responsible for the manipulation and encoding of visual and spatial information, as opposed to speech-based and verbal information. This subsystem has also been the topic of interest in WM research, though work in this area is scarce due to limitations in the techniques and paradigms used to study this subsystem (Baddeley, 2012; Logie & Pearson, 1997). One limitation of the visuospatial sketchpad literature is that several studies have claimed to tap into visuospatial memory but were employing tasks that also tapped into the phonological loop (Baddeley, 2012).

It has recently been proposed that individuals with different language impairments may experience a widespread impairment to overall WM capacity rather than verbal WM deficit (Vugs et al., 2013). A study by Hick et al., (2005) examined performance on various tasks including a visuospatial STM task, a visuospatial processing task, and a verbal STM task. It was found that only the visuospatial STM task showed slower development in young children with specific language impairment (SLI) compared to their typically developing (TD) peers. This is an interesting finding, because over time the SLI children received therapy that was focused on difficulties in the verbal domain and their performance improved when compared to TD peers as a result. However, the slower development in the visuo-spatial short-term memory task shows that children with SLI also experience difficulties that are non-language specific, as the difficulties persisted because the therapy did not address them.

Similarly, in another study, Hoffman, and Gillam (2004) investigated how children with a language impairment performed on two tasks: a verbal digit recall task and a spatial character location recall task. The tasks were designed to assess processes in the phonological loop and the visuospatial sketchpad, respectively (Hoffman & Gillam, 2004). It was found that in both tasks, the children with a language impairment performed more poorly than TD children. These results, along with the growing body of visuospatial STM literature, raise the question of whether language impairments such as SLI or dyslexia are truly just phonological in nature, and confined solely to the phonological loop (Hick, Botting, & Conti-Ramsden, 2005; Hoffman & Gillam, 2004; Vugs et al., 2013). Though the visuospatial sketchpad appears to be affected in language impairment, its involvement in language learning remains unclear (Ramus & Szenkovits, 2008). Thus, evaluating the performance of individuals with language impairments across domains might elucidate the extent to which impairments are due to a deficit in the phonological loop or an overall WM deficit.

This is given that some aspects of language acquisition may rely on modality-general mechanisms, where language learning is dependent on the ability to keep track of the order of phonemes, syllables, and words. While language follows a temporal order by nature, it is also possible that some language learning processes are mapped onto visuospatial order representations. In this way, representing serial order may be an aspect of language learning that uses both the phonological loop and the visual-spatial sketchpad, or it might depend on a cognitive process shared by different WM components.

The Issue of Serial Order. Keeping track of serial order is an important demand for numerous higher cognitive processes (Hurlstone & Hitch, 2015). For instance, retaining serial order in language is necessary for maintaining the order of sequences of sounds within words and of words within sentences (Baddeley, 2012). As discussed earlier, STM for a sequence of verbal items is reliant on the phonological loop, whereas STM for a sequence of either visual or spatial items is reliant on the visuospatial sketchpad (Baddeley, 2012). While the Baddeley and Hitch WM framework has proven useful for explaining some phenomena within serial tasks, it does not provide a specified mechanism for the memory of serial order and temporal properties of items in the two subsystems (Hurlstone & Hitch, 2015).

In the last two decades, various computational models have been proposed to explain the link between serial order and STM within the phonological loop and visuospatial sketchpad. These models postulate that memory for serial order is supported by a timing context signal, where each item is encoded with a temporal signal state and when the signal state is retrieved, a queuing mechanism in parallel reactivates the items (Henson, Hartley, Burgess, Hitch, & Flude, 2003; Hurlstone & Hitch, 2015). The item with the highest activation is selected for retrieval and then supressed, so that the next item becomes the highest activated (Gilbert, Hitch, & Hartley, 2017;

Henson et al., 2003). Other models propose a similar concept of a competitive queuing system that suggest a position mechanism, rather than a timing mechanism, to encode the order of items (Gilbert et al., 2017). Nonetheless, the main conclusion derived from these models is that STM for items does not retain the serial order of a sequence of these items (Burgess et al., 1999; Henson et al., 2003). Instead, a parallel positional code mechanism is responsible for coding the order information while item STM codes the item information (Burgess et al., 1999; Gilbert et al., 2017; Henson et al., 2003). I aim to further understand the extent to which verbal and nonverbal STM provide accurate memory of sequence timing, and whether this is needed for the creation of representations for new word forms in language.

Various findings have demonstrated that order errors in verbal sequence recall are sensitive to temporal qualities such as speech rate, duration, and rhythm (Frankish, 1985; Gilbert et al., 2017; Hartley, Hurlstone, & Hitch, 2016). Henson and colleagues (2003) studied the role of an ordering mechanism in verbal STM. The authors proposed two distinct processes within verbal STM, one for memory of verbal items and the other for memory of serial order. The purpose of this study was to find evidence that a timing signal promotes serial order memory, as assumed in the Burgess and Hitch (1999) phonological loop model. Two tasks that differed mainly in their serial order requirements were compared: an item probe task and a list probe task. The item probe task presented participants with a list of items in sequential order, followed by a single probe item for which participants had to indicate whether or not it had been in the list. The item probe task was designed to tap into individuals' item STM capacity. On the other hand, the list probe task presented participants with a list of items in sequential order, but in this case the probe that followed consisted of a second sequential list, where participants had to indicate whether or not the probe list was the same or different as the first list presented. The list probe task was created to tap into individuals' serial order STM capacity. Overall, the study demonstrated that when either task was paired with secondary tasks that relied on a temporal component, there were greater interference effects on the list probe task compared to the item probe task. These findings suggest that two distinct processes of information encoding within verbal STM exist, although support for the timing signal hypothesis remains limited (Henson et al., 2003). Further evidence for the interaction between phonological STM and an ordering mechanism can be found in studies using immediate serial recall tasks, where subjects have to recall a list of spoken items in their correct order. Immediate serial recall tasks are commonly used as a measure of the capacity of verbal STM (Burgess et al., 1999; Henson et al., 2003). These also may provide a measure of an individual's memory capacity for serial order and temporal properties of lists (Gilbert et al., 2017).

In the visuospatial domain, literature on how the serial order in visuospatial STM is represented remains unclear (Hurlstone, Hitch, & Baddeley, 2014). However, retaining serial order information in the visuospatial domain is necessary for a range of motor skills and social behaviours (Hurlstone et al., 2014). It has been suggested that the ordering mechanism in verbal STM may be similar to the one used in the visuospatial domain (Hurlstone et al., 2014). The concept of both verbal and visuospatial STM domains having similar ordering mechanisms stems from a large body of knowledge that suggests that both share many functional similarities (Hurlstone et al., 2014). These include item and order error distributions, sequence length effects, temporal groupings, item similarity, and sequence repetition (Jalbert, Saint-Aubin, & Tremblay, 2008; Jones, Farrand, Stuart, & Morris, 1995; Parmentier & Andrés, 2006; Smyth & Scholey, 1996).

Hurlstone and Hitch (2015) investigated whether verbal and visuospatial STM depend on common serial order phenomena. The authors compared the patterns of transposition errors in a spatial serial recall task to those obtained in a separate study that examined error patterns in a verbal serial recall task (Farrell, 2008; Farrell & Lewandowsky, 2004). This study demonstrated that the visuospatial serial recall task followed the same empirical pattern as the one demonstrated in the verbal serial recall task by Farrell & Lewandowsky (2004). Moreover, the data gathered was best explained by predictions in the aforementioned competitive queuing mechanism (Farrell, 2008; Farrell & Lewandowsky, 2004; Hurlstone & Hitch, 2015).

Given the consistent cross-domain similarities of serial order STM phenomena, several studies have investigated whether the ordering mechanism of STM is shared by the two subsystems. Depoorter and Vandierendonck (2009) provided evidence for the ordering mechanism in STM being domain-general by comparing performance in a primary task with a concurrent secondary task. The findings suggest that performance in the primary task was more impaired when both the primary and secondary task relied on serial order, regardless of the domain modality of the two tasks. In contrast, Hurlstone, Hitch, and Baddeley (2014) suggest that, although similar in nature, the verbal and visuospatial STM subsystems have their own ordering mechanisms and as such, the ordering mechanisms are domain-specific.

Further evidence for the domain-specific ordering mechanism of STM comes from a study conducted by Soemer and Saito (2016) that compared performance in a serial task when performing a concurrent serial task. The authors compared performance when tasks were both verbal serial tasks, both visuospatial serial tasks, and one was a verbal serial task and the other a visuospatial serial task, and vice versa. Their findings showed that performance in the embedded tasks was lower when subjects performed two serial tasks from the same domain compared to different domains. To summarize the findings on STM for order, it is established that the way in which the two WM subsystems remember and reproduce serial order is fundamentally similar

(Hurlstone et al., 2014). However, whether the ordering mechanism within STM is domain-general or domain-specific remains unclear.

The role of STM for items as opposed to their order in the development of various cognitive skills is presently not well understood. Some evidence suggests that, independent of item STM capacity, serial order STM capacity may be a predictor of lexical and reading development (Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, & Van Der Linden, 2006; Martinez Perez, Majerus, Poncelet, & Perez, 2012). Schraeyen and colleagues (2019), sought to find that adequate serial order recall in verbal STM is necessary for successful reading, and as a result, this ability is related to dyslexia. The authors compared performance in an NWR task between adults with dyslexia and controls. This study examined how well individuals identified phonemes, irrespective of their serial order, as well as how well individuals reproduced the phonemes' serial order. The findings showed that while the control and dyslexic group performed similarly in the phoneme identification task, the dyslexic group performed poorly when recalling the phonemes' serial order.

Further evidence for the disparity between item and serial order STM in dyslexia can be found in neuroimaging studies, where it has been shown that item STM and serial order STM are associated with different neural networks (Majerus et al., 2006; Majerus & Cowan, 2016; Martinez Perez, Poncelet, Salmon, & Majerus, 2015). Similarly, the involvement of serial order STM has been investigated in other modalities outside the verbal domain. Some studies have proposed that dyslexic individuals display difficulties in STM domains such as visuospatial and audio-tactile when asked to reproduce or recognize repeating and novel sequences (Bogaerts et al., 2015; Laasonen et al., 2012; Szmalec, Loncke, Page, Duyck, & Dunantlaan, 2011). Further evidence for this can be seen in a study by Romani, Tsouknida and Olson (2015), in which the investigators examined serial order STM capacity in the visuospatial domain. In this study, performance of reproducing visual characters in a simultaneous condition or a sequential condition was compared in individuals with dyslexia. They also examined how individuals performed in a simultaneous condition and a sequential condition in a visual recognition task. The visual recognition task involved having participants judge whether or not a second list of characters similar to the one used in the two conditions of the visual reproduction task were the same or different. For both types of visual tasks, the simultaneous condition presented characters in a line at the same time. The sequential condition presented the characters one at a time at fixation. The former condition tapped into individuals' spatial order abilities and the latter condition into their temporal order abilities. In their findings, they reported that dyslexic subjects performed worse than matched controls in the sequential condition, but not in the simultaneous condition when doing the visual reproduction task, showing that their dyslexic group experienced difficulties in using their temporal order abilities. In addition, the dyslexic and control groups performed normally in both conditions when doing the visual recognition task. The reviewed findings demonstrate that the observed verbal STM deficits cannot be a direct result of an underlying phonological processing impairment. It appears that not only is item STM impaired, but serial order STM is affected in dyslexic individuals as well, irrespective of the STM modality. In addition, these findings agree with the previously discussed body of evidence. It appears that an ordering mechanism within STM, while possibly separate from the verbal or visuospatial item STM systems, plays an important role in language learning and processing (Hurlstone et al., 2014).

The Rhythm of Language. In previous sections, the case has been made that phonological STM abilities are necessary in the acquisition of new words as well as the development of vocabulary in both native and foreign languages. These findings have been widely demonstrated in healthy individuals as well as individuals with language difficulties (Baddeley et al., 1998; Gathercole & Baddeley, 1989; Gathercole & Masoura, 2005; Papagno & Vallar, 1992; Service & Craik, 1993). The case has also been made for the idea that an ordering mechanism within STM is fundamental for language processing and learning (Gilbert et al., 2017). It is proposed that when individuals are trying to remember spoken material, they will rehearse the information in their heads. In doing this, they will model the time structure of the order in which they initially heard the information. Thus, the better individuals are at mimicking this time structure, the more likely they will be to remember the spoken material (Gilbert et al., 2017). The ability to use language is an integrative experience involved in a wide range of activities from verbal communication to conceptualizing and interpreting ongoing experiences (Bialystok, 2017). Still, each individual's experience of language use will vary fundamentally depending on the environment in which they find themselves. In this way, it could be said that understanding language learning and acquisition may also involve an understanding of how rhythm is manifested within language. Thus, if speech is perceived as a sequence of time, then rhythm can be described as how the order of events is distributed in time (Hausen, Torppa, Salmela, Vainio, & Särkämö, 2013).

Speech rhythm has been demonstrated to facilitate language learning within various levels of speech processing such as word segmentation, detection of intonational phrase boundaries, and processing of dynamic pitch changes and metrical information (Hausen et al., 2013; Langus, Mehler, & Nespor, 2017; Myers, Lense, & Gordon, 2019). This is possible because speech rhythm, also referred to as prosody, provides suprasegmental linguistic features across phonemes, syllables, and phrases (Myers et al., 2019; Pitt & Samuel, 1990). Given that there are certain prosodic cues that are found cross-linguistically, the linguistic rhythm literature has looked further into identifying grammatical properties that are acquired by language learners through rhythm

(Langus et al., 2017). As a result, three levels of a prosodic hierarchy have been universally established: a segmental level, a metrical feet level, and a phonological phrase level (Langus et al., 2017; Myers et al., 2019). At the lowest level, the segmental level, prosody is marked by the alternation between consonants and vowel-like structures (Langus el al., 2017). In the intermediate, metrical feet level, prosody is marked by stressed and unstressed elements alternating (Langus et al., 2017). At the highest level, the phonological phrase level, prosody is marked by the oscillation of words with phonological phrase stress and words with lexical stress but weakly stressed at the phonological phrase level (Langus et al., 2017).

Neurological evidence of this hierarchy of prosody can be seen in numerous studies that have investigated neural entrainment to the different elements of the acoustic signal of speech (Myers et al., 2019). These studies show that there is, indeed, neural oscillatory activity in specific frequency bands at the time-scales of specific linguistic functions. Basic level linguistic processing such as stress and syllable segmentation is suggested to be reflected in oscillations at lower frequencies at ~4 Hz, whereas semantic processing seems to be related to higher frequencies at ~13-50 Hz (Ding & Simon, 2014; Goswami, 2016; Mai, Minett, & S-Y Wang, 2016; Myers et al., 2019). While different levels of prosody represent different elements of speech, all levels group their respective speech elements following the same principle, the iambic trochaic law (Bolton, 1984; Langus et al., 2017). This principle is potentially able to explain how different acoustic correlates that determine prosody, such as duration, amplitude and fundamental frequency of the speech signal can be used to bootstrap language learning (Myers et al., 2019). Based on the iambic trochaic law (ITL), speech elements are grouped into iambs (weak-strong order) if they alternate in duration, meaning that the element that is longest and with highest prominence is placed in the final position (Peña, Bion, & Nespor, 2011). Conversely, speech elements are grouped into trochees (strong-weak order) if they alternate in pitch or intensity, meaning that the element with the highest pitch or intensity is placed in the initial position (Peña et al., 2011). The ITL is useful because it provides a universal pattern in rhythmic grouping. Furthermore, the ITL is worth considering when it comes to understanding how prosody is involved in language acquisition, as well as how rhythmic groupings might differ depending on the language (Henny Yeung, Bhatara, & Nazzi, 2018).

A classical theory of language rhythm proposed that all languages belong to one of three rhythm groups: stress-timed, syllable-timed and mora-timed (Abercrombie, 1967; Langus et al., 2017; Ordin & Polyanskaya, 2014; White, Mattys, & Wiget, 2012). Languages were categorized based on the idea of speech isochrony, where the speech signal was thought to be divided into equal intervals (Langus et al., 2017). Hence, languages belonging to a particular class were thought to show equal duration in a specific element of speech, such as syllable stress, for example (Langus et al., 2017; Ordin & Polyanskaya, 2014). In recent years, this theory has been disproven (Ordin & Polyanskaya, 2014; White et al., 2012). However, there is still no consensus on how child and adult language learners alike are able to differentiate the rhythmic patterns of languages that are traditionally labelled as stress-timed such as English, or syllable-timed such as Spanish (Nazzi & Ramus, 2003; Ordin & Polyanskaya, 2014; Ramus & Mehler, 1999).

A series of new theories have been developed to explain the phenomenon of languages belonging to the same rhythmic class sounding more similar to individuals than languages that belong to a different rhythmic class (Langus et al., 2017). In a review, Langus, Mehler and Nespor (2017) proposed that individuals can detect that different languages have different rhythms from the acoustic correlates of prosody at both the basic level and highest level of the prosodic hierarchy of speech. They proposed that at the basic level, individuals obtain the acoustic cues from the alternation of consonants and vowels within an utterance. The authors suggested that evidence for this can be found when comparing the percentage of vocalic space (%V) and the standard deviation of consonantal interval (ΔC) of the speech stream for each language (Langus et al., 2017; Ramus, Nespor, & Mehler, 1999). For instance, Ramus, Dupoux and Mehler (2003) used an AAX discrimination paradigm and resynthesized utterances to examine whether adult French listeners could discriminate between language pairs. The authors found that subjects could discriminate between language pairs when one language was "stress-timed" and the other was "syllable-timed" (i.e. English and Spanish) and vice versa. Conversely, subjects were unable to discriminate between language pairs when both languages belonged to the same "rhythm class" (Ramus, Dupoux, & Mehler, 2003). Hence, Ramus and colleagues (2003) found support for the idea that languages, indeed, can be classified based on rhythmic equivalency at the basic level by using the measures of %V and ΔC.

Langus and colleagues (2017) further theorized that prosody is also important at the level of phonological phrase. It can guide language learning, as its prominence can signal the word order of that language. For example, in head-initial languages like English, syntactic heads precede their complements (i.e. verbs precede object noun phrases in verb phrases), which means that head-initial languages display final phonological phrase prominence. In contrast, head-final languages, like Turkish, where the syntactic heads follow their complements, display initial phonological phrase prominence. This relates to the concept of the ITL, given that prominence in a phonological phrase alternates rhythmically between words that carry prominence and words that lack it (Langus et al., 2017). Hence, based on the ITL, if prominence is realized mainly through pitch or intensity, the phonological phrase will be stress initial and will have a head-final structure (Nespor et al., 2008). However, if prominence is realized mainly through duration, then the prominence will be stress final and the phrase will have a head-initial structure (Nespor et al., 2008).

In a study by Bion, Benavides-Varela and Nespor (2011), the realization of phonological phrases and its influence on the grouping of speech sequences in adults and infants was investigated. Bion and colleagues reported that adult native speakers of Italian (a head-initial language) followed the pattern as predicted by the ITL, where they were able to group syllables alternating in duration as iambs. On the other hand, adult native speakers of Turkish and Persian (head-final languages) were unable to do the same, and instead grouped the syllables alternating in duration as trochees, given that phonological phrase prominence is primarily signalled through pitch and intensity in those languages (Bion, Benavides-Varela, & Nespor, 2011). The authors proposed that these findings provide evidence for how the rhythmic structure of elements of speech, such as phrases, can facilitate language learning. They also proposed that these perceptual cues might be applied at different points of language learning and development (Bion et al., 2011; Nespor et al., 2008). Hence, biases for grouping based on pitch might appear early in language development, whereas biases for grouping based on duration might be dependent on linguistic experience (Bion et al., 2011).

Furthermore, other metrics and explanations have been put forth by other authors also challenging the idea of categorizing languages based on rhythmic classes or stress distributions. For instance, White, Mattys and Wiget (2012) proposed an alternative durational contrast hypothesis, according to which languages that differ in terms of their contrastive rhythm should always be distinguishable irrespective of their “rhythmic class”. This hypothesis was tested by assessing whether or not adult participants could distinguish between two languages that differed only in their durational characteristics. The results showed that participants could distinguish between languages of different rhythmic classes (English and Spanish) as well as between different accents of the same language (different accents of British English). This provides more evidence for the idea that language learners employ various cues in the speech stream, whether those cues are in the form of %V and ΔC, or durational contrasts. While the aforementioned studies employed different metrics to understand how listeners differentiate between languages of seemingly similar rhythmic classes, the amalgamated finding suggests the same essential conclusion – language learners use various acoustic cues based on the timing of information from the speech stream to segment the speech into words and phrases. Thus, all the models discussed thus far seem to reduce the variety of speech acoustic parameters to just durational cues or phoneme distribution of different languages, which is simplistic in a similar fashion as categorizing languages as stress-timed or syllable-timed. However, it appears that in language processing, the ease of encoding the time information and serial order of the phonological input might be closely related to the specific prosody of the language in question.

As previously mentioned, rhythm appears to play a central role in speech, as it cues the location of structural boundaries (Tierney, White-Schwoch, Maclean, & Kraus, 2017). The ITL principle was originally proposed in the literature of music perception (Bolton, 1984; Langus et al., 2017). Assuming that a similar mechanism of temporal order would be in place in language processing as in music perception appears possible. This is because music perception also involves the temporary storage of an entire sequence while encoding the sequence's rhythmic information across time (Bolton, 1984; Langus et al., 2017; Tierney et al., 2017). Mounting evidence has shown that music and speech might, indeed, share common mechanisms of rhythm processing (Hausen et al., 2013). Various neuroimaging studies using electroencephalography (EEG) and magnetoencephalography (MEG) have found that syntactic violations in music and speech elicit the same P600 responses in the brain, while discrimination of phrases induces similar closure positive shift (CPS) patterns (Hausen et al., 2013; Kraus & Chandrasekaran, 2010; Steinhauer, Alter, & Friederici, 1999).

Musical training has also been reported to improve speech-related skills, such as perception of acoustic cues in speech like pitch, intensity, and vowel duration. These are all prosodic elements that have been previously noted to assist language learners and infants bootstrap their language (Chartrand & Belin, 2006; Chobert, François, Velay, & Besson, 2014; Hausen et al., 2013; Moreno et al., 2009; Schön, Magne, & Besson, 2004). Further evidence for the interconnectedness of music perception and speech processing can be seen in studies of brain-damaged patients, where it has been found that individuals with Broca's aphasia experience difficulties with syntactic processing of language and these difficulties are related to challenges in processing structural relations in music (Hausen et al., 2013; Patel, 2005; Patel, Iversen, Wassenaar, & Hagoort, 2008). Hence, the case could be made that both music and speech depend on a serial ordering mechanism.

It has been proposed that rhythm is a temporal pattern that organizes events in time and as such it has similar roles in both music and speech, though in speech, rhythm is best described as part of prosody (Flaugnacco et al., 2014; Wolff, 2002). From this perspective, processing deficits across the music and speech domain could be related to some sort of temporal order judgment or serial order deficit (Flaugnacco et al., 2014; Wolff, 2002). In fact, rhythm deficits have been found to be linked to poor reading and phonological awareness (Flaugnacco et al., 2014). Perhaps the problems described are not due to processing deficits per se, but rather due to an underlying issue with the ordering of the events within the sequences being processed.

While there seems to be agreement that rhythmic abilities and language abilities are related, what is still unclear is which rhythm skills are necessary for which specific aspects of language (Tierney et al., 2017). For this reason, studies of individuals with dyslexia provide a unique opportunity to better understand this relationship. Dyslexia is currently considered a disorder of phonological processing. These individuals struggle with identifying syllable structure, and also appear to experience difficulties with rhyme schemes (Myers et al., 2019; Ziegler & Goswami, 2005). As a result, this phonological processing deficit has been theorized to be due to a deficit in the temporal coding in the verbal and other non-verbal modalities (Flaugnacco et al., 2014; Myers et al., 2019; Thomson & Goswami, 2008; Wolff, 2002). Support for this idea can be found in several studies, where it has been demonstrated that dyslexic individuals indeed have greater difficulty reproducing patterned rhythms of tones. It has also been shown that these individuals show impairments in musical temporal processing and struggle with grouping tasks in both speech and music when compared to controls (Flaugnacco et al., 2014; Overy, Nicolson, Fawcett, & Clarke, 2003; Petkov, O'connor, Benmoshe, Baynes, & Sutter, 2005; Thomson & Goswami, 2008; Wolff, 2002).

In summary, there is strong evidence for the relationship between serial order abilities and language abilities. Furthermore, strong evidence exists suggesting a relationship between language abilities and abilities across other modalities, though it is unclear whether the relationship between abilities across modalities is mediated by the same ordering mechanism. Thus, the question that remains is whether this ordering mechanism is domain-general or domain-specific. The answer could potentially suggest that related deficits across domains are due to a dysfunctioning ordering system that is shared.

Thus far, a case has been made that individuals vary in their rhythmic proficiency depending on their linguistic environment. This review has also established that these individual differences are important and may demonstrate evidence for an ordering mechanism responsible for tracking the prosody and phonology of a given language. I have also discussed a probable connection between music processing and this ordering mechanism and their possible relation with language abilities. Further evidence for the possibility of an ordering mechanism across domains comes from studies that have discovered that rhythm also plays an important role in visuospatial processing (Langus et al., 2017; Peña et al., 2011).

Segmentation of visual information is important for the encoding of visual sequences consisting of visual events containing an arrangement of different identical shapes (e.g. stars, triangles, squares), where better performance in visual information segmentation translates to better recall of visual sequences. (Peña et al., 2011; Swallow, Zacks, & Abrams, 2009). Various studies have investigated how segmentation of visual information relates to the perception and acquisition of nonverbal aspects of language, and whether the principle of ITL can be used to explain how individuals also segment visual input (Langus et al., 2017). In terms of the ITL, it has been demonstrated that this principle is used in the visual modality in a similar way as in music and speech. Individuals use duration, flicker frequency (visual analogue of pitch), and intensity as physical correlates of visual rhythm in order to segment continuous visual sequences (Kelly & Burbeck, 1984; Langus et al., 2017; Spillmann, 2006; Zacks, Kumar, Abrams, & Mehta, 2009).

In a study by Peña, Bion and Nespor (2011), the researchers examined whether the ITL could also explain sequences that were grouped in the visual domain. Similar to findings in music and speech processing, they found that whenever the events within a sequence varied in their duration, the events with the longest duration were in the final position. Likewise, whenever the events within a sequence varied in their temporal frequency or intensity, the events with the highest temporal frequency or intensity were in the initial position. Thus, it would appear that the better individuals are at tracking changes of different elements in a continuous stimulus sequence, the better they will become at learning the sequence. However, the ITL is only one aspect of how rhythmic and prosodic cues assist with the encoding of events in time across modalities.

While the ITL is helpful in explaining how humans exploit the acoustic or physical correlates of rhythm within a sequence of events (either verbal or visual), it does not explain how the timing of the events within a verbal or visual sequence is encoded into the short-term store. Thus, when individuals are presented with a sequence of verbal (or nonverbal) events, they are keeping track of the prosodic (or rhythmic) patterns of those events in time. To accomplish this task, individuals must rely on their serial STM to encode the changes in time or temporal order of the events in a sequence. Concomitantly, they must rely on their item STM to encode the individual events within that sequence. Therefore, when individuals are presented with a sequence, they are not solely trying to remember the order of the events or the events themselves. Instead, they are also using intrinsic characteristics of the sequence to signal how these events should be ordered or grouped in time.

Purpose

The current thesis aims to clarify some of the cognitive underpinnings that are involved in foreign-word learning and language processing. The primary goal of this thesis was to unravel the mechanisms by which individuals are able to make temporary phonological and serial order representations of new words. A secondary goal was to assess whether the mechanisms involved in maintaining temporary nonverbal and serial order representations are related to performance on tasks assessing phonological STM. This thesis investigates the hypothesis that an underlying mechanism for order in time affects the amount of information individuals can hold in their verbal STM store, nonverbal STM store, serial order STM store and rhythmic/prosodic STM store. The larger a capacity that individuals have in these forms of STM, the better they will be at creating and maintaining temporary representations of new words. This in turn would result in better long-term memory of the words and greater vocabulary size in a given language (Gathercole et al. 1994; Szewczyk et al., 2018).

We developed four tasks to determine how bilinguals process phonological item STM information, nonverbal item STM information, serial order STM information, and rhythmic STM information. The reasoning behind using bilinguals as the target population was to measure individual language abilities in two spoken languages by the same individual, thus allowing more fine-grained analysis. We also wanted to see how bilingual abilities (if any), and their proficiencies affected learning words in a foreign language with a different prosody and phonology.

As discussed earlier, the better individuals are at repeating nonwords (or pseudowords), the better their language-related abilities appear to be (Gathercole et al., 1994; Gupta et al., 2005; Szewczyk et al., 2018). The ability to recall item information and serial order information in STM, has been argued to involve distinct processes (Schraeyen et al., 2019). Nonwords, like new words, are processed like a sequence of sounds, and, as a result, nonword repetition relies on the verbal STM store as well as serial ordering mechanisms (Gupta et al., 2005). The NWR paradigm mimics the environment in which individuals find themselves when learning a new word in both native and foreign language acquisition, where in order to correctly reproduce a new word, individuals must correctly recall all of the phonemes that form the word as well as the phonemes' correct positions in the serial order (Schraeyen et al., 2019; Szewczyk et al., 2018). As a result, the NWR task has proven to be a sensitive measure of individual differences in phonological STM, as it is able to predict language learning ability and rapid learning of the phonology of new words in children, healthy adults and neuropsychological patients (Baddeley et al., 1998; Gathercole & Baddeley, 1990; Gathercole, et al., 1999; Gathercole et al., 1994; Gupta, 2003; Gupta et al., 2005; Service, 1992).

However, also LTM has been found to influence individuals' performance in NWR. Knowledge of the language has been found to influence the accuracy of nonword repetition (Szewczyk et al., 2018). Thus, if relying on lexical knowledge improves individuals' ability to temporarily represent a nonword (or a new word) in the phonological STM store, then individuals who are more proficient in the language should be better at repeating nonwords that resemble the words in that given language (Szewczyk et al., 2018). Developing a sensitive NWR test for adults is often difficult as pseudowords that are based on individuals' native language tend to lead to ceiling performance because of the numerous associated phonological and semantic representations available to them. Therefore, there is a need to explore more demanding tasks while avoiding articulatory difficulty from highly unusual phoneme combinations. In an attempt to achieve this, one of the tasks we developed is a sentence-repetition task, which employs similar cognitive mechanisms as the NWR. Similar to the NWR, the objective of the sentence repetition task is to see how individuals recognize and discriminate between sequences of speech sounds, store phonological representations into their verbal STM store, and retrieve the phonological representations from the verbal STM store in order to make an articulatory response (Munson, Swenson, & Manthei, 2005; Rispens & Parigger, 2010; Schraeyen et al., 2019).

The sentence-repetition tasks developed here involved three variants: an English pseudoword sentence-repetition task, a Turkish sentence-repetition task and a Spanish pseudoword sentence-repetition task. The concept behind having participants repeat back sentences consisting of pseudowords or foreign words is to develop a better test for individual language learning aptitude by assessing individuals' phonological STM and temporal order abilities in a sentence structure. As reviewed above, prosody of a language provides information for speech segmentation, independent of phoneme information (Becker, Schild, & Friedrich, 2018). Hence, when individuals are provided with full sentences, they are being provided with more prosodic information about the given language, which in turn should aid them in remembering the embedded pseudowords and foreign words. One of our research goals was to investigate the characteristics of the three variants of sentence repetition task, and the ability of the tasks to predict individual differences in adults learning new foreign-word forms.

We investigated individual language learning aptitudes in bilinguals. For a better understanding of the bilingual skills of the participants, measures of their first and second language abilities were taken. In the sample, 24 out of the 35 total participants recognized Spanish as their first language. For them, the Spanish pseudoword sentence-repetition task was expected to be supported by their first language abilities, while the English pseudoword sentence-repetition task relied on their second language abilities. In contrast, for the remaining 11 participants in the sample, the English pseudoword sentence-repetition task was expected to be supported by their first language abilities, whereas the Spanish pseudoword sentence-repetition task would reflect their second language abilities. The Turkish sentence-repetition task served as a measure of how individuals keep track of the phonological information and temporal order in a foreign language unrelated to either of the bilinguals' strong languages. Real words were used in the Turkish sentences. Because Turkish has a very different prosodic pattern and phonology from both English and Spanish, individuals were expected to have little support from the phonology or semantic associations from their two strong languages. Thus, Turkish sentences were expected to load more on individuals' phonological STM. Using pseudowords in the Turkish sentences could have resulted in less natural items that would have been even more unfamiliar than real foreign words (Roodenrys, Lethbridge, Hinton, Nimmo, & Hulme, 2002). Therefore, we propose that the pseudoword sentences assess phonological and serial order STM supported by familiar phonology and prosody, whereas the Turkish sentences assess phonological and serial STM unsupported by familiar prosody and phonology.

A second task that we developed is a foreign-word learning (FWL) task, adapted from the paired-associate recall paradigm used by Service and Craik (1993). In this task, an English word was paired up with a word in a foreign language. A list of such word pairs was auditorily presented four times to participants for probed recall of the foreign words after each presentation . This task was used as a measure of individuals' ability to learn phonological representations for LTM (Service & Craik, 1993). This task also used Turkish as the foreign language because it has unfamiliar prosody and phonology in comparison to the bilinguals' spoken languages. We propose that performance on this task, similarly to the three sentence repetition tasks, is also supported by the ability to form representation of temporal order in phonology and prosody. Therefore, performance in word form learning should be able to be predicted by the sentence-repetition tasks as well as other tasks that assess individuals' serial ordering abilities.

Given that language learning also relies on the perception and reproduction of patterns in time (Tierney et al., 2017), a third STM task was created to assess individuals' rhythmic skills. The reasoning behind having a task that assesses individuals' general rhythmic abilities was to see whether these abilities were related to individuals' ability to encode phonological and prosodic information in STM as well as their domain-general serial order abilities. Rhythm reproduction, like phonological reproduction, also requires WM capacity and an ability to group events into meaningful chunks (Flaugnacco et al., 2014; Gilbert, Boucher, Jemel, 2014; Gilbert, Hitch, & Hartley, 2017). In previous work, empirical evidence has found that rhythm reproduction is strongly associated with reading abilities, verbal memory, phonological awareness and pseudoword reading accuracy (Flaugnacco et al., 2014; Tierney et al., 2017). Hence, the temporal rhythm accuracy (TRA) task was created to specifically assess individuals' rhythm skills and serial order STM abilities. In this task, participants listened and reproduced by tapping on a keyboard ten different tone sequences, consisting of seven exemplars of a tone of constant pitch, each either short or long. The TRA task was scored based on individuals' reproduction accuracy by calculating the proportion of correctly reproduced tone lengths across all tone sequences. It was hypothesized that similarly to pseudoword repetition, the more accurate individuals would be in their responses, the greater their verbal STM capacity and serial order abilities would be.

In the literature it has been demonstrated through various studies that a phonological deficit may not be painting the full picture of the underlying processing impairment in cases of verbal STM impairment (Hurlstone et al., 2014; Lahti-Nuuttila et al., submitted). One possibility is that there is a serial order deficit within STM that leads to the language-level impairment, because individuals with phonological impairments often also show serial order deficits across STM modalities. Memory for serial order in verbal and visuospatial STM domains appears to be functionally similar, although the processes leading to the storage of item STM information are separate (Majerus, 2019; Lahti-Nuuttila et al., submitted). However, the relationship between language learning and serial order STM in domains other than the phonological modality remains unclear (Majerus & Cowan, 2016; Lahti-Nuuttila et al., submitted). Most of the research on how serial order is encoded within STM has been observed using measures assessing verbal STM, given the serial nature of language (Lahti-Nuuttila et al., submitted). Studying serial memory in the context of nonverbal tasks might also provide further evidence for the relationship between general serial STM and language abilities (Lahti-Nuuttila et al., submitted). Adams and Gathercole (2000) investigated whether the relationship between language and STM extended to other domains, such as the visuospatial modality. They used the Corsi Blocks task as one of the visuospatial measures to assess visuospatial STM. In this task, participants see a set of spatially distributed blocks and have to reproduce a path modelled to them by tapping them in a specific temporal order. The authors found that the Corsi Blocks task was the only visuospatial STM measure that showed a strong association with language performance. They attributed this finding to the fact that the processing mechanisms needed for the Corsi Blocks task have a more prominent serial component, which is functionally similar to the mechanisms employed in verbal span memory tasks. Thus, more studies are needed to unravel the associations between domain-general mechanisms of serial order and language abilities by investigating how serial order in nonverbal modalities relates to language performance (Lahti-Nuuttila et al., submitted).

I included a fourth task to get a measure of individuals' serial order STM abilities in nonverbal domains. The temporal order judgment (TOJ) task was originally created by Finnish collaborators, Lahti-Nuuttila et al. (submitted), to be presented to pre-school aged children. The aim of the task is to assess individuals' serial abilities in a nonverbal task presented in the auditory and visual modalities. The TOJ task engaged participants' serial order mechanisms by having them listen to or see a series of lengthening sequences of auditory or visual stimuli and asking them to keep track of the order of the stimuli in each sequence for a later comparison with another sequence. As the TOJ task was originally created with the aim of testing serial abilities in children, there was a possibility of ceiling effects when adult participants were tested. To prevent the participants from using skilled verbal strategies, they were additionally asked to perform articulatory suppression, i.e. continuously repeat a syllable "lah" while completing the tasks. This should also have prevented participants from using phonological rehearsal and ensure the tasks were, indeed, tapping into their serial nonverbal STM abilities. The two variants of the TOJ task were scored by calculating the proportion of correct comparison responses across trials. It was hypothesized that the better individuals are able to keep track of the order of events within the nonverbal modalities, the better their performance should be in tasks assessing serial order in verbal STM and in language learning tasks taxing order STM.

To summarize, the present thesis investigates cognitive mechanisms involved in STM for order and their relation to foreign-word learning. To do this, four behavioural tasks were developed that measured rhythm abilities as well as serial order in the verbal and nonverbal STM domains in adult bilingual speakers. In a first research question, we assumed that language learning abilities in a language with familiar prosody will be associated with the learning of phonological sequences in a language of unfamiliar prosody. We predict that participants that perform higher in the English and Spanish pseudoword sentences, will also perform higher in the Turkish sentences. In a second question, we investigate whether performance on the verbal language measures is related with that on the non-language (or nonverbal) measures. We expect that performance in the sentence-repetition tasks, which are measures of serial order in verbal STM and phonological memory abilities will be related to performance in the TRA and TOJ tasks, which are measures of rhythmic STM abilities and serial order in nonverbal STM, respectively. Thus, we expect that both measures should tap mechanisms that underpin performance in the phonological STM tasks, and, therefore, have a positive relationship with the tasks that assess phonological memory abilities. We also explore how sensitive the Spanish pseudoword sentence-repetition task is to Spanish proficiency. We predict that individuals who scored higher in the Spanish proficiency test will have greater support from Spanish prosody and their Spanish lexical knowledge, and, as a result will perform better in the Spanish pseudoword sentence-repetition task than individuals who scored lower on the test. Based on the assumption that prosody supports learning of phonological sequences such as those in new words, measures of language learning that are involved in speech prosody should predict learning of foreign words. In this way, we hypothesize that because the sentence-repetition tasks, the TRA task and the TOJ tasks tap on different elements of language learning aptitude they should predict performance on the FWL task.

METHODS

Participants

Thirty-five Spanish-English bilinguals aged between eighteen and forty-eight years old (Mean = 24.22 years, SD = 6.43 years) from the Hamilton Community took part in this study. Participants were recruited through the Psychology, Neuroscience and Behaviour Participant Pool, the Linguistics Participant Pool at McMaster University and via online postings and posters in the Hamilton Community. Participants had no prior familiarity with the Turkish language, reported normal hearing, normal or corrected-to-normal vision and no gross neurological conditions. The breakdown of the participants was as follows: Out of the thirty-five bilingual participants, eleven individuals were native speakers of English and the other twenty-four participants reported being native speakers of Spanish. The criterion we used to determine whether individuals were native speakers of Spanish or English was by asking whether participants had been born in a Spanish or English-speaking country and whether they grew up speaking either Spanish or English at home. All participants gave informed written consent prior to the beginning of the study and received course credit or $30 cash compensation and parking costs coverage for their participation. This protocol was cleared through the McMaster University Research Ethics Board.

Language Proficiency Tests

Two surveys were created on LimeSurvey to assess participants' English and Spanish proficiency levels. The two proficiency tests were derived from some of the subtests in the TOEFL (Test of English as a Foreign Language) and the SIELE (Servicio de Evaluación de la Lengua Española). We presented participants with four question types in both language tests to obtain a holistic measure of their proficiency level in all areas of both languages. Each test contained one question per each area of interest namely, reading comprehension, listening comprehension, written grammar and, oral expression.

The structure of all the questions was the same for both the English and Spanish proficiency tests. The reading comprehension question consisted of participants reading a text of approximately 450 words about a historical event or a natural phenomenon of the same complexity as a newspaper article and answering three multiple-choice questions about the information they had read on the text (see texts in **Figure i** and **Figure ii** in **Appendix B**). The listening comprehension question consisted of participants listening to an audio clip of about 3 minutes in length of a history lecture. The lecture in the English proficiency test was of the same difficulty as a Grade 10 lecture. In the Spanish proficiency test, the difficulty matched that of a freshman undergraduate history lecture. The differences in difficulty of the lectures between the two tests were decided based on the fact that all participants were predominantly post-secondary students and had reported attaining their highest level of education (post-secondary) in predominantly English-speaking institutions. Hence, we assumed that in order to attend higher education English-speaking institutions, students needed to have a certain level of English. Thus, we decided to increase the difficulty level of the Spanish lecture to assess participants' post-secondary levels in Spanish. As participants listened to the auditory lecture, they had to simultaneously choose six out of twelve statements that best described the information that they had just heard (see transcripts in **Figure iii** and **Figure iv** in **Appendix B**). Participants were not allowed to pause the sound file while it played but they were given the opportunity to listen to each lecture twice in its entirety. Both audio recordings were completed in a sound-treated room at the ARiEAL Research Centre using a linear frequency response Sennheiser ME62 omnidirectional capsule microphone, mounted on a Sennheiser K6 body and connected to a Focusrite Scarlett microphone amplifier at an approximate distance of 10 cm in front of the speaker's mouth. The microphone was placed at the centre of the room, slightly off-centre from the mouth of the speaker to avoid plosive noises. The audio files were digitally encoded with Soundforge Pro at a sampling rate of 44100 Hz and a 16-bit amplitude quantization and stored as WAV files. The signal was high pass filtered at 150 Hz to filter out unwanted low-frequency ambience noises. For the English sound file, the speaker was a female native speaker of standard Canadian English and for the Spanish one a female native speaker of standard South American Spanish (the filter did not affect the speaker's f0 because they were female and their f0 was around 220 Hz, below the frequency cut-off). The written grammar task was divided into two parts. The first part had participants read various sentences and fill out a blank with the options provided to make the sentences grammatical (**Figure v** in **Appendix B**). The second part had participants read sentences and choose which underlined word made a sentence ungrammatical (**Figure vi** in **Appendix B**). The oral expression task consisted of having participants describe for 2 minutes a picture of a family of four doing an outdoor activity (see **Figure vii** and **Figure viii** in **Appendix B**). Participants' responses were recorded using a Rode NT-USB pressure gradient microphone with a cardioid pick-up pattern in the testing room at an approximate distance of 20 cm. The voices were digitally encoded onto Audacity software with a mono audio track at a sampling rate of 41000 Hz and 16-bit amplitude quantization and stored as WAV files. Participants' responses were scored following the guidelines that mimic those used by the SIELE test administrators since all participants had had formal education in an English-dominant setting. Participants were scored based on their fluency, communicative ability, accuracy, vocabulary, pronunciation, and content (see **Figure ix** for guidelines in **Appendix B**). The primary investigator, a native Spanish speaker and English bilingual, scored all the participants' responses. A second and third rater, one a native speaker of standard South American

Spanish and the other a native speaker of standard Canadian English, also rated a sample of the participants to ensure that there was agreement in the rating of the oral fluency of participants' Spanish and English responses, respectively. We used Pearson's *r* correlations to determine the agreement between raters. There was a strong and significant positive correlation between the two raters' scoring for the Spanish oral expression $r(8) = 0.90$, (95% CI, 0.625 to 0.976), $p < 0.001$ and a perfect positive correlation between the two raters' scoring for the English oral expression $r(8) = 1$, (95% CI, 1 to 1) $p < 0.001$.

Phonological Memory Tasks

Sentence Repetition. This part of the experiment consisted of a task that assessed participants' phonological serial STM in the context of familiar or unfamiliar prosody. This task was programmed in Super Lab 5.0 and run on an iMac computer with a 21.5-inch screen with a resolution of 1920 x 1080 pixels.

The stimuli consisted of three types of sentences. Ten sentences were made up of English function words and English pseudowords instead of English content words (**Table i** in **Appendix A**), ten sentences made up of Turkish words (**Table ii** in **Appendix A**) and ten sentences made up of Spanish pseudowords, with real Spanish function words and bound morphemes (e.g. plural endings) as shown in **Table iii** in **Appendix A**. The words and pseudowords in the sentences for each task followed English, Turkish and Spanish phonotactic rules, respectively. The prosody (the variation in stress, intonation, and length) of both words and sentences followed the typical pattern of the language in question. The English sentences, the Turkish sentences, and the Spanish sentences, were spoken by a native speaker of standard Canadian English, a native speaker of standard Turkish and a native speaker of standard South American Spanish, respectively. In the three variants of the sentence-repetition task, participants saw the word LISTEN on the screen for

2377 milliseconds on average while immediately hearing the sentence from the headphones. After the audio had played, the word REPEAT appeared on the screen for as long as participants needed it, in order to prompt them to repeat back the sentence they had just heard. Participants' responses were recorded using a Rode NT-USB pressure gradient microphone with a cardioid pick-up pattern. They were recorded and scored by counting the number of correct syllables within each pseudoword (or word) recalled for each sentence. The scoring was performed online with double-checking completed offline.

Foreign-Word Learning. This part of the experiment consisted of a task that assessed participants' phonologically based word form learning. This task was also programmed using Super Lab 5.0 and run on an iMac computer with a 21.5-inch screen with a resolution of 1920 x 1080 pixels. The stimuli consisted of two lists of English-Turkish word pairs.

The FWL task was adapted from the task in Service and Craik (1993). The stimuli consisted of two lists, each with six English-Turkish word pairs (**Table iv** in **Appendix A**). Each list was auditorily presented to participants four times. After each presentation, participants were cued with each English word in the same order that they had heard them during the learning phase and asked to recall the foreign word that it had been paired with. In the learning phase, participants saw an **X** on the screen for 3328 milliseconds on average, while they instantly heard an English word immediately followed by a word in Turkish. In the recall phase, participants heard the English word again and immediately after, a hash-tag (**#**) appeared on the screen for as long as participants needed in order to prompt them to recall the Turkish word that had been paired with the English word they had just heard. For the learning phase stimuli, we used the same female native speaker of standard Canadian English as in the previous tasks to pronounce the English words. For the recall phase, we used a second female native speaker of standard Canadian English to say the

English words. For both the learning and recall phase of the task, we used the same female native speaker of standard Turkish to record the Turkish stimuli as had recorded the stimuli for the Turkish sentence-repetition task. Participants' responses for the word-learning task were recorded using a Rode NT-USB pressure gradient microphone with a cardioid pick-up pattern in the testing room at an approximate distance of 20 cm. The voices were digitally encoded onto Audacity with a mono audio track at a sampling rate of 41000 Hz and a 16-bit amplitude quantization and stored as WAV files. Participants' responses were transcribed by the primary investigator and scored by counting the correct number of syllables recalled for each Turkish word. The scoring was performed online with double-checking completed offline. The same recording and setup were used for recording the sentence repetition task and foreign-word learning as the sound files for the language proficiency tests.

For the English pseudoword sentences and the FWL learning tasks, the primary investigator, a standard Canadian English speaker, scored all the participants' responses. A second rater, a native speaker of English was also used to rate a sample of the participants to ensure that there was agreement in the rating for the responses on the English pseudoword sentence task and the FWL task. For the Turkish sentence task, a standard Turkish speaker scored all the participants' responses. A second rater, a native speaker of standard Turkish, also rated a sample of the participants to ensure that there was agreement in the rating for the responses on the Turkish sentences task. For the Spanish pseudoword sentence task, the primary investigator, a native Spanish speaker scored all the participants' responses. A second rater, a native speaker of standard South American Spanish also rated a sample of the participants to ensure the reliability of the scoring. For all four tasks, a Pearson's r was run to determine if the three pairs of raters agreed on how the participants scored on the sentence-repetition tasks and the FWL task. A series of

Pearson's correlations were run using a small sample of ten participants to assess the correlation between each two raters in each of the tasks. There was a high and statistically significant positive correlation between the two raters for the English pseudoword sentence task, $r(8) = 0.9213$ (95% CI, 0.695 to 0.982), $p < 0.001$ and the FWL task, $r(8) = 0.9960$ (95% CI, 0.983 to 0.999), $p < 0.001$. There was also a high and statistically significant positive correlation between the two raters for the Turkish sentence task, $r(8) = 0.9369$ (95% CI, 0.749 to 0.985), $p < 0.001$. Lastly, there was also a high and statistically significant positive correlation between the two raters for the Spanish pseudoword sentence task, $r(8) = 0.8242$ (95% CI, 0.405 to 0.957), $p = 0.003$.

We computed three phonological serial STM scores for each participant, one for the English pseudoword sentence task, one for the Turkish sentence task and one for the Spanish pseudoword sentence task. Each score consisted of the proportion of correctly recalled syllables in the ten trials that had been completed by participants in each task. The scores for all of the phonological STM tasks were standardized and transformed into z scores to compare performance on these tasks with performance on the other serial STM tasks below. We also computed one score for the FWL task for each participant, the proportion of correctly recalled syllables for each word. All the scores for this task were based on the number of recalled syllables from the two lists. These scores were also standardized and transformed into z scores to compare performance on this task with performance with the sentence-repetition tasks and the other serial STM tasks below. We looked at the proportion of correctly recalled syllables as opposed to words to avoid floor effects in the FWL task.

Temporal Rhythm Accuracy Task

This part of the experiment consisted of a tone reproduction task that assessed participants' serial rhythm abilities. The TRA task was programmed in Super Lab 5.0 and run on an iMac computer with a 21.5-inch screen with a resolution of 1920 x 1080 pixels. The stimuli consisted of ten tone intervals of two lengths, short and long. To create the events for each tone sequence, we used the "Generate" function in Audacity 2.3.0 and set the events at the default frequency of 525 Hz. We also added a fade-in and fade-out effect to each event. We created ten random sequences of seven events, where each sequence differed from one another by varying the order and number of short and long tones. Thus, each sequence of tones consisted of seven events that could have a short or a long event attached to it and the pitch of each event was always the same (see **Figure x** in **Appendix B**). The short tones were cut to be 200 milliseconds in length and the long tones were cut to be 800 milliseconds in length. In the task, participants saw the word LISTEN on the computer screen for 5152 milliseconds on average, and immediately heard one sequence of the tones. After the audio finished playing, the word REPEAT appeared on the screen for 8000 milliseconds and participants had to mimic the series of short and long tones they had heard by tapping with their index finger. Participants' responses to this tone reproduction task were recorded by asking them to press down on the spacebar on the computer keyboard when repeating back the number and the length of the tones that they had previously heard for each sequence. Participants' responses were scored based on the guideline that any "press down" of the spacebar for less than 500 milliseconds was considered a short tone and any "press down" of the spacebar for more than 500 milliseconds was considered a long tone.

We computed four sets of scores for every participant for this task. Scores were based on the proportion of correctly reproduced tone lengths in the ten trials that had been completed by participants. Scores were transformed into z scores to compare performance on this task with the phonological memory tasks and the other serial nonverbal STM tasks.

Temporal Order Judgement Task

This last part of the experiment consisted of a visual and an auditory variant that assessed participants' nonverbal serial STM for event order. These tasks were the same as the ones that were used in Lahti-Nuuttila et al. (submitted). Both tasks were performed on a Samsung Galaxy tablet, running Android 6.0-8.1 Tab A that had a 10.1-inch SM-T585 display with a resolution of 1200x1920 pixels (Samsung Electronics, 2016). The two tasks were custom-created application programs based on the Unity game engine (Unity Technologies) by our Finnish collaborators at the University of Helsinki. To ensure that the participants had to group the presented stimuli to a temporal sequence in their WM, they were presented each item one at a time in a sequence for a limited amount of time. During the task, participants were also required to perform articulatory suppression, where they had to repeatedly say "lah", while performing the two STM tasks. This was in order to hinder inner speech and binding of the presented stimuli to a verbal temporal sequence.

In the visual variant of the TOJ task, participants were presented with four barns on the screen, two at the back of the screen and two at the front of the screen. For this part of the task, participants saw animal-like characters move from the left upper barn to the right upper barn one at a time. Then, the same characters moved from the lower left barn to the opposite lower barn (**Figure 3**). Each character was visible for approximately 1500 milliseconds. The time between the first and second presentation was 3 seconds. The task was to say whether the animals moving between the lower barns had been in the same order as the animals moving between the upper barns. All the animal-like-characters were made-up of the same thirteen basic shapes of similar colouring. The only thing that distinguished the characters from one another was the varied proportions and position of their body parts and their distinct movements. Each sequence of characters was binary, meaning that it only consisted of two different stimuli, sampled from the pool of five possible characters. This was to minimize the effect of remembering animal details. The sequences were presented horizontally across the middle of the screen, where the first presentation of the sequence was slightly above the middle (upper barns) and the second presentation was slightly below (lower barns).

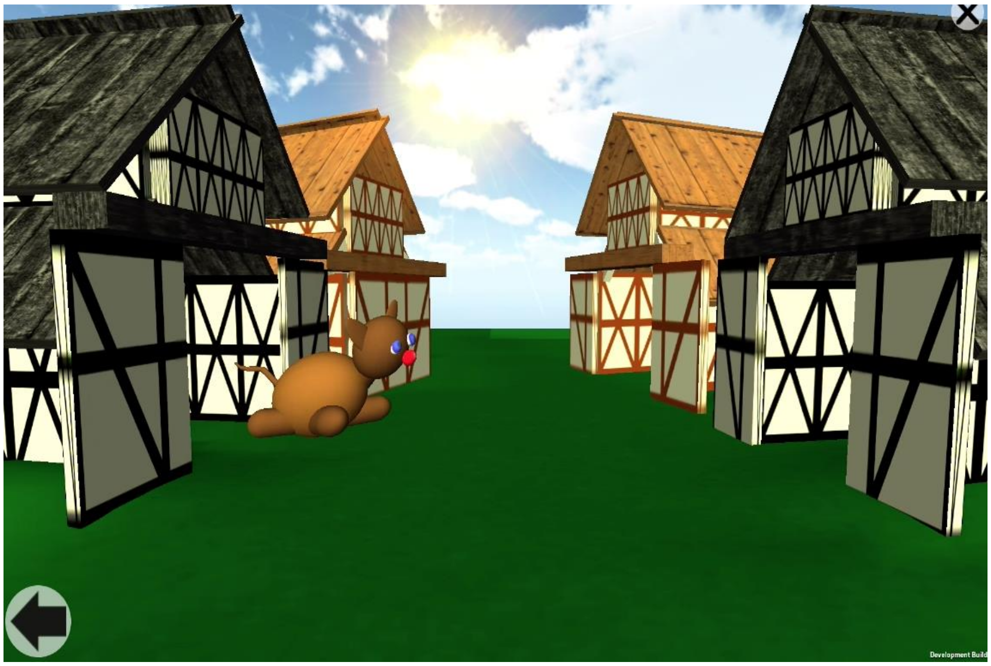

Figure 3: Illustrates the visual component of the TOJ task, where participants were asked to note sequences of fantasy animals that moved from the left side of the screen to the opposite side. Animals sequences were first presented in the upper barns (orange) and then in the lower barns (black).

In the auditory variant of the TOJ task, the stimuli consisted of five different sound files of animal-like calls (Lahti-Nuuttila et al., submitted). Each sound was approximately 1500 milliseconds in length. Similar to the visual task, participants were presented with the four barns on the screen, but this time the light of the screen was dimmed. During the first presentation of a sequence, the upper right barn lit up as the sounds came out. Then, the lower right barn lit up as the sounds came out for the second presentation of the sequence (**Figure 4**). The time between the first and second presentation was 3 seconds. Once again, each sequence of character sounds was binary, where it only consisted of two different stimulus sounds, sampled from the pool of five sounds.

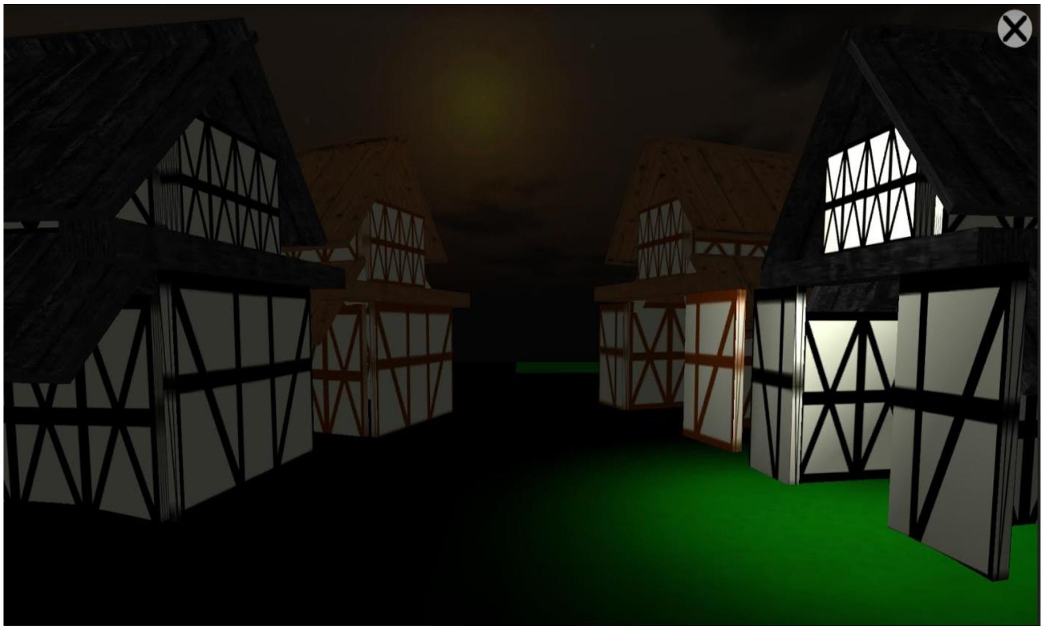

Figure 4: Illustrates the audio component of the TOJ task, where participants were asked to listen to sequences of fantasy animal sounds that came from the upper-right side of the screen and then from lower-right side of the screen. Sound sequences were first presented in the upper barn (orange) and then in the lower barn (black).

Both variants consisted of six trials for each sequence length, though participants were automatically given a perfect score for all six trials if they answered correctly in four trials in a row. Each trial sequence number grew in length from 2 to 7 events. The test ended dependent on participant's performance. Prior to the first test block, participants were presented with five practice trials to ensure they understood the premise of each task. The first three trials of the practice block consisted of a pair of sequences with two stimuli. The last two trials of the practice block consisted of a pair of sequences with three stimuli. The first test block began with a round of two binary stimulus sequences consisting of two stimuli. Then, the trials could go up to six, where participants were presented with a pair of binary sequences consisting of seven stimuli. Each trial within a sequence length involved two different stimuli randomly selected from the pool of five. During each trial, participants were presented with two binary stimulus sequences and asked to judge if the stimuli in the two presentations had followed the same or a different order. For each sequence length block, there were always three trials where the stimuli in the sequence pairs were in the same order and three trials where the stimuli in the sequence pairs were in a different order. The order of the same or different sequences were pseudorandom within a sequence length, but it was always the same for each participant. When the sequence length reached four or more stimuli, the stimuli that might be switched were in the middle positions. In order to move from one sequence length block to the next, participants had to answer correctly four to six trials within a certain length. Once a participant moved onto the next length block, the sequence was increased by one more stimulus. When a participant answered incorrectly on three trials in a row, the task was automatically terminated. Whenever the two sequences of stimuli in a trial were correctly judged as the "same", it meant that both sequences were presented exactly the same. Whenever the stimuli sequence pair was correctly judged as "different", it meant that two different consecutive stimuli had changed places in the second presentation of the sequence. For instance, if the first presentation in a trial consisted of characters A B B and the second sequence was characters B A B, when asked if the order was the same or different, participants were asked to judge by selecting from two options on the screen. The option on the right of the screen consisted of a green circle with a check mark "√" for SAME. The second option on the left of the screen consisted of a red circle with an "**X**" for DIFFERENT (**Figure 5** and **Figure 6**). In this case, participants were expected to tap on the red button to indicate that the two sequences were different. However, if the second sequence followed the same order as the first (characters A B B), then participants were expected to tap on the green circle to indicate that the two sequences were indeed the same. Both circles had a diameter of thirty millimeters and appeared approximately halfway to the screen's vertical midline and the left corner and right corner, respectively.

We computed two nonverbal STM scores for each participant, one for the visual variant of the TOJ task and one for the auditory variant of the TOJ task. Each score consisted of the proportion of correct responses of all the trials that had been completed by the participant in each variant. As such, participants automatically scored perfect in a length block, if they answered correctly in the first four trials out of six. If participants answered incorrectly in any of the first four trials, they were required to complete all six trials of a given length block. The scores for both TOJ tasks were standardized and transformed into z scores to compare performance on these tasks with performance on the serial phonological STM tasks and the TRA task.

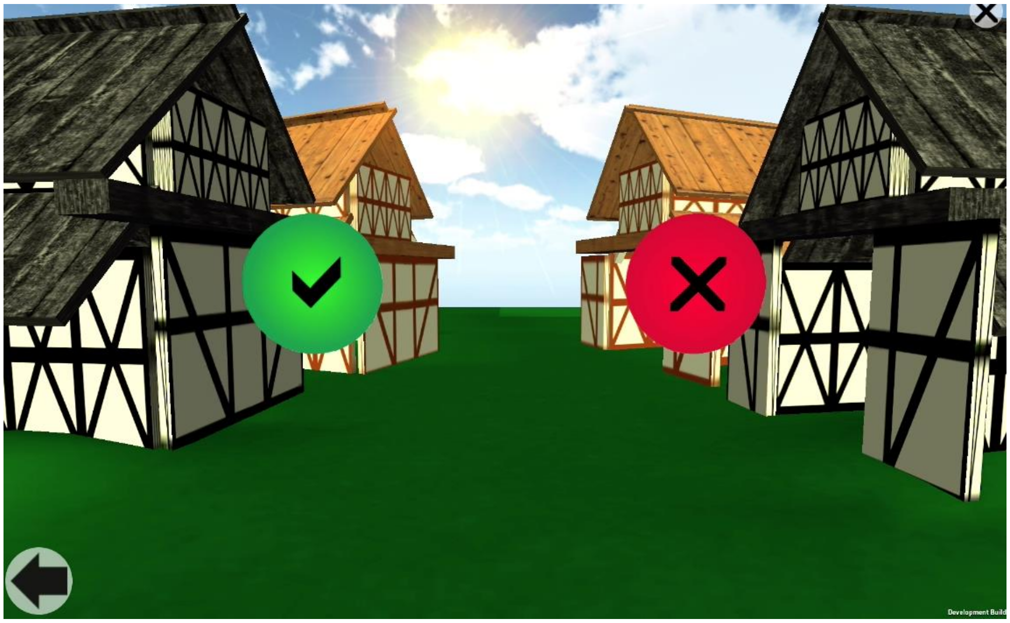

Figure 5: Green and red buttons on the screen used to prompt participants to answer whether the order of the fantasy animals matched the order or did not match the order of the first sequence presented.

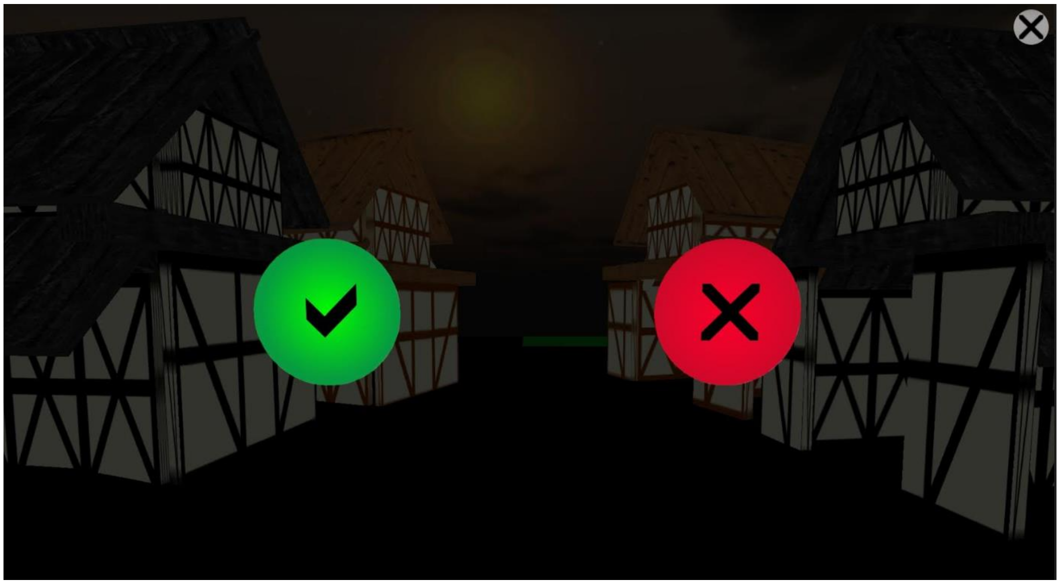

Figure 6: Green and red buttons on the screen used to prompt participants to answer whether the order of the sounds matched the order or did not match the order of the first sequence presented.

Procedure

Before the lab experiments, the participants' proficiency levels in English and Spanish were assessed using the two questionnaires described above on LimeSurvey – an online platform. The order in which participants completed the two language tests was counterbalanced, where half the participants completed the English questionnaire first and the other half completed the Spanish questionnaire first. Participants were also presented with the tasks programmed on Super Lab and the tasks on the Samsung tablet in a counterbalanced order, where half the participants completed the Super Lab tasks following the language tests and the other half completed the tasks on the Samsung tablet following the language tests. All participants completed the five Super Lab tasks in the same order: the English pseudoword sentence-repetition task, the TRA task, the Turkish sentence-repetition task, the FWL task and the Spanish pseudoword sentence-repetition task. For the TOJ tasks, the order in which participants completed the two variants was also counterbalanced, where half the participants completed the visual variant first and the other half completed the auditory variant first. The entire experiment was run with Sony MDRZX110NC noise-cancelling headphones in a in the presence of the experimenter.

Statistical Analyses

We identified and eliminated 6 participants as outliers from our sample using the interquartile range in each of our tasks. Our original sample consisted of 35 participants, leaving our current sample at 29. Using the interquartile range, we identified and removed 5 participants that had a performance below the lower bound limit (values under -0.372 and -1.58) in the visual variant of the TOJ and the Spanish pseudoword sentence-repetition task, respectively. We also removed 1 participant that was identified as an outlier, given that their performance was above the upper bound limit (value above 2.74) in the FWL task. However, this participant was only excluded for analyses involving the FWL task. The removal of outliers did not change any of the findings of this thesis and it was performed to prevent them from distorting our results.

We used various statistical analyses to assess the different predictions we made. Our main questions concerned the relationships between individual differences in our predictor and outcome variables. All variables in our data set were transformed into z-scores prior to analysis to reduce skewness and allow combination of variables. We performed five sets of correlation analyses using Pearson's *r* and Spearman's *rho* correlations, the latter for variables that did not pass the Shapiro-Wilk test of normality of distributions. The first analysis examined how performance on the three sentence-repetition tasks correlated with one another so that we could create a composite phonological STM measure. All the sentence repetition scores were proportions of correctly recalled syllables, which were converted to z-scores for the correlational analyses.

The second correlational analysis examined how our measures of phonological and non-phonological STM are related to each other. It examined how the repetition variable, an audio-visual TOJ variable and the TRA task correlated with one another. The auditory and visual TOJ scores are based on the proportion of correct responses of the theoretical maximum of 36 trials in the tasks. An audio-visual STM variable was created as a composite of the scores in the auditory and visual variants of the TOJ task. A tones score was based on the proportion of correctly reproduced tone lengths.

The purpose of the third correlational analysis was to explore how the scores that participants achieved on the Spanish Proficiency test were related to performance on the Spanish pseudoword sentence-repetition task. The aim of this correlation was to see the relationship between long-term knowledge of Spanish (measured by proficiency score) and individual differences in phonological STM for Spanish pseudoword sentences. To create groups of lower and higher scorers, we performed a median split on participants' scores on the Spanish proficiency test.

We then investigated how the different STM measures predicted foreign word form learning. The fourth correlational analysis examined how the phonological sentence repetition measure, the audio-visual STM measure and the tones STM measure correlated with the FWL task. We also performed two multiple linear regression analyses to further examine the relationship between FWL learning and the other three STM measures. In the analyses, the outcome variable was the proportion of correctly recalled syllables in the FWL learning task and the predictor variables consisted of the repetition variable, the audio-visual variable, and the tones measure. For one of the analyses, the audio-visual STM measure was excluded from the model. Holm's corrections to the *p*-values were made for all correlations to avoid inflating our Type I error rate.

All statistical analyses were carried out using Jamovi software version 0.9.5.12, with the exception of the Holm's corrections, which were carried out using R software version 3.6.1.

RESULTS

Correlation Analyses

All correlations were computed on data for 30 participants, except for the correlations that involved the FWL task performance. These were based on data for 29 participants. We performed five sets of correlation analyses using Pearson's *r* and Spearman's *rho* correlations, the latter for variables that did not pass the Shapiro-Wilk test of normality of distribution. The first correlation examined how performance on the three sentence-repetition tasks correlated with one another so that we could create a composite phonological STM measure. All the sentence repetition scores were proportions of correctly recalled syllables, which were converted to z-scores for the correlational analyses. Basic descriptive statistics of the variables for the first correlational analysis performed among the three sentence-repetition tasks are shown in **Table 1** and the correlations are shown in **Table 2**. As expected, there is a strong positive correlation that is statistically significant between performance in the English pseudoword sentences and the Turkish sentences, $r(28) = 0.759$, $p = \leq 0.001$, $R^2 = 0.576$ (**Figure 7**). The results also found a statistically significant moderate positive correlation between performance in the Spanish pseudoword sentences and the Turkish sentences, $r(28) = 0.462$, $p = 0.020$, $R^2 =0.213$ and $rho(28) = 0.467$, $p = 0.019$, $R^2 = 0.218$ (**Figure 8**). On the other hand, performance in the English and Spanish pseudoword sentences showed only a weak positive correlation, which is not statistically significant, $r(28) = 0.241$, $p = 0.199$, $R^2 = 0.058$ $rho(28) = 0.262$, $p = 0.162$, $R^2 = 0.069$ (**Figure 8**).

Table 1: Descriptive statistics for repetition tasks. Note. Scores represent number of correctly recalled syllables across all sentences presented. All scores portrayed are raw scores.

	English pseudoword sentences	**Spanish pseudoword sentences**	**Turkish sentences**
N	30	30	30
Mean	0.718	0.914	0.519
Standard deviation	0.170	0.077	0.175
Minimum	0.329	0.597	0.227
Maximum	0.959	0.987	0.879
Shapiro-Wilk p	0.258	<.001	0.631

Table 2: Correlations among the sentence repetition tasks. Note. N = 30. * $p < .05$, ** $p < .01$, *** $p < .001$. All scores were standardized and transformed into z scores. All *p*-values in the table are Holm's corrected.

			1	**2**	**3**
1.	**English pseudoword sentences**	Pearson's *r*	—	0.241	0.759 ***
		p-value	—	0.199	<.001
		Spearman's *rho*	—	0.262	0.772 ***
		p-value	—	0.162	<.001
2.	**Spanish pseudoword sentences**	Pearson's *r*		—	0.462*
		p-value		—	0.020
		Spearman's *rho*		—	0.467*
		p-value		—	0.019
3.	**Turkish sentences**	Pearson's *r*			—
		p-value			—
		Spearman's *rho*			—
		p-value			—

.

Based on these correlations, a repetition variable was created as a composite of the English pseudoword sentence-repetition task scores and the Turkish sentence-repetition task scores, leaving out the Spanish pseudoword sentence-repetition task scores which did not have enough variability to be used as a predictor variable. The majority of participants in our sample described Spanish as their first language, and the Spanish pseudoword sentence-repetition task seemed to function as a measure of first language abilities.

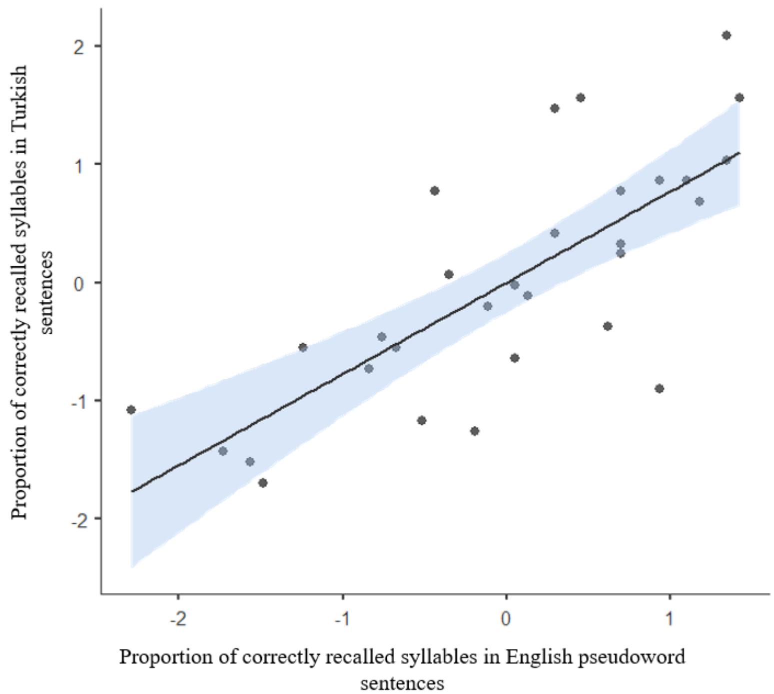

Figure 7: Scatterplot illustrating the strong positive correlation between performance on the English pseudoword sentences and the Turkish sentences.

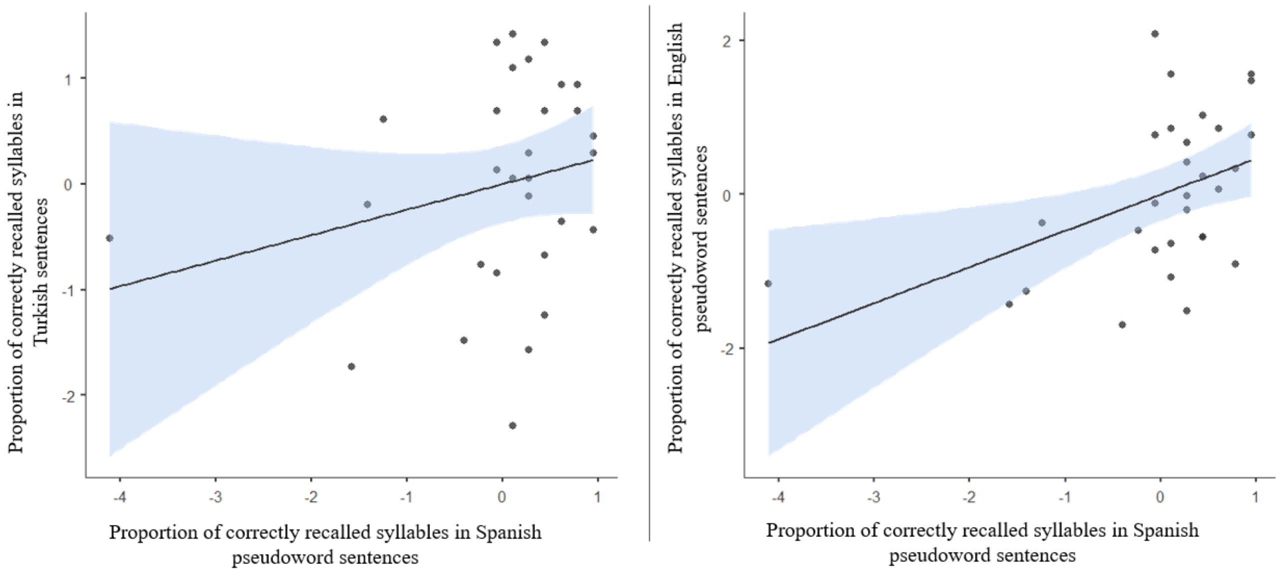

Figure 8: Scatterplots illustrating the correlations between performance on the Turkish sentences and English pseudoword sentences and the Spanish pseudoword sentences.

The second correlational analysis examined how our measures of phonological and non-phonological STM are related to each other. This correlation was performed by examining the relationship between scores on a composite repetition measure, a composite audio-visual TOJ measure, and scores on the TRA task. The English pseudoword and Turkish sentence repetition scores were based on the proportions of correctly recalled syllables. The composite repetition variable combined the scores from the English pseudoword and Turkish sentence-repetition tasks. The auditory and visual TOJ scores are based on the proportion of correct responses of the theoretical maximum of 36 trials in the two variants. The composite audio-visual STM variable was created as a composite of the scores in the auditory and visual variants of the TOJ task. The tones variable was based on the proportion of correctly reproduced tone lengths. Basic descriptive statistics of the variables are shown in **Table 3** and the correlations between the three measures are shown in **Table 4**.

Table 3: Descriptive statistics for repetition tasks, TOJ task, and TRA task. Note. All scores are the raw scores for proportion of correctly recalled syllables, proportion of maximum correct responses and proportion of correctly reproduced tasks, respectively.

	Repetition variable	Audio-visual variable	Tones variable
N	30	30	30
Mean	0.619	0.900	0.720
Standard deviation	0.162	0.113	0.110
Minimum	0.331	0.611	0.529
Maximum	0.912	1.000	0.943
Shapiro-Wilk p	0.406	< 0.001	0.750

Table 4: Correlations among the two repetition tasks, the TOJ tasks and the TRA task. Note. N=30. * $p < .05$, ** $p < .01$, *** $p < .001$. All scores were standardized and transformed into z scores. All *p*-values in the table are Holm's corrected.

		1	2	3
1. Repetition variable	Pearson's *r*	—	0.264	0.381
	p-value	—	0.243	0.136
	Spearman's *rho*	—	0.179	0.407
	p-value	—	0.562	0.087
2. Audio-visual variable	Pearson's *r*		—	0.302
	p-value		—	0.206
	Spearman's *rho*		—	0.217
	p-value		—	0.486
3. Tones variable	Pearson's *r*			—
	p-value			—
	Spearman's *rho*			—
	p-value			—

The results appear to show that a moderate positive correlation, though not statistically significant, exists between performance in the English pseudoword and the Turkish sentence-repetition tasks (repetition variable) and performance in the TRA task (tones variable), $r(28) = 0.368$, $p= 0.087$, $R^2 = 0.135$ (**Figure 9**). In addition, the results also found a weak positive correlation between performance in the English pseudoword and Turkish sentence-repetition tasks (repetition variable) and performance in the auditory and visual variants of the TOJ task (audio-visual variable) and performance, that is not statistically significant, $r(28) = 0.220$, $p= 0.243$, $R^2 = 0.048$ or $rho(28) = 0.110$, $p= 0.562$, $R^2 = 0.012$ (**Figure 10**). Similarly, a positive correlation between performance on the TRA task (tones variable) and the auditory and visual variants of the TOJ task (audio-visual variable) also did not achieve statistical significance, $r(28) = 0.304$, $p= 0.206$, $R^2 = 0.092$ or $rho(28) = 0.224$, $p= 0.486$, $R^2 = 0.050$ (**Figure 11**).

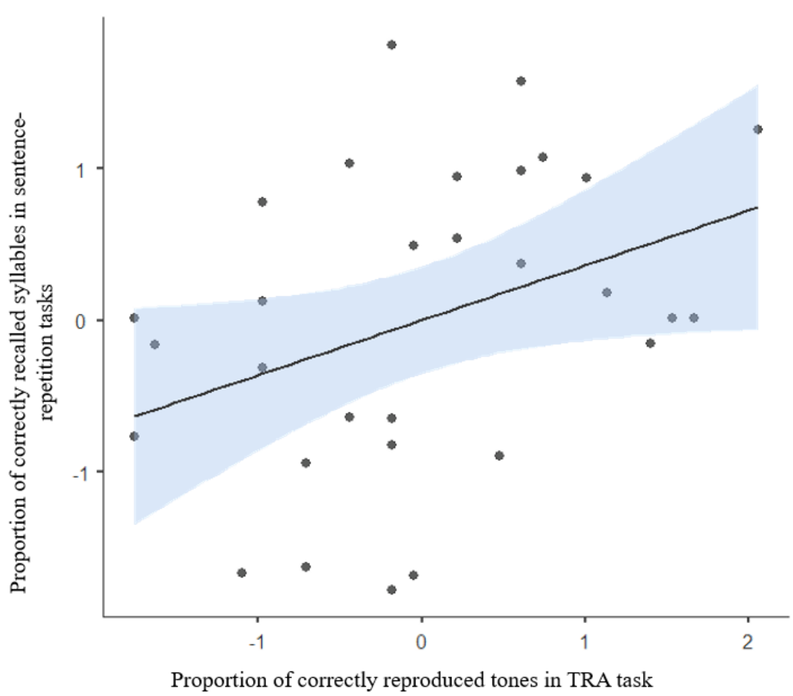

Figure 9: Scatterplot illustrating the positive correlation between performance on a measure of the TRA task (tones variable) and performance on the English pseudoword and Turkish sentences (repetition variable).

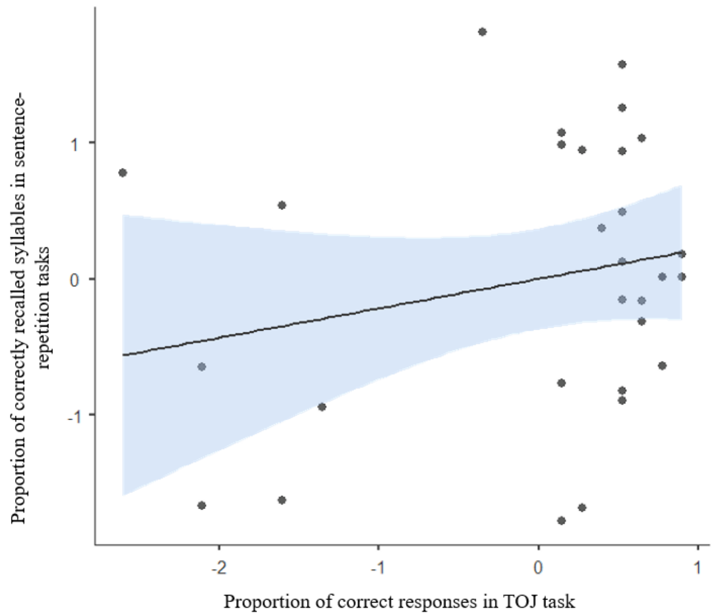

Figure 10: Scatterplot illustrating the weak positive correlation between the audio-visual variable and the repetition variable. The audio-visual variable measures serial nonverbal STM by accounting for the proportion of correct responses in the two variants of the TOJ task. The repetition variable measures phonological serial order by accounting for the proportion of correctly recalled syllables in the English pseudoword and Turkish sentences

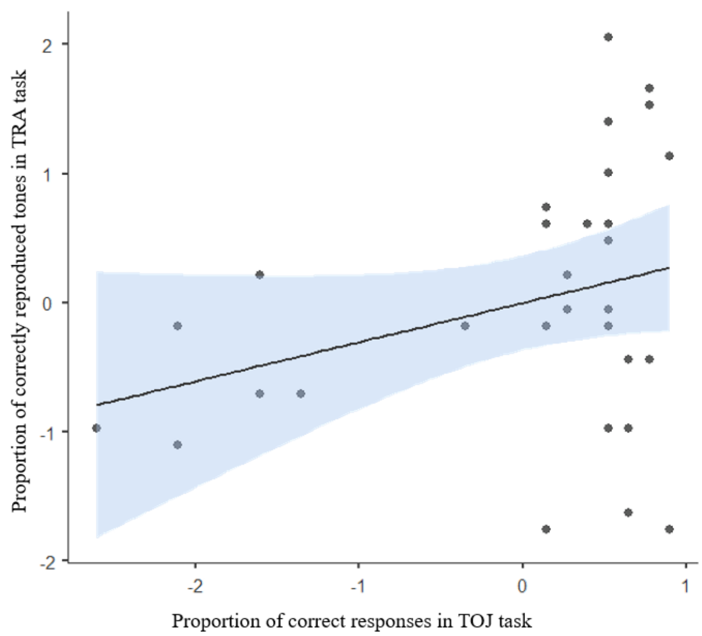

Figure 11: Scatterplot illustrating the correlation between performance on the two variants of the TOJ task (audio-visual variable) and a measure of the TRA task (tones variable).

In a third correlational analysis, we examined whether individuals' score on the Spanish proficiency test, a measure of individuals lexical knowledge in Spanish, predicted their performance on the Spanish pseudoword sentence-repetition task, a measure of their phonological STM abilities. Before comparing the performance on the proficiency test and the sentence-repetition task, we performed a median split on participants' scores on the Spanish proficiency test in order to divide individuals into two groups: lower and higher scorers. Basic descriptive statistics of the variables are shown in **Table 5** and the correlations can be seen in **Table 6** and **Table 7**.

The findings reveal a weak positive correlation between the individuals that had the lower scores on the Spanish proficiency test and their performance on the Spanish pseudoword sentence-repetition sentence task, which did not attain statistical significance, $r(15) = 0.228$, $p = 0.378$, $R^2 = 0.052$ or $rho(15) = 0.244$, $p = 0.345$, $R^2 = 0.060$ (**Figure 12**). The findings also reveal a non-statistically significant and weakly positive correlation between the individuals that had the higher scores on the Spanish proficiency test and their performance on the Spanish pseudoword sentences, $r(11) = 0.164$, $p = 0.592$, $R^2 = 0.027$ or $rho(11) = 0.230$, $p = 0.449$, $R^2 = 0.053$ (**Figure 12**).

Table 5: Descriptive statistics for the Spanish proficiency test and Spanish pseudoword sentence-repetition task for two participant groups based on a median split of the Spanish proficiency test scores. The scores for Spanish pseudoword sentence-repetition are raw scores based on proportion of correctly recalled syllables. Scores for Spanish proficiency test are based on the raw total score on the test.

	Spanish Proficiency scorers' group	**Spanish Proficiency scores**	**Spanish pseudoword sentence scores**
N	Lower scorers	17	17
	Higher scorers	13	13
Mean	Lower scorers	19.1	0.932
	Higher scorers	22.6	0.928
Standard deviation	Lower scorers	1.73	0.038
	Higher scorers	0.768	0.037
Minimum	Lower scorers	15	0.818
	Higher scorers	22	0.870
Maximum	Lower scorers	21	0.987
	Higher scorers	24	0.987
Shapiro-Wilk *p*	Lower scorers	0.038	0.018
	Higher scorers	0.002	0.277

Table 6: Correlations between performance of the lower group of scorers on the Spanish pseudoword sentences and the Spanish proficiency test. Note. * $p < .05$, ** $p < .01$, *** $p < .001$. All scores were standardized and transformed into z scores. All *p*-values in the table are Holm's corrected.

			1	2
1.	**Spanish pseudoword sentence scores**	Pearson's *r*	—	0.228
		p-value	—	0.378
		Spearman's *rho*	—	0.244
		p-value	—	0.345
2.	**Spanish Proficiency Test scores**	Pearson's *r*		—
		p-value		—
		Spearman's *rho*		—
		p-value		—

Table 7: Correlations between performance of higher group of scorers on the Spanish pseudoword sentences and the Spanish proficiency test. Note. * $p < .05$, ** $p < .01$, *** $p < .001$. All scores were standardized and transformed into z scores. All *p*-values in the table are Holm's corrected.

			1	2
1.	**Spanish pseudoword sentence scores**	Pearson's *r*	—	0.164
		p-value	—	0.592
		Spearman's *rho*	—	0.230
		p-value	—	0.449
2.	**Spanish Proficiency Test scores**	Pearson's *r*		—
		p-value		—
		Spearman's *rho*		—
		p-value		—

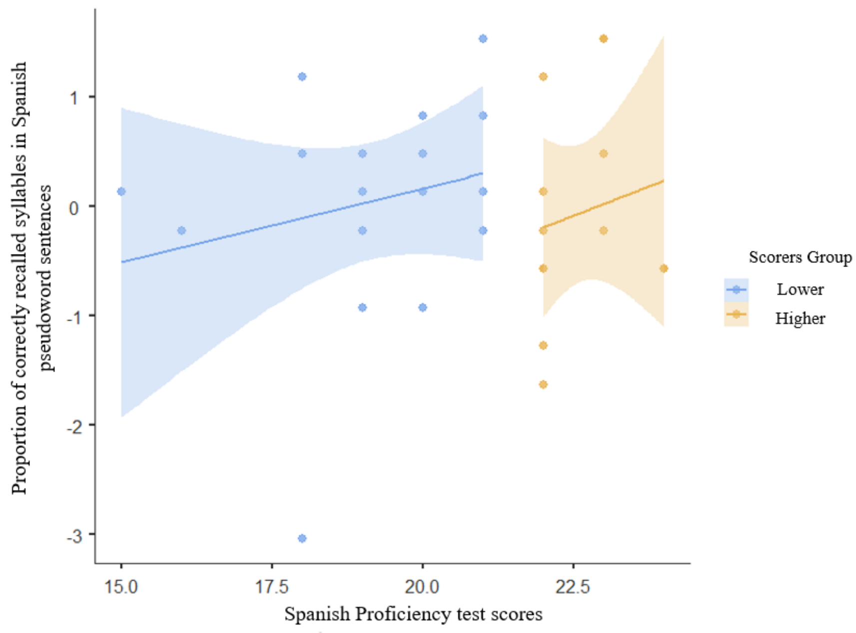

Figure 12: Scatterplots illustrating separate correlations for two participant groups. In blue, we see the correlation between individuals that attained the lower scores in the Spanish Proficiency test and their performance on the Spanish pseudoword sentences task. In orange we see the correlation between individuals that attained the higher scores in the Spanish Proficiency test and their performance on the Spanish pseudoword sentence task.

Multiple Linear Regression Analyses

Two multiple linear regression models were explored for predicting participants' performance in the FWL task from their performance in all or a subset of three measures: the repetition variable (English pseudoword and Turkish sentence-repetition), audio-visual variable (the two variants of the TOJ task) and the tones variable (score of the TRA task). Performance on the FWL task is the outcome variable for this model, based on the proportion of correctly recalled syllables for each word recalled in the task. The repetition variable was measured as the proportion of correctly repeated syllables in the English pseudoword and Turkish sentences, the audio-visual variable as the proportion of correct responses in the two variants of the TOJ task, and the TRA score as the proportion of correctly reproduced tone lengths. Statistics about the regression coefficients is shown in **Table 8**. The initial regression model predicting participants' performance in FWL had the equation FWL score = $-7.99*10^{-7}$ + 0.321 (repetition variable) + 0.066 (audio-visual variable) + 0.160 (TRA score) tones. This three-predictor model was able to account for 19.4% of the variance in performance in the FWL task, ($F(3,25) = 2.00$, $p = 0.139$), $R^2 = 0.194$, 95% CI [-0.363, 0.363]. However, the linear regression equation did not achieve statistical significance. None of the predictor variables had a significant partial Pearson's *r* or Spearman's *rho* correlation with performance in the FWL task (**Figure 13**). The partial correlation between the repetition variable and the FWL task was moderately positive, but it is not statistically significant, $r(27) = 0.404$, $p = 0.178$ $R^2 = 0.163$. There was a weak positive partial correlation between the tones variable and the FWL task, which is not statistically significant, $r(27) = 0.312$, $p = 0.396$ $R^2 = 0.097$. Similarly, the partial correlation between the audio-visual variable and the FWL task was small and positive, and statistically non-significant, $r(27) = 0.199$, $p = 0.396$, $R^2 = 0.040$ or $rho(27) = 0.228$, $p = 0.879$, $R^2 = 0.052$.

Table 8: Three-factor multiple regression model predicting performance in FWL task. N=29. * $p < .05$, ** $p < .01$, *** $p < .001$.

	Coefficient	***t*-value**	***p*-value**
Intercept	-7.99e-7	-4.53e-6	1.000
Repetition variable	0.321	1.613	0.119
Audio-visual variable	0.066	0.353	0.727
Tones variable	0.160	0.807	0.428

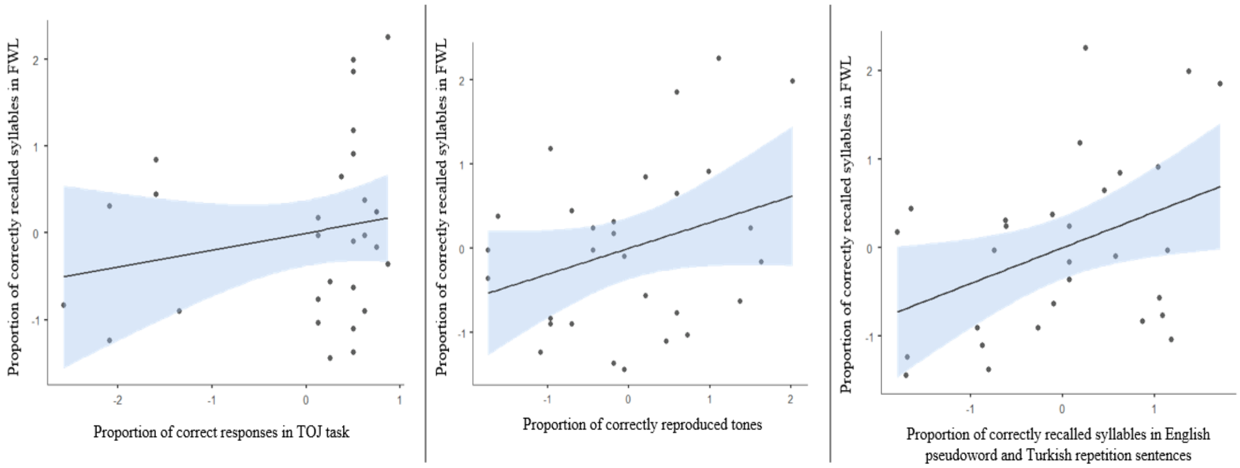

Figure 13: Scatterplots illustrating the three correlations between the FWL task and the TOJ task (left), the TRA task (middle) and the two sentence repetition tasks (right).

A second, simpler, multiple linear regression model for predicting participants' performance in the FWL task from their performance on the repetition variable and the tones variable score of the TRA task was explored next. The audio-visual variable was not normally distributed because of ceiling effects as shown in **Figure 11** and **Figure 13**. As the coefficient for the audio-visual variable did not approach significance, it was dropped out of the model to increase statistical power. Statistics about the regression coefficients are shown in **Table 9**. The regression model predicting participants' performance in FWL had the equation FWL score = $-7.82*10^{-7}$ + 0.332 (repetition variable) + 0.175 (TRA score) tones. This model was able to account for 19.0% of the variance in performance in the FWL task, ($F(2,26) = 3.04$, $p = 0.065$), $R^2 = 0.190$, 95% CI [-0.357, 0.357] but did not reach statistical significance either. In the two-predictor model, the tones variable was not a significant unique predictor of performance on the FWL task. Thus, this model may be a more reliable model for predicting performance on the FWL task in comparison to the previous, three-factor model, but it may still not be an accurate model for predicting performance on the FWL task.

In the previous model, when including the audio-visual variable as a predictor variable, none of the predictor variables had a significant partial Pearson's *r* or Spearman's *rho* correlation with performance in the FWL task (**Figure 13**). For the second model, we determined the partial correlations based on Pearson's *r* rather than Spearman's *rho* since now all the variables passed the Shapiro-Wilk test. The tones variable continued to show a moderate, but statistically non-significant partial correlation with the FWL task, $r(27) = 0.312$, $p = 0.099$, $R^2 = 0.010$. Likewise, a moderate unique positive relationship between the repetition variable and the FWL task did not achieve statistical significance, $r(27) = 0.404$, $p = 0.089$, $R^2 = 0.163$.

Table 9: Two-factor multiple regression model predicting performance in FWL task. N=29. * $p < .05$, ** $p < .01$, *** $p < .001$.

	Coefficient	***t*-value**	***p*-value**
Intercept	-7.823-7	-4.51e-6	1.000
Repetition variable	0.332	1.719	0.098
Tones variable	0.175	0.924	0.364

DISCUSSION

The current thesis investigated how bilinguals performed on a series of tasks that assessed various abilities that have been reliably found to be related with foreign language learning such as processing of serial verbal STM information, serial nonverbal STM information, and general rhythmic STM information (Flaugnacco et al., 2014; Gathercole et al., 1994; Hurlstone et al., 2014; Majerus & Cowan, 2016; Szewczyk et al., 2018). The idea behind studying a bilingual population was to see how these abilities and their proficiencies translated when learning words in a foreign language with a distinct prosody and phonology. In the literature it has been suggested that when bilinguals perform a foreign language learning task such as the FWL task, the task becomes a test of their vocabulary abilities rather than their phonological STM abilities (Majerus et al., 2006; Papagno, Valentine, & Baddeley, 1991; Service & Craik, 1993). Having two languages allows bilinguals to have an extensive word repertoire and so when performing such tasks, a foreign word may remind them of a word or sound in this repertoire (Papagno et al., 1991). We used sentences to assess bilinguals' phonological STM and isolated English word – Turkish word pairs to assess foreign word learning. Sentence repetition requires individuals to keep track of the serial order of words, so it adds an extra level of complexity that still assesses STM processing abilities, even if some words within a sentence may be familiar. Hence, individuals' abilities to keep track of the serial order of events in time, whether of verbal, nonverbal or rhythmic nature, should also give us an indication of their overall language learning abilities.

We first examined whether performance on the three sentence-repetition tasks was related to one another. We predicted that language abilities in a language with familiar prosody would be associated with the learning of phonological sequences in a language of unfamiliar prosody and phonology. We found that both pseudoword sentence-repetition tasks were reliable predictors of performance on the Turkish sentence-repetition task. Based on this, it would seem that there is a relationship between phonological STM of a language of familiar prosody and phonology and phonological STM of a language of an unfamiliar prosody and phonology. Hence, our interpretation is that individuals make use of their vocabulary repertoire to assist them in creating and maintaining phonological representations of the unfamiliar language. Interestingly, there was a weak relationship between performance on the two pseudoword sentence-repetition tasks. This could be explained by the fact that bilinguals are used to making the distinction between the prosody and phonology of Spanish and English and as a result only rely on the respective repertoire for that language when exposed to it.

Another interpretation for these correlations irrespective of individuals' knowledge of their spoken languages being transferred over to knowledge of the unfamiliar language, is individuals' general serial phonological STM abilities. All sentence-repetition tasks, regardless of the prosodic and phonological familiarity required that participants stored within their phonological STM item information (the individual phonemes and words presented) and order information (the order in which the phonemes and words were presented). While both processes may reflect distinct cognitive processes within STM, they are closely connected systems and our sentence-repetition tasks tapped into individuals' abilities to keep track of both types of information (Gupta, 2003; Henson et al. 2003; Majerus et al., 2006). The idea that a direct transfer of knowledge from familiar languages to the unfamiliar language facilitates storage and recall of the Turkish sentences is more related to individuals' processing of item information within verbal STM independent of serial order. However, this second interpretation presents that the correlations amongst the sentence-repetition tasks in our findings are also explained by individuals' processing of serial information within STM. In this way, individuals who were better at storing the order information of the phonemes or words within sentences in the English pseudoword condition, were also better at maintaining and recalling the order information of the phonemes or words within sentences in the Spanish pseudoword and Turkish conditions. Thus, the correlations are also indicative of individuals' capacities for retention and retrieval of serial order information within the verbal STM domain.

We also addressed the question of how serial order STM abilities compared across domains. We compared performance between three variables: verbal repetition, audio-visual event order memory and memory for patterns of short and long tones. We predicted that performance on tasks that measured serial phonological STM abilities would be related to performance on measures of serial nonverbal and rhythmic STM abilities. Overall, we found two positive relationships: between the repetition variable and the tones variable ($r(28) = 0.368$), and between the audio-visual variable and the tones variable ($r(28) = 0.304$ or $rho(28) = 0.224$). However, neither of these relationships attained significance. Thus, performance on the TRA task did not reliably predict performance on the English pseudoword and Turkish sentence-repetition tasks. Likewise, performance on the two variants of the TOJ task also did not reliably predict performance on the English pseudoword and Turkish sentence-repetition tasks. In addition, we found a positive relationship between the tones variable and the audio-visual variable that was also not significant. One plausible explanation for the weak relationships between the repetition variable and the tones variable and with the tones variable and the audio-visual variable, may be that the TOJ tasks were not challenging enough for an adult population, as participants performed at ceiling. The TOJ tasks were originally adapted as measures of serial nonverbal STM for children ages 3-6 (Lahti-Nuuttila et al., submitted). As a result, the TOJ tasks failed to create enough variability in adult performance to reliably detect individual differences as were found in the two sentence-repetition tasks. However, the tones variable also failed to show a reliable relationship with the repetition variable. Thus, it is also plausible that there is no relationship between individuals' serial phonological STM abilities and serial STM abilities in the nonverbal and rhythm domains.

We also investigated whether the level attained based on the scores on the Spanish proficiency test could determine how participants performed on the Spanish pseudoword sentence-repetition task, granted that individuals would vary in their lexical knowledge of Spanish. We predicted that more proficient Spanish speakers would perform better in the Spanish proficiency test and then in turn in the Spanish pseudoword sentence-repetition task because they would have greater support from Spanish prosody and lexical knowledge. Our findings suggest that this may not be the case as we found no significant relationships between the scorer's level on the Spanish proficiency test and performance on the Spanish pseudoword sentence-repetition task. As such, the scorer's level of proficiency attained based on performance on the Spanish proficiency test did not predict performance on the Spanish pseudoword sentence task. It appears that there was not enough variability in individuals' performance on Spanish proficiency test, and, as a result, the Spanish pseudoword sentence task may have been too easy to be an accurate assessment of phonological STM of a language with a familiar prosody and phonology. In fact, bilinguals in our sample all had a very similar language profile, where the majority reported Spanish as their first language and English as their second and dominant language.

We examined whether the same three variables, verbal repetition, audio-visual event order memory and memory for patterns of short and long tones could predict performance on the FWL task. Our reasoning was that assuming that the three variables are measures of individuals' serial STM abilities in the verbal, nonverbal and rhythmic domains, respectively, they should in turn predict individuals' foreign word learning abilities. The findings of a first linear regression model analysis demonstrated that none of the three variables were strong unique predictors of performance on the FWL task. As such, the three variables considered, accounted for 19.4 per cent of the variance in performance on the FWL task. This means that the variables that are measuring individuals' serial STM abilities in the verbal, nonverbal and rhythmic modalities account for a fifth of the variance in how individuals performed on the FWL task. However, these findings have to be interpreted with caution because the model's precision may not be accurate given that none of the variables were found to be reliable predictors.

In a second linear regression model analysis, the results also found that none of the variables were reliable predictors of performance on the FWL task. This model excluded the audio-visual variable as it was the one predictor that was least related to performance on the FWL task, probably because many participants' performance was at ceiling. No significant amount of variance that was accounted for by the previous model was lost when we removed the audio-visual variable, as this second model still accounted for 19 per cent of the variance in performance on the FWL task. Thus, the repetition and tones variables may be the factors that truly account for a fifth of the variance in how individuals performed on the FWL task. This model's precision of the variance accounted for and the reliability of the factors should also be taken with caution as the model failed to reach statistical significance.

Previous work in the literature of verbal and phonological memory has established that a relationship exists between serial verbal STM as measured by the NWR task and vocabulary development in first and second languages (Gupta, 2003; Gupta et al., 2005; Szewczyk et al., 2018). This is given that LTM and individuals' knowledge of their language is thought to lead to better pseudoword repetition because such a task relies on individuals' lexical knowledge and phonological representations of whole word forms in order to represent the nonwords in verbal STM (Roodenrys, Lethbridge, Hinton, Nimmo, & Hulme, 2002; Szewczyk et al., 2018). It has been proposed that individuals who are more proficient in a language should be better at repeating nonwords which remind them of real words in that language (Szewczyk et al., 2018). The NWR task has also been found to tap into individual's abilities to keep track of the serial order of the individual sounds that make up the pseudowords (Gupta, 2003; Schraeyen et al., 2019). Hence, the ability to repeat pseudowords is an important indicator of individuals' verbal and serial abilities within STM. The findings of the present thesis are consistent with this body of literature as we found that performance on the English and Spanish pseudoword sentence-repetition tasks predicted performance in the Turkish sentence repetition task. We propose that this finding suggests that individuals that have a greater repertoire of phonological representations in their two languages displayed better learning of phonological representations of a language of unfamiliar prosody and phonology.

However, our findings did not show that individuals who were more proficient in Spanish performed significantly better when repeating Spanish pseudoword sentences. This inconsistency of our findings with previous work could be attributed to the fact that individuals in our sample did not differ significantly in terms of their proficiency levels in Spanish. Perhaps, if we had compared performance on the Spanish pseudoword sentence-repetition task between learners of Spanish and native Spanish speakers, the results would have shown significant differences in performance given that the lexical knowledge and phonological representations of the native Spanish speakers would be more extensive than that of the Spanish learners. In our sample, all individuals were proficient Spanish speakers and while we categorized them based on how they scored on the Spanish proficiency test, the scores were not significantly different between the lower and higher scorers.

In the field of developmental language disorders, a body of literature has focused on demonstrating how non-linguistic factors and abilities contribute to different deficits and impairments such as dyslexia or SLI (Vugs et al., 2013). This body of literature hinges on the idea that deficits in verbal STM cannot be explained by phonological processing impairments alone, although both item and serial STM in the verbal domain have been found to be impaired in individuals with developmental language disorders (Majerus & Cowan, 2016). Evidence for this view is found in studies that have shown that serial order in the visuospatial domain of STM also appears to be affected in language impairments such as dyslexia and SLI (Majerus & Cowan, 2016; Vugs et al., 2013). Some researchers have suggested that poorly developed serial order STM abilities increase the risk of learning difficulties irrespective of the cognitive domain (Jaroslawska, Gathercole, Logie, & Holmes, 2016; Leclercq & Majerus, 2010; Majerus & Cowan, 2016). For this reason, assessing how individuals perform on serial STM tasks that employ individual's abilities in domains other than verbal is important for understanding exactly how sequential mapping of prosody and phonology is involved in language learning and impairment (Majerus & Cowan, 2016).

In our study, we surprisingly did not find a significant relationship between serial STM abilities in the nonverbal domain and serial STM abilities in the verbal domain. This weak relationship in the present thesis could be attributed to the fact that performance on the two variants of the TOJ task reflected individuals' verbal mediation of auditory and visual information and as such the tasks do not reflect their true abilities of storage and processing of serial auditory and visual information. While we attempted to block any use of verbal coding by having individuals employ an articulatory suppression technique, it is quite possible that this task was not challenging enough for the individuals in our sample. As a result, the articulatory suppression may not have been enough to prevent individuals from using a phonological code for each event while keeping track of the serial order of the visual or auditory events in the task. This is further made evident by our findings showing that individuals performed at ceiling on both the auditory and visual portions of the TOJ task.

One plausible explanation as to why the TOJ task may not have been challenging enough for an adult bilingual population is the fact that bilingualism may confer an advantage to our participants that allows them to perform two tasks simultaneously (the articulatory suppression and the TOJ variants). Evidence for such a bilingual advantage has been found in various studies, where the findings suggest that a bilingual advantage might exist in different cognitive tasks such as inhibition tests, cognitive control, WM and spatial processing (Bialystok et al., 2008; Brito et al., 2016; Grundy & Timmer, 2017; Morales, Calvo, & Bialystok, 2013). However, this view has been heavily contested by other findings. For instance, a recent study by Nichols et al. (2020) failed to show any differences between bilinguals and monolinguals in performance on a series of cognitive tasks. Thus, it appears that bilingualism might not confer cognitive advantages in terms of WM abilities, after all. Consequently, perhaps it is not the case that the task was not challenging enough due to a bilingual advantage. Nonetheless, it is possible that the task did not properly assess individuals’ serial STM abilities in a non-linguistic domain because the task is overall too simple for an adult population, which was not the population for which the task had been intended. We suspect that the adult participants picked up on patterns within the tasks, such as the fact that they only had to pay attention to the two events that would change in situations where the sequence was “different”. In this view, it is no surprise that we failed to find a strong correlation between serial

STM in the nonverbal domains and the tasks that assessed individuals' serial STM abilities in the verbal domain.

While we were unable to find any significant results with the TOJ task in the present thesis, the task itself has proven to be a good measure for teasing apart individual differences in children's serial nonverbal STM abilities (Lahti-Nuuttila et al., submitted). For this reason, we were interested in exploring how an adult population would perform on this task, even if that meant they would perform at ceiling. Given its proven ability to be a measure for serial nonverbal STM abilities, the TOJ task could still be a sensitive measure to individual differences in an adult population upon changing the length of time between the events in a sequence. In turn, this would make the pattern of the items that are incorrect in the "different" sequences unpredictable. Reducing the time that passes between each event in a given sequence should make it difficult for adults to create individual phonological labels for each character or sound. However, individuals may still need to perform the articulatory suppression, to prevent them from creating phonological codes to retain visual information (Kroll et al., 1970). This is given that studies have found that visual information is often easier to maintain via subvocal rehearsal (Baddeley, 2012; Kroll et al., 1970). Another important element to upgrade in the TOJ task would be to have the number of incorrect characters or sounds in a sequence be at random and at varying positions. This is in contrast to the current pattern for the "different" sequences, where once individuals reach four events in a sequence, all the sequences contain only two events that are in the incorrect order and these two events are always in the middle position. As reported by various participants, this pattern in the current version of the TOJ task became very predictable and as a result, individuals may not have employed their nonverbal serial STM abilities in the way we had intended.

We also explored the relationship between individuals' general rhythm abilities and their serial STM abilities across domains. Understanding how general rhythm abilities relate to language processing is important in order to understand the elements that are key for language learning. Findings in the literature of speech and rhythm have shown that a relationship exists between individuals' rhythm skills and language development, where specific rhythm abilities have been found to be related with verbal and nonverbal information processing (Tierney et al., 2017). Thus, rhythm has been reported to help individuals organize events in time and may play a major role in the phonology and prosody of a language. In our findings, we failed to find a relationship between rhythm abilities and serial STM abilities in other domains. This is inconsistent with previous work in the literature of rhythm, where rhythm reproduction has been found to be related with phonological processing (Flaugnacco et al., 2014). In this thesis, the TRA task assessed individuals' rhythmic abilities by having them listen to different sequences of short and long tones and coordinating their motor movements to reproduce the same sequences. The TRA task required motor coordination while grouping events into meaningful chunks (Flaugnacco et al., 2014). One potential reason for why performance in the TRA task did not correlate with performance on the other serial order STM tasks could be attributed to the fact that the TRA task did not create enough variability in individuals' performance as the number of events remained consistent throughout all sequences. This is important to note because in both the sentence-repetition tasks and the TOJ tasks, individuals had to keep track of a varying number of events in the different sequences they completed. This would also explain why our findings differed from those found in previous work, given that such work had individuals reproduce rhythmic sequences that varied in the number of events instead of the length of the events (Flaugnacco et al., 2014). In the case of the TRA task, individuals only had to keep track of the order and length of seven events for each sequence. Thus, it is entirely possible that the task became predictable and individuals employed other strategies to remember the order of the tones and their varying lengths. This view is consistent with the idea that individuals may have created a code to keep track of the order of the long tones and perhaps filled the gaps of the remaining number of tones with short tones, creating a random pattern that would have them performing at chance level. Therefore, the TRA task perhaps tapped into individuals' sensorimotor synchronization abilities, which have not been found to be related with language abilities, and as a result did not tap into individuals' abilities of building a temporal structures of the events as it had been intended (Flaugnacco et al., 2014).

The weak relationship we found between individuals' general rhythm abilities and serial STM abilities could also be explained by theories provided by the event-based models for how serial order information is represented and processed in STM. These models propose that time does not play a role in representing serial order in STM (Gorin, 2020). Instead, serial order is represented in STM through the binding of the items and states of a contextual signal as the items occur in a sequence (Gorin, 2020). Under this view, it is not surprising for individuals' general rhythmic abilities in STM to not predict their abilities to represent serial order in STM. In a study, Gorin (2020) manipulated the timing of the verbal material (digits) that were presented as a sequence. One condition involved regular timing, where the sequences were made-up of verbal material that followed a regular and predictable rhythmic pattern. The second condition involved irregular timing, where the verbal material in the sequence had an irregular and unpredictable rhythmic pattern. The results showed that irrespective of the regularity of the digits presented on the list, no effect was found on individuals' abilities to reconstruct the serial order of the digits presented to them. Our results in a way mirror the findings in Gorin (2020), where it is not necessarily the case that rhythmic abilities in STM are not related to language learning or foreign-word learning, but they may not be related with the serial order aspect of language as we had originally assumed.

As aforementioned, verbal STM reflects the capacity of maintaining phonological representations and as such it is necessary for foreign-word learning (Service & Craik, 1993; Service & Kohonen, 1995). The FWL task in this thesis was developed with the aim of measuring individuals' ability to create and maintain phonological representations of the foreign words in WM (Service & Craik, 1993). We hypothesized that performance on the FWL task would be related to individuals' performance on the STM tasks. We assumed that together the STM tasks created a profile of individuals' abilities to link foreign words to existing lexical items, their phonological support from familiar word forms in their LTM, their abilities to build temporal structures of the events, and their abilities to exploit prosodic and phonemic information to process speech (Becker, Schild, & Friedrich, 2018; Flaugnacco et al., 2014; Service & Craik, 1993). Together, these abilities have been found in several studies to be important for representing unfamiliar phonological information in WM, which is essential for the acquisition of new words in foreign-language learning (Service, 1992; Service & Craik, 1993; Service & Kohonen, 1995). In our study, we did not find any significant correlations between performance on the FWL task and performance on the three STM tasks. However, performance on the English pseudoword and Turkish sentence-repetition tasks appears to be the best predictor of the proportion of correctly recalled syllables in the FWL task. It is plausible that serial STM abilities in the nonverbal and rhythm domains are not related to the ability of creating and maintaining phonological representations of the Turkish words in WM given that they could be distinct processes. However, as mentioned earlier we are not certain that the TOJ tasks and the TRA task captured individuals' true serial STM abilities in the nonverbal and rhythm domains as we had intended. Therefore, we are unable to reject with certainty the null hypothesis that abilities underlying these tasks are truly unrelated with foreign-word learning.

The interpretations of the findings in this thesis should be taken with caution as the tasks that we developed do not appear to measure with certainty what we had meant them to measure. The current thesis was intended as exploratory research as we had aimed to see how these various mechanisms might be linked when individuals are asked to learn new words in a foreign language. Future studies should include a monolingual group to compare their performance with the bilingual group and help us see if our suggestions in regard to lexical knowledge are correct and whether these tasks actually measure what we want them to measure. Also, a monolingual sample might help us see with certainty if there are truly no correlations between the four behavioural tasks we developed. Similarly, future work should consider testing a different population of bilinguals, such as French-English bilinguals, where French is their first and dominant language and English is their second language. In this way, one could examine how these bilinguals perform on tasks that use a language with familiar prosody and phonology (English pseudoword), a language with prosody and phonology that is unfamiliar but similar to one of their languages (Spanish pseudoword) and a language with very unfamiliar prosody and phonology (Turkish). Perhaps, this would give us more information about the relationship between the ability to create and keep phonological representations of a foreign language and how varying degrees of lexical knowledge in different languages might assist in this process. In addition, future versions of the TOJ task and the TRA task should avoid having predictable patterns for individuals which would make the tasks less about recognition of the serial events and more about maintaining the serial order and chunking the events into meaningful units. For instance, a future version of the TOJ task could have more than two events being incongruent or having the incongruent events in positions other than the middle. Similarly, a future version of the TRA task could involve having individuals reproduce tone sequences that not only have two events that vary in length but have the sequences of tones vary in the number of events as well. Finally, future work should see how the suggested changes for the three other behavioural tasks affect the relationship and their predictive ability over foreign-word learning.

Conclusion

This thesis investigated different cognitive mechanisms assumed to be involved in language processing as well as the effects on foreign word learning. We used a series of behavioural tasks to measure adult bilinguals' serial order STM abilities in the verbal, nonverbal and rhythm domains, and abilities to learn new words in a foreign language with an unfamiliar prosody and phonology. In general, maintenance and learning of phonological sequences that followed an unfamiliar prosody and phonology was predicted by an individual's ability to link unfamiliar sounds to existing lexical items in their spoken languages as indexed by the sentence repetition tasks. This can be best explained by the fact that individuals must be obtaining phonological support from LTM in their spoken languages. However, the findings also suggest that lexical knowledge in their first language alone may not have been as indicative of individuals' abilities to link new words to existing lexical items, as the participants in our sample were not likely to vary much in their lexical knowledge.

Another surprising finding was that serial order STM abilities in the verbal domain was not significantly related to serial order STM abilities in other domains, a finding that is inconsistent with other work in the literature. However, it is quite possible that the nonverbal and rhythmic serial order tasks used in this project failed to tap into the serial ordering mechanism of STM as the tasks were not challenging enough for the participants in our sample to make use of such mechanisms. We also found that individual's serial order STM abilities in the nonverbal and rhythm domain did not effectively predict their ability to create and maintain phonological representations of foreign words.

However, phonological STM abilities as measured in the sentence-repetition task may be the most reliable predictor of individuals' abilities to form phonological representations of foreign words. As such, the relationship between the nonverbal or rhythm STM behavioral tasks that aimed to tap into the various cognitive mechanisms in language processing and foreign language learning remains unclear. Further development of the nonverbal serial and rhythm STM tasks will be needed for more conclusive evidence either for or against a relationship between the domain general forms of STM and foreign word learning aptitude.

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 89–195.

Attout, L., Grégoire, C., & Majerus, S. (2020). How robust is the link between working memory for serial order and lexical skills in children? *Cognitive Development*, 53, 1-11.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation: Advances in Research and Theory*, *8*, 47–89.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), 136–140.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158–173.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*, *36A*, 233–252.

Becker, A., Schild, U., & Friedrich, C. K. (2018). Tracking independence and merging of prosodic and phonemic processing across infancy. *Developmental Science*, *21*(2), 1–11.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*(4),

Bion, R. A. H., Benavides-Varela, S., & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, *54*(1), 123-140.

Bogaerts, L., Szmalec, A., Hachmann, W. M., Page, M. P. A., Woumans, E., & Duyck, W. (2015). Increased susceptibility to proactive interference in adults with dyslexia? *Memory*, *23*(2), 268-277.

Bolton, T. L. (1984). Rhythm. *The American Journal of Psychology*, *6*(2), 145–238.

Brito, N. H., Murphy, E. R., Vaidya, C., & Barr, R. (2016). Do bilingual advantages in attentional control influence memory encoding during a divided attention task? *Bilingualism (Cambridge, England)*, *19*(3), 621-629.

Burgess, N., Hitch, G. J., Henson, R., Shallice, T., Houghton, G., Vallar, G., … Yeo, C. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551-581.

Chartrand, J.-P., & Belin, P. (2006). Superior voice timbre processing in musicians. *Neuroscience Letters*, *405*, 164–167.

Chobert, J., François, C., Velay, J.L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956-967.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, *55*(1), 75–84.

Depoorter, A., & Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *The Quarterly Journal of Experimental Psychology*, *62*(3), 531–549.

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience, 8*(311), 1–7.

Farrell, S. (2008). Multiple roles for time in short-term memory: Evidence from serial recall of order and timing. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*, 128-145. https://doi.org/10.1037/0278-7393.34.1.128

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*, 115–135.

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., … Schon, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*(392), 1-14.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language*, *24*, 200–209.

Gathercole, S. E. (2006). Nonword repetition and nonword learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513–543.

Gathercole, S. E., & Baddeley, A. D. (1995). Short-term memory may yet be deficient in children with language impairments: A comment on van der Lely & Howard (1993). *Journal of Speech, Language, and Hearing Research*, *38*(2), 463-466.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*(4), 439-454.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological memory and new word learning in children. *Developmental Psychology,* 33, 966-979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*(1), 65-77.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103–127.

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

Gorin, S. (2020). The influence of rhythm on short-term memory for serial order. *Quarterly Journal of Experimental Psychology*, 1–22.

Goswami, U. (2016). Educational neuroscience: Neural structure-mapping and the promise of oscillations. *Current Opinion in Behavioral Sciences*, *10*, 89–96.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*(3), 325-340.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology*, *56A*(7), 1213-1236.

Gupta, P., Lipinski, J., Abbs, B., & Lin, P.-H. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language*, *53*, 141–162.

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

Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology*, *4*(566), 1–16.

Henny Yeung, H., Bhatara, A., & Nazzi, T. (2018). Learning a phonological contrast modulates the auditory grouping of rhythm. *Cognitive Science*, *42*(6), 2000–2020.

Henson, R., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *The Quarterly Journal of Experimental Psychology*, *56A*(8), 1307-1334.

Hick, R., Botting, N., & Conti-Ramsden, G. (2005). Cognitive abilities in children with specific language impairment: Consideration of visuo-spatial skills. *International Journal of Language & Communication Disorders*, *40*, 137-149.

Hoffman, L. M., & Gillam, R. B. (2004). Verbal and spatial information processing constraints in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *47*, 114-125.

Hurlstone, M. J., & Hitch, G. J. (2015). How is the serial order of a spatial sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 295-324.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140*(2), 339-373.

Jalbert, A., Saint-Aubin, J., & Tremblay, S. (2008). Visual similarity in short-term recall for where and when. *The Quarterly Journal of Experimental Psychology*, *61*(3), 353–360.

Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., & Holmes, J. (2016). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, *44*, 580-589.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 1008-1018.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language*, *54*(2), 265-281.

Kelly, D. H., & Burbeck, C. A. (1984). Critical problems in spatial vision. *Critical Reviews in Biomedical Engineering*, *10*(2), 125–177.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*, 599–605.

Kroll N.E., Parks T., Parkinson S.R., Bieber S.L., Johnson A.L. (1970). Short-term memory while shadowing: recall of visually and aurally presented letters. *Journal of Experimental Psychology*. *85*, 220–24.

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

Lahti-Nuuttila, P., Service, E., Kunnari, S., Arkkila, E., & Laasonen, M. (in preparation). Nonverbal serial short-term memory moderates vocabulary growth in children with developmental language disorder.

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience and Biobehavioral Reviews*, *81*, 158–166.

Leclercq, A.-L., & Majerus, S. (2010). Serial-order short-term memory predicts vocabulary development: Evidence from a longitudinal study. *Developmental Psychology*, *46*(2), 417–427.

Logie, R. H., & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of Cognitive Psychology*, *9*(3), 241-257.

Mai, G., Minett, J. W., & S-Y Wang, W. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, 516–528.

Majerus, S. (2019). Verbal working memory and the phonological buffer: The question of serial order. *Cortex*, *112*, 122–133.

Majerus, S., & Boukebza, C. (2013). Short-term memory for serial order supports vocabulary development: new evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology*, 116(4), 811-828.

Majerus, S., & Cowan, N. (2016). The nature of verbal short-term impairment in dyslexia: The importance of serial order. *Frontiers in Psychology*, *7*(1522), 1–8.

Majerus, S., Poncelet, M., Greffe, C., & Van Der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology*, *93*, 95-119.

Martinez Perez, T., Majerus, S., Poncelet, M., & Perez, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, *111*, 708-723.

Martinez Perez, T., Poncelet, M., Salmon, E., & Majerus, S. (2015). Functional alterations in order short-term memory networks in adults with dyslexia. *Developmental Neuropsychology*, *40*(7–8), 407-429.

Miller, G., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior*. New York: Henry Holt and Co.

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202.

Moreno, S., Marques, C., Santos, A., Santos, M., Luis, S., Castro, L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*, 712-723.

Munson, B., Swenson, C. L., & Manthei, S. C. (2005). Lexical and phonological organization in children: Evidence from repetition tasks. *Journal of Speech, Language, and Hearing Research*, *48*, 108-124.

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*(70), 1-17.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*, 233–243.

Nespor, M., Shukla, M., Van De Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realizations in VO and OV languages. *Lingue e Linguaggio*, *7*(2), 139-168.

Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science*, *31*(5), 1-20.

Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System*, *42*(1), 244–257.

Overy, K., Nicolson, R. I., Fawcett, A. J., & Clarke, E. F. (2003). Dyslexia and music: Measuring musical timing skills. *Dyslexia*, *9*, 18–36.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *The Quarterly Journal of Experimental Psychology*, *444*(I), 47-67.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(1), 331–347.

Parmentier, F. B. R., & Andrés, P. (2006). The impact of path crossing on visuo-spatial serial memory: Encoding or rehearsal effect? *The Quarterly Journal of Experimental Psychology*, *59*(11), 1867-1874.

Patel, A. D. (2005). The Relationship of Music to the Melody of Speech and to Syntactic Processing Disorders in Aphasia. *Annals of the New York Academy of Sciences*, *1060*(1), 59–70.

Patel, A. D., Iversen, J. R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology*, *22*(7–8), 776–789.

Peña, M., Bion, R. A. H., & Nespor, M. (2011). How modality specific is the iambic-trochaic law? Evidence from vision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1199.

Petkov, C. I., O'connor, K. N., Benmoshe, G., Baynes, K., & Sutter, M. L. (2005). Auditory perceptual grouping and attention in dyslexia. *Cognitive Brain Research*, *24*, 343–354.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology; Human Perception and Performance*, *16*(3), 564–573.

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

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America*, *105*(1), 512–521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology*, *61*(1), 129–141.

Rispens, J., & Parigger, E. (2010). Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. *British Journal of Developmental Psychology*, *28*(1), 177-188.

Romani, C., Tsouknida, E., & Olson, A. (2015). Encoding order and developmental dyslexia: A family of skills predicting different orthographic components. *The Quarterly Journal of Experimental Psychology*, *68*(1), 99–128.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349.

Schraeyen, K., Van Der Elst, W., Geudens, A., Ghesquière, P., & Sandra, D. (2019). Short-term memory problems for phonemes' serial order in adults with dyslexia: Evidence from a different analysis of the Nonword repetition task. *Cambridge University Press*, *40*(3), 613–644.

Service, E. (1992). Phonology, Working Memory, and Foreign-language Learning. *The Quarterly Journal of Experimental Psychology*, *45*(1), 21–50.

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

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, *16*(2), 155–172.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, *33*, 5–42.

Smyth, M. M., & Scholey, K. A. (1996). Serial order in spatial immediate memory. *The Quarterly Journal of Experimental Psychology*, *49A*(1), 159–177.

Soemer, A., & Saito, S. (2016). Domain-specific processing in short-term serial order memory. *Journal of Memory and Language, 88,* 1–17.

Spillmann, L. (2006). From perceptive fields to Gestalt. *Progress in Brain Research*, *155*, 67–92.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236.

Szewczyk, J. M., Marecka, M., Chiat, S., & Wodniecka, Z. (2018). Nonword repetition depends on the frequency of sublexical representations at different grain sizes: Evidence from a multi-factorial analysis. *Cognition*, *179*, 23-36.

Szmalec, A., Loncke, M., Page, M. P. A., Duyck, W., & Dunantlaan, H. (2011). Order or disorder? Impaired Hebb learning in dyslexia. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37*, 1270-1279.

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*, 120–129.

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.

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour*, *23*, 151-161.

Vugs, B., Cuperus, J., Hendriks, M., & Verhoeven, L. (2013). Visuospatial working memory in specific language impairment: A meta-analysis. *Research in Developmental Disabilities*, *34*, 2586–2597.

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.

Wolff, P. H. (2002). Timing precision and rhythm in developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal*, *15*, 179–206.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, *112*, 201–216.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29.

APPENDIX A

Table i: List of sentences used as stimuli for the English sentence repetition task. Sentences varied from five to six pseudowords and varied from six to eight syllables. All function words were in English.

English Jabberwocky Sentences					
THE	MUTHTIN	FANED	A	GRASHPIT.	
A	DREG	PRIBED	IN	THE	BLANTIS.
THESE	SMORKETS	NOOLED	A	GUZDIN.	
THE	FLOGIN	JANED	TO	A	KINTO.
SOME	FLOTTERS	PEFT	THEIR	LAUMSES.	
A	PONDLE	BOYED	TO	THE	TROMPERS.
THE	NOLK	WINTED	THE	PRILES.	
SOME	STORDINS	RIMPLE	WITH	MIPTERS.	
THESE	GILPERNS	KARNAYED	THE	WUPS.	
A	JOWLER	FRINED	WITH	A	VULKIT.

Table ii: List of sentences used as stimuli for the Turkish sentence repetition task. Sentences varied from three to four words and varied from six to eight syllables.

Turkish Sentences			
SANA	CICEK	ALDIM.	
SENDE	BU	TOPLA	OYNA.
KOPEK	EVE	KACTI.	
UGUR	BAHCEYE	ATLADI.	
KEDIYI	GUZEL	YEDIR.	
O	ATI	BANA	AL.
CAN	KIZA	YEMEK	VERDI.
BU	KELEBEK	COK	GUZEL.
SIZ	ARABAYA	BININ.	
HEPSI	ET	YEDILER.	

Table iii: List of sentences used as stimuli for the Spanish sentence repetition task. Sentences varied from three to five pseudowords and varied from seven to eight syllables. All function words were in Spanish.

Spanish Jabberwocky Sentences				
LAS	NEBAS	RADAN	UN	ZARLO.
UN	CHALO	CON	ESTAGO.	
LAS	PIQUIÑORAS	UNISAN.		
ALGUNOS	TURREN	LA	FUA.	
EL	RUME	DE	UN	FIDETE.
ALGUNAS	SECUNAS	GUEPEN.		
LOS	CILEBOS	DE	TOQUI.	
LA	GAPATA	ROTULLOSA.		
LOS	CHAPUJOS	SENDARON.		
UNA	BESICA	SIN	PAMUL.	

Table iv: List of English-Turkish word pairs that were used as stimuli for the foreign-word learning task. The Turkish words consisted of three syllables and the location of the stressed syllable (bolded for each word) varied between the second and third syllable of the Turkish word.

LIST A	**LIST B**
HAND - LO**KAN**TA	CRAB - DONDUR**MA**
GIRL - E**NIŞ**TE	DOOR - KALEM**LIK**
PLUM - SEPET**LER**	FEET - SAL**YAN**GOZ
NOSE - KA**RA**YEL	ROCK - MAYDA**NOZ**
SCAR - TA**BAN**CA	SILK - A**NAH**TAR
BOAT - ARKA**DAŞ**	POND - ÇER**KIR**GE

APPENDIX B

In the southeastern Pacific Ocean, on the piece of land known as Easter Island (now a territory of Chile), stand several hundred massive stone monoliths. These carvings, called "moai," are recognizable by their oversized heads, with their heavy brows, long noses, elongated ears, and protruding lips. While they average four meters in height and 12.5 tonnes, the largest is almost 10 meters tall and the heaviest weighs a full 86 tons. The upright sculptures are scattered around Easter Island, many installed on platforms called "ahu" along the coast, while others are more inland and several stand near the main volcanic quarry of Rano Raraku. The Rapa Nui people of the island built a total of 887 of these impressive statues between the 12th and 16th centuries. They were, it is said, symbols of religious and political authority, embodiments of powerful chiefs or ancestors which faced inland toward the island's villages, perhaps watching over their creators, keeping them safe. While the very creation of such monoliths – most out of volcanic ash with stone hand chisels – is an impressive feat, what is more remarkable (not to mention mysterious) is how they were transported to their resting places. In the past, most researchers associated the building and transportation of the moai with widespread deforestation on the island and eventual collapse of the Rapa Nui civilization. This hypothesis is based, in part, on the fact that the pollen record suddenly disappears at the same time as the Rapa Nui people stopped constructing the moai and transporting them with the help of wooden logs. How exactly would logs facilitate the movement of the statues? Most proponents of this method believe that the people created "rollers" by arranging parallel logs on which the prone statues were pulled, or pushed. They would not have required an entire roadway of logs, since logs from the back could be placed at the front, creating a moving platform of sorts. To make it easier to roll, and keep in position, the statue would be placed on two logs arranged in a V shape. One proponent of this idea of rolling the statues in a prone position is Jo Anne Van Tilburg, of UCLA. Van Tilburg created sophisticated computer models that took into account available materials, routes, rock, and manpower, even factoring in how much the workers would have to have eaten. Her models supported the idea that rolling prone statues was the most efficient method. As further evidence, Van Tilburg oversaw the movement of a moai replica by the method she had proposed. They were successful, but evidence that it was possible is not necessarily evidence that it actually happened.

Figure i: Illustrates the English text that was used to determine participants' reading comprehension ability.

Recientes investigaciones del Instituto de Ciencias del Mar y Limnología de la UNAM han demostrado que los efectos del cambio climático producen la pérdida de los arrecifes de coral, informó Roberto Iglesias Prieto, académico de esta entidad. El incremento de tan sólo un grado centígrado en la temperatura de los océanos ocasiona el blanqueamiento masivo de las colonias de coral y les causa la muerte, lo que significa la pérdida de miles de especies, debido a que en estos ecosistemas se encuentran las más variadas del planeta. Además estos arrecifes son la más importante fuente turística de México, pues producen la arena blanca del mar Caribe, preferida por el turismo internacional, que representa la tercera fuente de divisas, solo por debajo de la industria petrolera y de las remesas provenientes de Estados Unidos. El investigador menciono que a pesar de que solo ocupan el dos por ciento de la superficie terrestre, estos ecosistemas son una gran fuente de riqueza económica y de servicios ambientales para más de cien millones de personas en el mundo y capturan aproximadamente 50 por ciento del carbono planetario. Asimismo, informó son las estructuras geológicas de origen biológico más grandes del orbe y han dominado las zonas poco profundas de los mares tropicales en los últimos 200 millones de años. Por eso, argumentó, sería necesario que se generara conciencia entre la población para el cuidado de estas fuentes de riqueza, ya que tan perjudiciales son el calentamiento global y el efecto invernadero como la contaminación marítima producida por la población. Como ya se dijo, los arrecifes resultan del crecimiento de las colonias de corales que depositan sus esqueletos en la roca calcárea, y a través de miles de años forman una estructura definida que se incrementa mientras no haya erosión. Si bien este último fenómeno explica las playas blancas como las de Cancún, especificó el investigador, en caso de que la erosión sea mayor a la incorporación de carbonato de calcio, se seguirá produciendo la arena blanca pero el arrecife no podrá prosperar. Las acciones para el cuidado de los arrecifes son escasas e insuficientes, algo no privativo del país, señaló el especialista, pues en todo el mundo persiste el desconocimiento sobre el estado de ese ecosistema o se minimiza su importancia, inclusive en países que tienen políticas de protección del medio ambiente.

Figure ii: Illustrates Spanish text that was used to determine participants' reading comprehension ability.

So, we've been discussing sixteenth-century Indigenous life, and today we're going to focus on Iroquois and Huron peoples. They lived in the northeastern Great Lakes region of North America. Now, back then, their lives depended on the natural resources of the forest, especially the birch tree. The birch tree can grow in many different types of soils and is prevalent in that area. The birch tree has white bark. And this tough protective outer layer of the tree, this white bark, is waterproof, and this waterproof quality of the bark, it made it useful for making things like cooking containers, a variety of utensils. If you peel birch bark in the winter, we call it the "winter bark". Another layer, a tougher inner layer of the tree adheres to the bark, producing a stronger material … so the "winter bark" was used for larger utensils and containers. One of the great things about birch bark is that the taste of the birch tree doesn't get transferred to the food—so it was perfect for cooking containers, but the most use of the bark was the canoe. Since the northeast region of North America is interconnected by many streams and waterways, water transportation by vessels like a canoe was most essential. The paths through the woods were often overgrown, so water travel was much faster. And here's what the Indigenous people did … they would peel large sheets of bark from the tree to form lightweight yet sturdy canoes. The bark was stretched over frames made from tree branches, stitched together and sealed with resin—you know, that sticky liquid that comes out of the tree—and when it dries, it's watertight. One great thing about these birch bark canoes was, that they could carry a large amount of cargo. For example, a canoe weighing about 50 pounds could carry up to 9 people and 250 pounds of cargo. The Indigenous people made canoes of all types, for travel on small streams or on large open ocean waters. For small streams they made narrow, maneuverable boats, while larger canoes were needed for the ocean. They could travel throughout the area, only occasionally having to portage to carry the canoe over land a short distance to another nearby stream. And since the canoes were so light … this wasn't a difficult task. Having an efficient means of transportation helped the Iroquois to form a federation, linked by natural waterways, and this federation expanded from what is now southern Canada all the way south to the Delaware River. This efficiency of the birch bark canoe also made an impression on newcomers to the area. French traders in the seventeenth century adopted the design of the Iroquois birch bark canoes and they found that they could travel great distances—more than 1500 kilometers a month.

Figure iii: Illustrates the transcript of the English audio that was recorded and used to determine participants' listening comprehension ability.

En filosofía, la palabra emoción no aparece mucho; aparecen otros conceptos como sentimientos, afectos, pasiones... Por ello, no voy a distinguir entre una cosa u otra, es decir, yo me referiré indistintamente a las pasiones, a los sentimientos, a los afectos, a las emociones porque creo que no es importante matizar para lo que voy a decir aquí. Y también quiero señalar de entrada que, por lo general, la ética ha sido entendida sobre todo como la expresión de la actividad racional. Nos ha llegado, hemos heredado una ética que llamamos ética de principios, que es una ética basada en deberes, en normas, en obligaciones, en derechos incluso, ¿no? Es una ética muy racionalista. La ética concebida como que es bueno aquello que produce mayor felicidad en el mundo. Para llegar a esa conclusión, también se necesita elaborar un cálculo que es, sobre todo, racional. Se ha olvidado un poco, ¿eh?, un olvido que se está corrigiendo en los últimos años, la ética de las virtudes, que es la que está más relacionada el tema de las emociones. Y es absolutamente fundamental, ¿no?, no como sustituto a las otras éticas, sino como complemento. Hay algunos filósofos que desde el principio pensaron que no era posible construir una ética sin tener en cuenta los sentimientos. Uno de ellos es Aristóteles. Aristóteles construye una ética de las virtudes. Piensa que la ética debe ser una tarea práctica. Él, sobre todo, en la Ética a Nicómaco hace esta reflexión; dice: estoy hablando de lo que es la virtud y lo que nos importa es hacer personas virtuosas. Eso es lo que debería importarnos, ese paso de la teoría a la práctica. Pero es que además Aristóteles considera que las virtudes morales se asientan en lo que él llama el alma sensitiva de la persona. Por lo tanto, están muy ligadas a los sentimientos, a los afectos, porque de esas virtudes depende nuestra manera de ser. Y eso forma…, acaba formando parte de nuestra manera de ser y, por lo tanto, está muy ligado a nuestra forma de sentir, también. Otro filósofo que no concibe la ética separada de los afectos es Spinoza. Él los llama así: “afectos” y entiende que hay dos grandes tipos de afectos: la alegría y la tristeza. Todos los afectos se pueden resumir en estos dos: o son alegres o son tristes, y lo que hay que conseguir, es decir, el imperativo ético para Spinoza, es potenciar al máximo los afectos alegres y evitar los tristes, porque eso será vivir libremente. Y, finalmente, un tercer filósofo que tiene que ser destacado también por la valoración que hace de los sentimientos es Hume. De Hume es muy conocida su frase de La Razón, donde dice que la razón es esclava de las pasiones y que no es la razón lo que mueve al mundo sino la pasión. Son los sentimientos los que mueven el mundo.

Figure iv: Illustrates the transcript of the Spanish audio that was recorded and used to determine participants' listening comprehension ability.

Selecciona la respuesta más apropiada.

1. ¿Qué __________ ustedes?
   a) hacéis
   b) hago
   c) hacen
   d) haces

2. Hace un año que trabajo en __________ fábrica.
   a) esto
   b) esta
   c) ese
   d) este

3. Carlos dijo que __________ al correo pero no tuvo tiempo.
   a) va
   b) iría
   c) iban
   d) ir

4. ¿Dónde vivían los aztecas a __________ venció Cortés?
   a) quienes
   b) que
   c) las cuales
   d) quien

5. Todos mis amigos __________ a la fiesta de Marcos.
   a) han sido invitados
   b) han sido invitado
   c) están invitado
   d) esta invitados

6. Todos mis amigos __________ a la fiesta de Marcos.
   a) han sido invitados
   b) han sido invitado
   c) están invitado
   d) esta invitados

Figure v: Illustrates the questions in the first part of the grammar section that were used to determine participants' written grammar abilities.

Selecciona la palabra subrayada que esté incorrecta
1. Ayer yo el escribí un correo electrónico. a) Ayer b) el c) un d) correo 2. Nosotros dirigimos uno negocio importante. a) Nosotros b) dirigimos c) uno d) importante 3. El maestro era un hombre dedicaba a los niños. a) era b) hombre c) dedicaba d) a 4. Pagó diez dólares para el libro. a) Pagó b) para c) el d) libro 5. Si tenía mucho dinero, me compraría un coche nuevo. a) tenía b) me c) compraría d) un

Figure vi: Illustrates the questions in the second part of the grammar section that were used to determine participants' written grammar abilities.

Figure vii: Illustrates the picture that was shown to participants to determine their English oral expression.

Figure viii: Illustrates the picture that was shown to participants to determine their English oral expression.

Level 6: 18-21 points Level 5: 14-17 Level 4: 9-13 points
Level 3: 6-8 points Level 2: 3-5 points Level: 0-2 points

SPANISH ORAL EXPRESSION		
Fluency	- Suitable speed and pauses - How well does the participant describe the picture?	**0 1 2 3**
Communicative Ability	- Speech contained long utterances, flexible use of language and complex responses	**0 1 2 3**
Accuracy	- How accurate and appropriate was the participant's grammar? - How are the verbs, sentence structure and individual grammar points?	**0 1 2 3 4 5**
Vocabulary	- Does the participant use a wide variety of words and phrases? - The choices of words, idioms and phrasal verbs are appropriate each time	**0 1 2 3 4**
Pronunciation	- Effort is made to use correct intonation, stress and individual sounds	**0 1 2 3 4**
Content	- Topic elaboration, organization, coherence and suitable linkers and connects are used	**0 1 2**
		TOTAL:

Level 6: 18-21 points Level 5: 14-17 Level 4: 9-13 points
Level 3: 6-8 points Level 2: 3-5 points Level: 0-2 points

ENGLISH ORAL EXPRESSION		
Fluency	- Suitable speed and pauses - How well does the participant describe the picture?	**0 1 2 3**
Communicative Ability	- Speech contained long utterances, flexible use of language and complex responses	**0 1 2 3**
Accuracy	- How accurate and appropriate was the participant's grammar? - How are the verbs, sentence structure and individual grammar points?	**0 1 2 3 4 5**
Vocabulary	- Does the participant use a wide variety of words and phrases? - The choices of words, idioms and phrasal verbs are appropriate each time	**0 1 2 3 4**
Pronunciation	- Effort is made to use correct intonation, stress and individual sounds	**0 1 2 3 4**
Content	- Topic elaboration, organization, coherence and suitable linkers and connects are used	**0 1 2**
		TOTAL:

Figure ix: Illustrates the guidelines that were used to determine the oral expression for participants in both English (bottom) and Spanish (top).

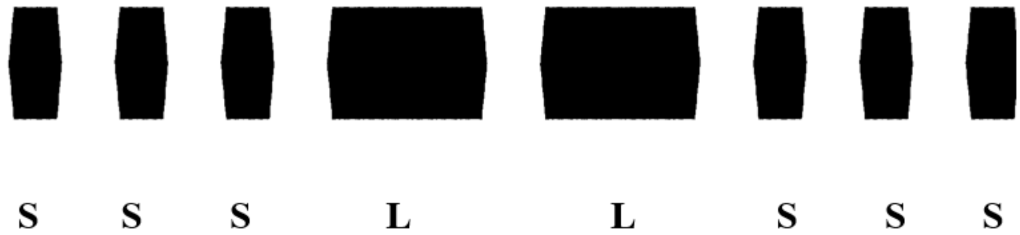

Figure x: Illustrates an example of tone sequence of 7-alternating short and long tones used as stimuli for the Tones task.